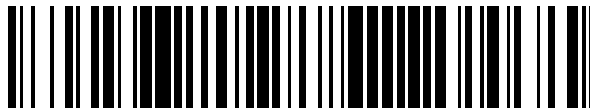


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 627 117**

51 Int. Cl.:

**C12N 9/10** (2006.01)

**C12N 15/62** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.12.2011 PCT/EP2011/072387**

87 Fecha y número de publicación internacional: **14.06.2012 WO2012076715**

96 Fecha de presentación y número de la solicitud europea: **09.12.2011 E 11791602 (3)**

97 Fecha y número de publicación de la concesión europea: **15.03.2017 EP 2649178**

54 Título: **Método basado en MGMT para obtener un rendimiento elevado de expresión de proteínas recombinantes**

30 Prioridad:

**09.12.2010 EP 10306389**  
**08.07.2011 US 201161505694 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**26.07.2017**

73 Titular/es:

**INSTITUT PASTEUR (100.0%)**  
**25-28, rue du Docteur Roux**  
**75015 Paris, FR**

72 Inventor/es:

**DESPRES, PHILIPPE;**  
**PAULOUS, SYLVIE y**  
**CRUBLET, ELODIE**

74 Agente/Representante:

**CURELL AGUILÁ, Mireia**

**Observaciones:**

**Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes**

ES 2 627 117 T3

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

## DESCRIPCIÓN

Método basado en MGMT para obtener un rendimiento elevado de expresión de proteínas recombinantes.

5 La presente invención se refiere al campo de la ingeniería genética y la biología molecular. En particular, la presente invención se refiere a un nuevo aumentador de la producción de proteínas en las células hospedadoras. Además, la presente invención se refiere a vectores que contienen la secuencia de ADN codificante de dicha proteína aumentadora y también a la utilización de la misma para la expresión de proteínas recombinantes, tales como enzimas industriales o proteínas para la utilización farmacéutica, incluyendo proteínas eucarióticas (por ejemplo de mamífero, tal como humanas) y proteínas víricas.

## Antecedentes de la invención

15 Los sistemas de producción de proteínas, en los que se producen polipéptidos o proteínas de interés en organismos o células recombinantes, son la columna vertebral de la biotecnología comercial.

A los primeros sistemas, basados en la expresión bacteriana en huéspedes tales como *E. coli*, les han seguido sistemas basados en huéspedes eucarióticos, en particular células de mamífero en cultivo, células de insecto tanto en cultivo como en forma de insectos completos, y mamíferos transgénicos tales como ovejas y cabras.

20 Los sistemas de cultivo de células procarióticas son fáciles de mantener y de explotación económica. Sin embargo, las células procarióticas no son capaces de modificar postraduccionalmente las proteínas eucarióticas. Además, muchas proteínas se pliegan incorrectamente, requiriendo procedimientos específicos para plegarlas nuevamente, lo que se añade al coste de producción.

25 Se han descrito sistemas de cultivo de células eucarióticas para varias aplicaciones. Por ejemplo, algunas células de mamífero son capaces de realizar modificaciones postraducionales y generalmente producen proteínas que están correctamente plegadas y son solubles. Entre las desventajas principales de los sistemas de células de mamífero se incluyen la necesidad de instalaciones de cultivo especializadas y caras, el riesgo de infección, que puede llevar a la pérdida del cultivo completo y a un riesgo de contaminación del producto final con proteínas de mamífero potencialmente peligrosas. Las células de insecto se utilizan alternativamente para la expresión de polipéptidos. El sistema de expresión más extendido que se utiliza en células de insecto se basa en vectores baculovíricos. Se construye un vector de expresión baculovírico mediante la sustitución del gen polihedrina de baculovirus, que codifica una proteína estructural mayor del baculovirus, con un gen heterólogo, bajo el control del promotor de polihedrina nativo fuerte. Las células hospedadoras de insecto en cultivo resultan infectadas por el virus recombinante y la proteína producida de esta manera puede recuperarse a partir de las células mismas o partir del medio de cultivo en caso de que se utilicen señales de secreción adecuadas.

40 Sin embargo, ambos sistemas adolecen de problemas asociados a la reproducibilidad del nivel y calidad de expresión de las proteínas recombinantes, a la infección del cultivo y a que pueden requerir de instalaciones de cultivo especializadas. Además, las reservas de baculovirus, que para la producción de determinadas proteínas podrían requerir la preparación bajo condiciones GMP, no son estables durante el tiempo en todos los casos.

45 Las células de *Drosophila*, en particular las células de *Drosophila melanogaster* S2, para la expresión de proteínas se han dado a conocer en las patentes US nº 5.550.043, nº 5.681.713 y nº 5.705.359. En contraste con el sistema baculovírico de la técnica anterior, en el que la proteína de interés se proporciona únicamente tras la lisis de las células de insecto infectadas, el método basado en células S2 proporciona un sistema de expresión celular continua para las proteínas heterólogas y, por lo tanto, conduce a niveles de expresión más elevados.

50 Se han propuesto algunos otros medios para intensificar la expresión de proteínas heterólogas en células hospedadoras: por ejemplo, la patente US nº 5.919.682 describe un método de sobreproducción de ácido nítrico sintasa funcional en un procariota utilizando un vector pCW bajo el control del promotor *tac* y la coexpresión de la proteína con chaperones. Además, la patente US nº 4.758.512 se refiere a la producción de células hospedadoras que presentan mutaciones específicas dentro de las secuencias de ADN de las mismas que provocan que el organismo muestre una capacidad reducida de degradación de productos foráneos. Estos organismos huésped mutados pueden utilizarse para incrementar los rendimientos de proteínas foráneas manipuladas genéticamente.

60 Las células de vertebrado, en particular las células de mamífero, también han sido utilizadas ampliamente en la expresión de proteínas recombinantes. El nivel de producción de proteínas durante el tiempo de las células en cultivo depende de varios factores, tales como, por ejemplo, la densidad celular, la fase del ciclo celular, las tasas de biosíntesis celular de las proteínas, la condición del medio utilizado para la viabilidad y crecimiento celulares, y la longevidad de las células en cultivo (es decir, cuánto tiempo transcurre antes de que sucumban a la muerte celular programada, o apoptosis). Se han desarrollado diversos métodos para mejorar la viabilidad y el periodo de vida de las células en cultivo, conjuntamente con métodos de incremento de la productividad de una proteína deseada mediante, por ejemplo, el control de los nutrientes, la densidad celular, el contenido de oxígeno y

dióxido de carbono, la lactato deshidrogenasa, el pH, la osmolaridad, los catabolitos, etc.

Pueden utilizarse otras células hospedadoras para producir proteínas recombinantes heterólogas, notablemente células vegetales y células de levadura.

5 Se han sintetizado en plantas muchas proteínas farmacéuticas de origen mamífero. Entre ellas se incluyen productos sanguíneos, tales como albúmina de suero humano, para la que existe una demanda anual superior a 500 toneladas, y citoquinas y otras moléculas de señalización que se necesitan en cantidades mucho menores. La mayoría de proteínas derivadas de plantas han sido producidas en plantas del tabaco transgénicas y se han  
10 extraído directamente de las hojas. En general estas proteínas se producen a niveles bajos, típicamente menos de 0,1% de la cantidad total de proteínas solubles. Este nivel bajo de producción probablemente refleja una combinación de factores, siendo los más importantes un plegamiento y estabilidad pobres de las proteínas. Más recientemente, se ha utilizado el sistema de cloroplastos del tabaco para expresar proteínas humanas a niveles mucho más elevados (MA JKC *et al.*, 2004).

15 Los sistemas de levadura han resultado fundamentales para producir grandes cantidades proteínas para la utilización industrial y biofarmacéutica durante muchos años. Las levaduras pueden cultivarse hasta generar densidades de masa celular muy elevadas en un medio bien definido. Las proteínas recombinantes en levaduras pueden sobreexpresarse de manera que el producto se secrete de la célula y se encuentra disponible para la  
20 recuperación en la solución de fermentación. Las proteínas secretadas por las levaduras se encuentran fuertemente glucosiladas en sitios de glucosilación de consenso. De esta manera, la expresión de las proteínas recombinantes en los sistemas de levadura históricamente se ha restringido a proteínas en las que los patrones postraduccionales de glucosilación no afectan a la función de las proteínas. Se utilizan varios sistemas de expresión de levadura para la expresión de proteínas recombinantes, entre ellos *Saccharomyces*,  
25 *Schizosaccharomyces pombe*, *Pichia pastoris* y *Hansenula polymorpha*. Recientemente ha aparecido un nuevo sistema con la capacidad de producir glucoproteínas recombinantes en levaduras que presenta secuencias de glucosilación similares a las de las glucoproteínas humanas secretadas que se producen en las células de mamífero. La ruta de glucosilación de *Pichia pastoris* se modificó mediante la eliminación de enzimas endógenos, que añaden cadenas ricas en manosa a intermediarios de N-glucosilación. Además, se transfirieron  
30 específicamente al interior de *P. pastoris* por lo menos cinco enzimas activos participantes en la síntesis de cadenas oligosacáridas humanizadas. La capacidad de producir grandes cantidades de glucoproteínas humanizadas en levaduras ofrece la ventaja de que las estructuras glucosiladas pueden ser muy uniformes y fáciles de purificar. Además, la contaminación cruzada por virus de mamífero y glucoproteínas de otros huéspedes mamíferos puede eliminarse mediante la utilización de una producción semicontinua en una levadura,  
35 con tiempos de fermentación más cortos que en las células de mamífero.

Sin embargo, mediante la utilización de dichos sistemas, se producen proteínas heterólogas a razón de aproximadamente 1 a 2 mg/l en el sobrenadante de las células en cultivo, que es un nivel bastante bajo para los  
40 objetivos de la producción industrial.

De esta manera, existe una necesidad urgente de proporcionar un sistema que permita alcanzar un nivel significativamente elevado de expresión de proteínas heterólogas.

45 La presente invención satisface dicha necesidad y proporciona métodos de expresión de proteínas que alcanzan un nivel de producción hasta 100 veces superior a los medios existentes de producción de proteínas (es decir, hasta 200 mg/l de proteínas en el sobrenadante).

50 Los presentes inventores en efecto han demostrado que la utilización de un vector de nucleótidos codificante de una proteína derivada de la proteína humana 6-metilguanina-ADN metiltransferasa (MGMT<sub>h</sub>), estando unida dicha proteína derivada MGMT<sub>h</sub>, directamente o no, a una proteína de interés incrementa la producción de dicha proteína de interés hasta un rendimiento de 40 mg/l a 200 mg/l de promedio.

### Leyendas de las figuras

55 la figura 1 da a conocer (A) una vista esquemática del ARNm codificante de una secuencia de proteína de fusión MGMT de la invención, que contiene, de 5' a 3', un péptido de señal, la secuencia de ARNm de MGMT, un espaciador, un sitio de corte de proteasa, un gen de proteína recombinante (gen foráneo), un espaciador y un marcaje de etiqueta (His<sub>6</sub>), y (B) las secuencias de ADN y de aminoácidos de la misma parte del vector, que comprende: i) el péptido de señal BiP monocatenario (BiP<sub>mc</sub>) de insecto (en cursiva), ii) la secuencia de  
60 intensificador codificante de SNAP (en gris), iii) una secuencia de espaciador de ADN, iv) la secuencia codificante de sitio de enterocinasa (en negrita), v) los sitios de clonación EcoRV/XmaI (subrayados) y vi) el ADN codificante del marcaje de etiqueta de His (cursiva y negrita) (ver también la SEC ID nº 5).

65 la figura 2 da a conocer (A) la secuencia de aminoácidos de la proteína de fusión SNAP (en gris) y la nucleoproteína N del virus de la fiebre del valle del Rift (FVR.N., en negrita) unida a un marcaje de etiqueta de His, separando ambas proteínas con un espaciador GGGS, (B) ensayo de inmunotransferencia en

sobrenadante celular de células S2 transfectadas con el vector de ADN de SEC ID nº 19 (SNAP-FVR) estimuladas o no con cadmio durante 10 días, utilizando anticuerpos anti-His<sup>etiqueta</sup>, y (C) un ensayo de inmunotransferencia llevado a cabo con anticuerpos anti-His, en fracciones de proteína insolubles (INS) o solubles (SOL) de lisados de *E. coli* B21, portando dichas bacterias un plásmido pET302/FVR.N+proTEV+GST. (D) Un ensayo de inmunotransferencia que muestra la cantidad de SNAP-FVR.N en las muestras de fracción sucesivas, obtenidas tras una purificación en dos etapas de proteínas quimérica secretada SNAP-FVR.N a partir de células S2 estimuladas durante 10 días, utilizando columnas Talon y Superdex 75.

la figura 3 da a conocer (A) las secuencias de ADN y de aminoácidos de la proteína de fusión SNAP (en cursiva) y la forma soluble de la proteína de cubierta E del virus del Nilo Occidental (en gris), unido a un marcaje de etiqueta de His (en negrita), estando separadas las proteínas por un espaciador GGGS (SEC ID nº 20) y (B) ensayo de inmunotransferencia con anticuerpos anti-etiqueta His, que muestra la secreción de la forma soluble de la proteína de cubierta E del virus del Nilo Occidental en el sobrenadante de las células S2 transfectadas con el vector de ADN de la invención codificante de SNAP-WNsE (SEC ID nº 20) y estimuladas o no con cadmio durante 10 días.

la figura 4 da a conocer (A) un esquema del casete de ADN que contiene una señal de péptido BiP, una secuencia codificante de tipo MGMT (tipo SNAP), dos sitios de corte de pro-TEV (proteasa del virus del grabado del tabaco) en cada lado de la secuencia de IFN $\alpha$  (IFNAI\_hu) y un marcaje de etiqueta His, (B) las secuencias de ADN y de aminoácidos de la proteína de fusión SNAP (en gris, precedidas por una señal de péptido de insecto, en cursiva) e IFN $\alpha$  (en negrita), seguido de un marcaje de etiqueta de His (en cursiva y negrita), estando separadas las proteínas SNAP e IFN $\alpha$  por el sitio de corte de enterocinasa (subrayado) y una secuencia espaciadora GGGS. (C) Ensayo de inmunotransferencia con anticuerpos anti-etiqueta His, para detectar la expresión de IFN $\alpha$  en el sobrenadante de células S2 transfectadas con el vector de la invención codificante de IFN $\alpha$  (S2/SNAP-IFN) o un vector de control, estimulado o no con Cd<sup>2+</sup>. (D) Ensayo de inmunotransferencia con anticuerpos anti-SNAP en 10  $\mu$ l de sobrenadante de células S2/DeSNAPuniv-IFN $\alpha$  inducidas o no inducidas durante 10 días con cadmio, (E) actividad de luciferasa en células HeLa infectadas por virus Chikungunya expresantes de una luciferasa de *Renilla*, en donde se tratan dichas células con diferentes dosis de IFN $\alpha$ , de proveedor comercial (Intergen), o con la IFN $\alpha$  producida mediante el método de la invención. (F) Actividad de luciferasa en células HeLa infectadas por virus Chikungunya expresante de una luciferasa de *Renilla*, en donde se tratan dichas células con diferentes dosis de la proteína SNAP-IFN $\alpha$  obtenida mediante el procedimiento de producción de la invención.

la figura 5 representa las diferentes etapas del procedimiento de producción de proteínas recombinantes de la invención.

la figura 6 da a conocer (A) las secuencias de ADN y de aminoácidos de la proteína de fusión SNAP (en gris, precedida de una señal peptídica de insecto) y el granzima M, seguido de un marcaje de etiqueta His, estando separadas las proteínas SNAP y granzima M por el sitio de corte de enterocinasa y una secuencia espaciadora GGGS. (B) Vista esquemática de la proteína de fusión quimérica SNAP-GrM, que subraya los tres sitios de corte potenciales de la GrM proteasa en SNAP. (C) Ensayo de inmunotransferencia con anticuerpos anti-SNAP y anti-etiqueta His para la detección de la expresión de SNAP-GrM en el sobrenadante de células S2 transfectadas con el vector de la invención codificante de GrM (S2/SNAP-GrM, SEC ID nº 55).

la figura 7 da a conocer (A) un esquema del casete universal de ADN que contiene una señal peptídica de tipo BiP, una secuencia codificante de MGMT, dos sitios de corte pro-TEV en cada lado de la secuencia de IFN $\alpha$  (IFNAI\_hu) y un marcaje de etiqueta His, (B) las secuencias de ADN y de aminoácidos de la proteína de fusión SNAP (en gris, precedida de una señal peptídica de tipo BiP de insecto) e IFN $\alpha$ 1 humana (aminoácidos en negrita), seguido de un marcaje de etiqueta de His, estando separadas las proteínas SNAP e IFN $\alpha$  por el sitio de corte de pro-TEV y una secuencia espaciadora GGGS. (C) Ensayo de inmunotransferencia con anticuerpos anti-SNAP para la detección de la expresión de SNAP-IFN $\alpha$  en el sobrenadante de células HeLa transfectadas con un vector codificante de SNAP solo sin señal peptídica (vector pSNAPf) o un vector codificante de SNAP solo, precedido de la señal peptídica del virus Dengue (pDV1ssprM-SNAP) o el vector de la invención codificante de IFN $\alpha$ , que comprende la secuencia de ADN tal como se define en (A) (pDeSNAP-4/SNAP-IFNA1, SEC ID nº 57).

la figura 8A da a conocer un casete universal de ADN que contiene una señal peptídica de tipo BiP, una secuencia codificante de SNAP, dos sitios de corte de pro-TEV, un marcaje de etiqueta de His, cuatro sitios de clonación únicos, BamH1, EcoRV, XmaI y ApaI para la clonación de un gen de interés y una secuencia espaciadora GGGS (DeSNAP Univ, SEC ID nº 59 y nº 60). Los sitios únicos NheI en el extremo 5' y NotI/HindIII en el extremo 3' resultan necesarios para la etapa de subclonación en los vectores de expresión de mamífero (por ejemplos el plásmido pcDNA3 o pCI-neo) y los sitios únicos Bgl II en el extremo 5' y Age I en el extremo 3' resultan necesarios para la etapa de subclonación en el sistema DES no de vertebrado. El esquema en (B) da a conocer el casete universal de ADN que contiene una señal peptídica de tipo BiP, una

secuencia codificante de MGMT, dos sitios de corte de pro-TEV, un marcaje de etiqueta de His, cuatros sitios de clonación única: *BamH1*, *EcoRV*, *SmaI* y *ApaI* para la clonación de un gen de interés y una secuencia espaciadora GGGs (DeMGMT univ, SEC ID n° 69 y n° 70).

5 la figura 9 da a conocer un medio para insertar un gen foráneo de interés en DeMGMT Univ.

la figura 10 da a conocer la estabilidad térmica de las proteínas de fusión de SNAP llamadas CHIK.sE2-SNAP, SNAP-WN.EDIII y SNAP-IFN $\alpha$ I incubadas durante 4 días a -80°C, a 4°C, a 25°C o a 37°C (A) o durante dos meses a -80°C, a 4°C, a 25°C o 37°C (B).

10 la figura 11 da a conocer la producción de las proteínas de fusión SNAP-SSX2 y SNAP-sFasL por los vectores de la invención introducidos en células S2, tras 10 días de inducción con cadmio (+) o sin cadmio (-) en sobrenadante completo (A) o en diferentes fracciones (B).

15 la figura 12 da a conocer (A) un esquema del casete universal de ADN que contiene una señal peptídica de tipo BiP, una secuencia codificante de MGMT (de tipo SNAP), dos sitios de corte de pro-TEV, en cada lado del antígeno de cáncer SSX2 y un marcaje de etiqueta His, (B) un esquema del casete universal de ADN que contiene una señal peptídica de tipo BiP, una secuencia codificante de MGMT, dos sitios de corte de pro-TEV en cada lados de la proteína NERMCSL y un marcaje de etiqueta His, (C) un ensayo de inmunotransferencia en células HeLa transitoriamente transfectadas durante dos días utilizando anticuerpos de ratón anti-SNAP, que mostraba producción extracelular o intracelular de IFN $\alpha$ , SSX2 y NERMCSL.

25 la figura 13 da a conocer (A) un esquema del casete universal de ADN que contenía una señal peptídica de tipo BiP, una secuencia codificante de MGMT (de tipo SNAP), dos sitios de corte de pro-TEV en cada lado del polipéptido SULF\_h-2 $\Delta$ TMD, y un marcaje de etiqueta de His, (B) el ADN y las secuencias de aminoácidos de la proteína de fusión SNAP (en gris oscuro, precedido de una señal de péptido de tipo BiP de insecto) y hSULF-2 $\Delta$ TMD, seguido por un marcaje Histag, estando separadas las proteínas SNAP y SULF\_h-2 $\Delta$ TMD por el sitio de corte de pro-TEV y una secuencia espaciadora GGGs y (C) la actividad enzimática de DeSNAP-SULF\_h-2 $\Delta$ TMD quimérico secretado por las células HEK 293 transfectadas transitoriamente durante dos días con pcDNA3/DeSNAPuniv-SULF\_h-2 $\Delta$ TMD.

30 la figura 14 da a conocer (A) un esquema del casete de ADN que contenía una señal peptídica BiP, una secuencia codificante de MGMT (de tipo SNAP), dos sitios de corte de pro-TEV en cada lado de la proteína NERMCSL y un marcaje de etiqueta de His, y (B) un ensayo de inmunotransferencia utilizando anticuerpos anti-SNAP, para la detección de la expresión de la proteína NERMCSL en el sobrenadante de células S2 transfectadas con el vector de la invención codificante de la proteína NERMCSL (S2/SNAP-NERMCSL) o con un vector codificante de la proteína soluble E2 del virus Chikungunya (CHIK.sE2-SNAP) estimulada o no con Cd $^{2+}$ .

#### 40 Descripción detallada de la invención

En el contexto de la presente invención se aprecia que la coexpresión del enzima 6-metilguanina-ADN-metiltransferasa (MGMT) conjuntamente con una proteína recombinante de interés mejoraba mucho la producción de dicha proteína recombinante en células de insecto, tales como células S2, así como en células de mamífero, tales como en células HeLa.

El enzima 6-metilguanina-ADN-metiltransferasa (MGMT, también conocido como ATasa o AGT, y en lo sucesivo denominado "MGMT") presenta la referencia EC 2.1.1.63 en la nomenclatura de enzimas del IUBMB. Es un enzima de reparación del ADN 6-alkilguanina-ADN-alkiltransferasa de 207 residuos aminoácidos la función del cual en las células es reparar el ADN alquilado. Más exactamente, MGMT actúa sobre la guanina O $^6$ -metilada en el ADN mediante la transferencia del grupo metilo en una reacción S $_N$ 2 a un residuo de cisteína reactiva (Cys 145). El mecanismo de reparación es inusual, ya que la proteína resulta inactivada irreversiblemente (Pegg A.E. *et al.*, *Mutat. Res.* 462:82-100, 2000). Este enzima se utiliza actualmente en biología molecular para el marcaje de proteínas in vivo con moléculas informadoras, mediante una reacción de marcaje irreversible con derivados O $^6$ -bencilguanina (Juillerat A. *et al.*, *Chemistry & Biology* 10:313-317, 2003 y documento n° WO 2005/085470).

Hasta el momento se han descrito diferentes enzimas derivados de MGMT (Lim A. *et al.*, *EMBO J.* 15:4050-4060, 1996; Daniels D.S. *et al.*, *EMBO J.* 19:1719-1730, 2000, Juillerat A. *et al.*, *Chemistry & Biology* 10:313-317, 2003, documentos n° WO 2005/085470, n° WO 2004/031405). En particular, se ha obtenido una proteína mutante de 20 kDa que contiene las mutaciones Cys62Ala, Lys125Ala, Ala127Thr, Arg128Ala, Gly131Lys, Gly132Thr, Met134Leu, Arg135Ser, Cys150Ser, Asn157Gly, Ser159Glu truncada en el aminoácido 182 (el mutante denominado "AGT26" en el documento n° WO 2005/085470, también denominado "SNAP 26" en el documento n° WO 2006/114409). Se ha demostrado que el mutante particular "SNAP26" presenta actividad de marcaje incrementada. Sin embargo, nunca se ha demostrado o sugerido que pueda incrementar la expresión de las proteínas recombinantes a las que se acopla.

En la presente memoria los presentes inventores proponen por primera vez la utilización del enzima MGMT (EC 2.1.1.63), un mutante, un dominio catalítico del mismo o subfragmentos del mismo, para incrementar la producción de proteína en las células hospedadoras, en particular en células hospedadoras no de vertebrado y de vertebrado. El efecto de potenciación se observa al expresar las células hospedadoras un polipéptido de fusión que comprende por lo menos: i) una señal peptídica de secreción que es funcional en dichas células hospedadoras, ii) el enzima MGMT, mutante, dominio catalítico o subfragmentos del mismo, e iii) la proteína de interés. Para que se produzca el efecto de potenciación, el enzima MGMT debe unirse físicamente, de manera directa o indirecta (pueden introducirse espaciadores y otros aminoácidos), a la proteína de interés. Sin restringirse a ninguna teoría, se encuentra contemplado que el enzima MGMT pueda servir como proteína chaperona, por ejemplo favoreciendo la secreción a partir de la célula hospedadora y estabilizando el polipéptido de fusión sintetizado en el sobrenadante de las células hospedadoras, o evitando que resulte metabolizado durante y después de la síntesis y secreción del mismo a partir de las células hospedadoras.

Además, se ha observado que el MGMT presenta una estructura globular 3D que comprende una hélice  $\alpha$  (Wibley J.E.A. *et al.*, 2000) que es compatible con una función de andamiaje de MGMT.

En el contexto de la presente exposición, las células "huésped" son cualesquiera células que pueden utilizarse para producir proteínas recombinantes, tales como células "no de vertebrado" (o de invertebrado), células de vertebrado, células vegetales, células de levadura o células procarióticas. Preferentemente son células no de vertebrado y de vertebrado.

Los no vertebrados (también conocidos como invertebrados) comprenden diferentes divisiones, siendo las más conocidas: Insectos, Arácnidos, Crustáceos, Moluscos, Anélidos, Cirrípedos, Radiados, Celenterados e Infusorios. En la actualidad se clasifican en más de 30 divisiones, desde organismos simples tales como esponjas de mar y platelmintos, hasta animales complejos tales como artrópodos y moluscos. En el contexto de la invención, se utilizan células no de vertebrado y dichas células son células de insecto, tales como células de *Drosophila* o de Mosquito, más preferentemente células de *Drosophila* S2.

Entre los ejemplos de células derivadas de organismos vertebrados que se dan a conocer en la presente memoria como líneas de células hospedadoras se incluyen células madre embrionarias no humanas o derivados de las mismas, por ejemplo células EBX aviares, la línea CVI de riñón de mono transformada por secuencias de SV40 (COS-7, ATCC nº CRL 1651), la línea renal embrionaria humana (293), las células renales de hámster neonato (BHK, ATCC nº CCL 10), células de ovario de hámster chino (CHO), células de Sertoli de ratón [TM4], células renales de mono (CVI, ATCC nº CCL 70), células renales de mono verde africano (VERO-76, ATCC nº CRL-1587), células de carcinoma cervical humano (HeLa, ATCC nº CCL 2), células renales caninas (MDCK, ATCC nº CCL 34), células hepáticas de rata búfalo (BRL 3A, ATCC nº CRL 1442), células pulmonares humanas (W138, ATCC nº CCL 75), células hepáticas humanas (Hep G2, HB 8065), células de tumor mamario de ratón (MMT 060562, ATCC nº CCL51), células de hepatoma de rata [HTC, MI.5], y células YB2/O (ATCC nº CRL1662), NIH3T3, HEK y TRI. En el contexto de la invención, las células de vertebrado preferentemente son células EBX, CHO, YB2/O, COS, HEK, NIH3T3 o derivados de las mismas.

Las células vegetales que pueden utilizarse en el contexto de la presente exposición son los cultivares del tabaco Bright Yellow 2 (BY2) y *Nicotiana tabaccum* 1 (NT-1).

Las células de levadura que pueden utilizarse en el contexto de la presente exposición son: *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* y *Hansenula polymorpha*, así como levaduras metilotróficas como *Pichia pastoris* y *Pichia methanolica*.

Las células procarióticas que pueden utilizarse en el contexto de la presente exposición son típicamente bacterias de *E. coli* o bacterias *Bacillus subtilis*.

La presente exposición se refiere a un vector de expresión de nucleótidos codificantes de por lo menos a) una señal de secreción peptídica, que preferentemente es funcional en células no de vertebrado o en células de vertebrado, y b) un enzima 6-metilguanina-ADN-metiltransferasa, un mutante, un subfragmento o un dominio catalítico de la misma. El vector de expresión de la presente invención se describe en mayor detalle posteriormente.

El término "vector" en la presente memoria se refiere al vehículo en el que puede introducirse una secuencia de ADN o ARN de un gen foráneo en una célula hospedadora de manera que se transforma y se promueve la expresión de la secuencia introducida. Entre los vectores pueden incluirse, por ejemplo, plásmidos, fagos y virus y se comentan en mayor detalle posteriormente. En efecto, puede utilizarse cualquier tipo de plásmido, cósmido, YAC o vector vírico para preparar un constructo de ácidos nucleicos recombinantes que pueden introducirse en una célula hospedadora en la que se desea la expresión de la proteína de interés. Alternativamente, en el caso de que se desee la expresión de la proteína de interés en un tipo particular de célula hospedadora, pueden utilizarse vectores víricos que infectan selectivamente el tipo celular o tipo tisular deseado. También resulta importante en el contexto de la invención los vectores para la utilización en la terapia génica (es decir, los que

son capaces de transportar la molécula de ácidos nucleicos a un organismo huésped).

Por ejemplo, vectores víricos, tales como lentivirus, retrovirus, virus herpes, adenovirus, virus adenoasociados, virus Vaccinia, baculovirus y otros virus recombinantes con tropismo celular deseable. Los métodos para la construcción y la utilización de vectores víricos son conocidos de la técnica (ver Miller y Rosman, BioTechniques 7:980-990, 1992).

Los vectores víricos que actualmente resultan preferentes en la presente invención y exposición son aquellos que resultan idóneos para la utilización en células de vertebrado y de no vertebrado.

Para las células no de vertebrado de la presente invención, los vectores preferentes son los arbovirus, resultando particularmente preferido el virus del Nilo Occidental, los cuales son vectores artrópodos. Otros vectores que es conocido que son expresados eficientemente en las células no de vertebrado son los baculovirus.

Para las células de vertebrado de la presente exposición resultan preferidos los vectores lentivíricos, VAA, baculovíricos y adenovíricos. Los vectores adecuados para la expresión en las células hospedadoras de mamífero también pueden ser de origen no vírico (por ejemplo ADN plasmídico). Entre los vectores plasmídicos adecuados se incluyen, aunque sin limitación, pREP4, pCEP4 (Invitrogene), pCI (Promega), pCDM8 y pMT2PC, pVAX y pgWiz.

Como células procarióticas de la presente exposición resultan preferidos los vectores plásmido, bacteriófago y cósmido. Entre los vectores adecuados para la utilización en sistemas procarióticos se incluyen, aunque sin limitación, los vectores pBR322 (Gibco BRL), pUC (Gibco BRL), pBluescript (Stratagene), pPoly, pTrc, pET 11d, pIN y pGEX.

Para las células vegetales de la presente exposición, resultan preferentes los vectores de expresión plasmídicos, tales como los plásmidos Ti y los vectores de expresión víricos, tales como el virus del mosaico de la coliflor (CaMV, por sus siglas en inglés) y el virus del mosaico del tabaco (TMV, por sus siglas en inglés).

La expresión de proteínas recombinantes en las células de levadura puede llevarse a cabo utilizando tres tipos de vector: vectores de integración (YIp), plásmidos episómicos (YEpl) y plásmidos centroméricos (YCp). Entre los vectores adecuados para la expresión en levaduras (por ejemplo *S. cerevisiae*) se incluyen, aunque sin limitación, pYepSec1, pMFa, pJRY88, pYES2 (Invitrogen Corporation, San Diego, Calif.) y pTEF-MF (Dualsystems Biotech código de producto P03303).

Los vectores que pueden utilizarse para la terapia génica son bien conocidos de la técnica. Son, por ejemplo, lentivirus, retrovirus, adenovirus, poxvirus, virus herpes, virus del sarampión, virus espumoso o virus adenoasociados (VAA). Los vectores víricos pueden ser competentes para la replicación o pueden discapacitarse genéticamente de manera que sean defectuosos para la replicación o de replicación deteriorada.

Los vectores de terapia génica preferentes son los vectores de ADN solapado ("DNA flap"), tal como se indica en el documento nº WO 1999/055892, la patente US nº 6.682.507 y el documento nº WO 2001/27300.

Una secuencia "codificante" de un producto de expresión, tal como un ARN, polipéptido, proteína o enzima, es una secuencia de nucleótidos que, al expresarse, resulta en la producción de dicho ARN, polipéptido, proteína o enzima; es decir, la secuencia de nucleótidos "codifica" dicho ARN o codifica la secuencia de aminoácidos para dicho polipéptido, proteína o enzima.

En el contexto de la presente exposición, el "dominio catalítico" de un enzima se refiere al sitio activo del enzima o, en otras palabras, la parte de una molécula de enzima en la que se produce la catálisis del sustrato (en la presente memoria, la transferencia del grupo metilo en una reacción  $S_N2$  a un residuo de cisteína reactivo). Por lo tanto, la expresión "un dominio catalítico del mismo" se refiere a cualquier fragmento o secuencia homóloga del polipéptido MGMT, preferentemente que presenta una actividad catalítica de por lo menos 80% la del enzima MGMT nativo. Estos fragmentos (también denominados "subfragmentos") pueden comprender entre 20 y 180, preferentemente entre 30 y 100 aminoácidos. La secuencia homóloga de dicho dominio catalítico puede presentar una o más mutaciones, resultando en la pérdida parcial o total de dicha actividad catalítica.

En el contexto de la invención, el enzima MGMT puede ser la MGMT humana (con la referencia NP\_002403.2) de secuencia SEC ID nº 4, la MGMT de ratón identificada como NP\_032624.1 (SEC ID nº 45), la MGMT de rata identificada como NP\_036993.1 (SEC ID nº 46) o una secuencia homóloga que presenta una identidad de por lo menos 70% respecto a la SEC ID nº 4. También se dan a conocer en la presente memoria secuencias homólogas de SEC ID nº 4, SEC ID nº 45 y SEC ID nº 46.

El término "homólogo" se refiere a secuencias que presentan similitud de secuencia. La expresión "similitud de secuencia", en todas las formas gramaticales de la misma, se refiere al grado de identidad o correspondencia entre las secuencias de ácidos nucleicos o de aminoácidos. En el contexto de la invención, dos secuencias de aminoácidos son "homólogas" en el caso de que por lo menos aproximadamente 80%, alternativamente por lo

5 menos aproximadamente 81%, alternativamente por lo menos aproximadamente 82%, alternativamente por lo  
 menos aproximadamente 83%, alternativamente por lo menos aproximadamente 84%, alternativamente por lo  
 menos aproximadamente 85%, alternativamente por lo menos aproximadamente 86%, alternativamente por lo  
 menos aproximadamente 87%, alternativamente por lo menos aproximadamente 88%, alternativamente por lo  
 5 menos aproximadamente 89%, alternativamente por lo menos aproximadamente 90%, alternativamente por lo  
 menos aproximadamente 91%, alternativamente por lo menos aproximadamente 92%, alternativamente por lo  
 menos aproximadamente 93%, alternativamente por lo menos aproximadamente 94%, alternativamente por lo  
 menos aproximadamente 95%, alternativamente por lo menos aproximadamente 96%, alternativamente por lo  
 10 menos aproximadamente 97%, alternativamente por lo menos aproximadamente 98%, alternativamente por lo  
 menos aproximadamente 99% de los aminoácidos son similares. Preferentemente, las secuencias polipeptídicas  
 similares u homólogas se identifican mediante la utilización del algoritmo de Needleman y Wunsch.

15 Preferentemente, la secuencia homóloga respecto al enzima 6-metilguanina-ADN-metiltransferasa comparte una  
 identidad de secuencia de aminoácidos de por lo menos 64%, preferentemente una identidad de secuencia de  
 aminoácidos de por lo menos 65%, alternativamente una identidad de secuencia de aminoácidos de por lo  
 menos aproximadamente 66%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 67%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 68%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 20 aproximadamente 69%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 70%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 71%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 72%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 73%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 25 aproximadamente 74%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 75%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 76%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 77%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 78%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 79%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 30 aproximadamente 80%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 81%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 82%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 83%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 84%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 35 aproximadamente 85%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 86%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 87%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 88%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 89%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 40 aproximadamente 90%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 91%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 92%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 93%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 94%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 45 aproximadamente 95%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 96%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 97%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 aproximadamente 98%, alternativamente una identidad de secuencia de aminoácidos de por lo menos  
 50 aproximadamente 99% respecto a SEC ID nº 4. La secuencia homóloga respecto a la SEC ID nº 4 según la  
 invención es por lo menos 70%, y más preferentemente 80% idéntica a la SEC ID nº 4.

Una secuencia de MGMT homóloga más preferida contiene las mutaciones indicadas en el documento  
 nº WO 2005/085470, las posiciones de las cuales pueden transponerse fácilmente en vista de la SEC ID nº 4, el  
 residuo de metionina de partida de SNAP26 correspondiente al residuo de metionina en la posición 32 de la SEC  
 55 ID nº 4 (por lo tanto, deberían añadirse 31 aminoácidos a las posiciones dadas a conocer en el documento  
 nº WO 2005/085470 con el fin de obtener las correspondientes en la SEC ID nº 4).

Preferentemente, la secuencia homóloga de MGMT útil en la invención corresponde a la secuencia de MGMT de  
 tipo salvaje de SEC ID nº 4, en la que se han eliminado entre 1 y 30, preferentemente entre 6 y 25, y en  
 60 particular 14, 15, 16, 17, 18, 19, 20, 21, 22 o 23 aminoácidos son sustituidos por otros aminoácidos, y/o 1 a 40,  
 preferentemente 1 a 20, en particular 10 a 20 aminoácidos, más preferentemente 15 aminoácidos en el extremo  
 C-terminal.

En una forma de realización preferida, la secuencia homóloga de MGMT contiene las mutaciones siguientes  
 65 respecto a la SEC ID nº 4:



- 5 (A) Lys31 sustituido por Arg, o Met32 sustituido por Ser, o Cys93 sustituido por Ala, o Lys156 sustituido por Ala, o Ala158 sustituido por Thr, o Arg159 sustituido por Ala, o Gly162 sustituido por Lys, o Gly163 sustituido por Thr, o Met165 sustituido por Leu, o Arg166 sustituido por Ser, o Cys181 sustituido por Ser, o Asn188 sustituido por Gly, o Ser190 sustituido por Glu, o Gly214 sustituido por Pro, o Ser215 sustituido por Ala, o Ser216 sustituido por Gly, o Gly217 sustituido por Ile, o Leu218 sustituido por Gly, o Gly220 sustituido por Pro, o Ala221 sustituido por Gly, o Trp222 sustituido por Ser, o
- 10 (B) Lys31-Met32 sustituido por Arg-Ser, o Ala158-Arg159 sustituido por Thr-Ala, r Gly162-Gly163 sustituido por Lys-Thr, o Met165-Arg166 sustituido por Leu-Ser, o Gly162-Gly163/Met165-Arg166 sustituido por Lys-Thr/Leu-Ser, o Asn188/Ser190 sustituido por Gly/Glu, o Gly214-Ser215-Ser216-Gly217-Leu218 sustituido por Pro-Ala-Gly-Ile-Gly, o Gly220-Ala221-Trp222 sustituido por Pro-Gly-Ser, preferentemente en combinación con cualesquiera otras sustituciones de aminoácidos citadas en (A), o
- 15 (C) Truncado después de Leu223 (los aminoácidos 224 a 238 han sido eliminados), preferentemente en combinación con cualquier otra sustitución de aminoácido citada en (A) o (B).

Las secuencias homólogas de MGMT preferidas son las truncadas después de Leu223.

20 Las secuencias homólogas de MGMT preferidas son aquellas con dos de las modificaciones de (B) y opcionalmente con truncado después de Leu223.

Las secuencias homólogas de MGMT preferidas son aquellas con tres de las modificaciones de (B) y opcionalmente con truncado después de Leu223.

25 Las secuencias homólogas de MGMT preferidas son aquellas con cuatro de las modificaciones de (B) y opcionalmente con truncado después de Leu223.

Las secuencias homólogas de MGMT preferidas son aquellas con cinco de las modificaciones de (B) y opcionalmente con truncado después de Leu223.

30 Las secuencias homólogas de MGMT preferidas son aquellas con seis de las modificaciones de (B) y opcionalmente con truncado después de Leu223.

35 Otras secuencias homólogas de MGMT preferidas son las que contienen una combinación de 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 mutaciones seleccionadas de entre las modificaciones dadas a conocer en (A), y opcionalmente truncadas después de Leu223.

40 Resultan particularmente preferidas las secuencias homólogas que contienen las mutaciones Lys31Arg, Met32Ser, Cys93Ala, Lys156Ala, Ala158Thr, Arg159Ala, Gly162Lys, Gly163Thr, Met165Leu, Arg166Ser, Cys181Ser, Asn188Gly, Ser190Glu, Gly214Pro, Ser215Ala, Ser216Gly, Gly217Ile, Leu218Gly, Gly220Pro, Ala221Gly, Trp222Ser y truncado después de Leu223 (es decir, la secuencia SNAP de la SEC ID nº 2).

45 En una forma de realización todavía más preferente, el enzima MGMT es la proteína mutante SNAP de SEC ID nº 2 o un homólogo de la misma tal como se define a continuación. El mutante SNAP de SEC ID nº 2 comparte una homología de 77% con la secuencia de aminoácidos de la 6-metilguanina-ADN-metiltransferasa humana (NP\_002403.2, SEC ID nº 4) y una homología de 70% respecto a la secuencia de aminoácidos de la 6-metilguanina-ADN-metiltransferasa de ratón (NP\_032624.1, SEC ID nº 45).

50 Dicha secuencia homóloga respecto a la proteína SNAP es por lo menos idéntica a más del 80%, preferentemente al 81%, más preferentemente al 82%, más preferentemente al 83%, más preferentemente al 84%, más preferentemente al 85%, más preferentemente al 86%, más preferentemente al 87%, más preferentemente al 88%, más preferentemente al 89%, más preferentemente al 90%, más preferentemente al 91%, más preferentemente al 92%, más preferentemente al 93%, más preferentemente al 94%, más preferentemente al 95%, más preferentemente al 96% y todavía más preferentemente al 97% respecto a la proteína SNAP de secuencia SEC ID nº 2.

60 Preferentemente, el vector de expresión de nucleótidos de la invención comprende además sitios de clonación que permiten la inserción dentro del marco de una secuencia de ADN heteróloga codificante de una proteína de interés.

La expresión "señal de secreción peptídica" tal como se utiliza en la presente invención se refiere a una cadena peptídica corta (de 3 a 60 aminoácidos de longitud) que dirige el transporte de una proteína.

65 Entre los ejemplos de señales de secreción apropiadas para la presente invención se incluyen, aunque sin limitación, las secuencias peptídicas de señal del factor de apareamiento (MF, por sus siglas en inglés) alfa (patente US nº 5.879.926), la invertasa (documento nº WO84/01153), PHO5 (patente DK nº 3614/83), YAP3

(documento nº WO95/02059) y BAR1 (documento nº WO87/02670).

5 En el contexto de la invención, dicha señal de secreción peptídica es funcional en células de insecto, que son células de no vertebrado. Las señales de secreción peptídica que son funcionales en las células de vertebrado también se dan a conocer en la presente memoria. Las señales de secreción peptídica que son funcionales tanto en células de insecto como en células de vertebrado también se encuentran comprendidas en la presente memoria.

10 Son ejemplos de señales de secreción peptídica que son funcionales en las células de insecto: BiPmc de insecto (SEC ID nº 48, que presenta por ejemplo la secuencia de ADN SEC ID nº 11), la señal peptídica de tipo BiP de SEC ID nº 51 (por ejemplo que presenta la secuencia de ADN SEC ID nº 50) y cualquier señal peptídica presente en un arbovirus, por ejemplo la proteína E de cubierta del virus del Nilo Occidental (SEC ID nº 15).

15 Resulta interesante que la señal peptídica de tipo BiP anteriormente indicada es funcional en células tanto de no vertebrado como de vertebrado. Esta señal de tipo BiP corresponde a la señal peptídica BiP de SEC ID nº 48 en la que el último aminoácido glicina ha sido sustituido por la secuencia de aminoácidos Pro Thr Ala Leu Ala (SEC ID nº 61), que corresponde al sitio de corte de la proteína E del virus Dengue. De acuerdo con lo anteriormente expuesto, la señal de tipo BiP es cortada ventajosamente tras la traducción de la proteína y la secreción al sobrenadante de las células hospedadoras.

20 Tal como se da a conocer en la presente memoria, también se encuentra disponible una diversidad de señales de secreción para la expresión en células hospedadoras de levadura, por ejemplo en *S. cerevisiae*. Entre ellas se incluyen el factor alfa Prepro, HSp150, PHO1, SUC2, KILM1 (toxina asesina tipo 1) y GGP1.

25 Un sitio de clonación es una secuencia que facilita la clonación de un gen codificante de una proteína de interés en el sistema de expresión. Contiene sitios de restricción, o sitios de reconocimiento de restricción, es decir localizaciones en una molécula de ADN que contiene secuencias específicas de nucleótidos, los cuales son reconocidos por enzimas de restricción (ver, por ejemplo, en las figuras). Éstas son secuencias generalmente palindrómicas (debido a que los enzimas de restricción habitualmente se unen como homodímeros) y un enzima de restricción particular puede cortar la secuencia entre dos nucleótidos dentro de su sitio de reconocimiento, o en una localización próxima. Los sitios de clonación son bien conocidos por el experto en la materia.

30 Más preferentemente, el vector de expresión de nucleótidos comprende además una secuencia de ADN heterólogo codificante de una proteína de interés heteróloga o un polipéptido heterólogo insertado en dichos sitios de clonación.

35 El término "heterólogo" se refiere a una combinación de elementos no natural. Por ejemplo, la presente invención incluye "secuencias de ADN heterólogas" codificantes de "proteínas/polipéptidos de interés", estando dichas secuencias de ADN no situadas naturalmente en o dentro de un sitio cromosómico de la célula hospedadora que se utiliza para la expresión de proteínas.

40 En el caso de que se inserte en el vector de nucleótidos de la invención una secuencia de ADN heteróloga codificante de una proteína o polipéptido heterólogo de interés, preferentemente codifica un polipéptido de fusión que comprende dicha señal peptídica, dicho enzima MGMT, mutante u homólogo del mismo, y dicha proteína/polipéptido heterólogo de interés.

45 En una forma de realización preferida de la invención, la secuencia de ADN codificante de dicho enzima MGMT se encuentra situada en el extremo 5' o 3' de la secuencia de ADN codificante de dicha proteína heteróloga de interés, preferentemente en el extremo 5'. Por lo tanto, el enzima MGMT se une directa o indirectamente a la proteína/polipéptido heterólogo de interés y preferentemente se localiza en el extremo N-terminal de la proteína/polipéptido heterólogo de interés.

50 Resulta particularmente preferido que la secuencia de ADN codificante de dicho enzima MGMT de la misma se encuentre situada en el extremo 5' de la secuencia de ADN codificante de dicha proteína/polipéptido heterólogo de interés, en el caso de que el dominio de actividad de la proteína/polipéptido heterólogo de interés se encuentre localizado en su parte C-terminal, tal como IFN $\alpha$ . De la misma manera, podría resultar particularmente preferido que la secuencia de ADN codificante de dicho enzima MGMT se encuentre situada en el extremo 3' de la secuencia de ADN codificante de dicha proteína/polipéptido heterólogo de interés, en el caso de que el dominio de actividad de la proteína/polipéptido heterólogo de interés se encuentre localizado en su parte N-terminal.

55 Más exactamente, en un primer aspecto, la presente exposición se refiere a un vector para la expresión de proteínas recombinantes en células hospedadoras, preferentemente en células hospedadoras no de vertebrado y/o de vertebrado, más preferentemente en células de insecto, que comprende una secuencia de nucleótidos codificante en un único marco de lectura abierto, de extremo 5' a 3':

- a) una señal de secreción peptídica que es funcional en dicha célula hospedadora,
- b) el enzima 6-metilguanina-ADN-metiltransferasa (MGMT, EC 2.1.1.63), un mutante o un dominio catalítico del mismo y
- c) una proteína recombinante.

Todavía más exactamente, la presente invención se refiere a un vector para la expresión de proteínas recombinantes en células de insecto, tal como se define en cualquiera de las reivindicaciones 7 a 9.

En el contexto de la invención, la expresión "proteína recombinante" o "proteína de interés" se refiere a productos génicos o polipéptidos que son foráneos respecto a la célula productora de proteína y que se seleccionan preferentemente de entre el grupo que consiste en una o más proteínas o polipéptidos diagnósticos y terapéuticos.

Más preferentemente, dicha proteína o proteínas o polipéptido o polipéptidos diagnósticos y terapéuticos se seleccionan de entre el grupo que consiste de:

- proteínas inmunogénicas bacterianas o víricas, más preferentemente proteínas víricas (infecciosas, patógenas), por ejemplo la proteína EDIII de los virus Dengue, de la encefalitis japonesa, de la encefalitis transmitida por garrapatas, de la fiebre amarilla, Usutu, Rocio, de la encefalitis de Murray, Wesselbron, Zika o del Nilo Occidental, o la nucleoproteína N del virus de la fiebre del valle del Rift o del virus Toscana, o la forma soluble de la proteína de cubierta E2 del virus Chikungunya, o la forma soluble de la proteína de cubierta E del virus del Nilo Occidental, y
- factores sanguíneos, anticoagulantes, factores de crecimiento, hormonas, vacunas, enzimas terapéuticos, anticuerpos monoclonales y citoquinas (tales como IFN $\alpha$ , granzima M y FasL),
- antígenos, por ejemplo antígenos del cáncer tal como el antígeno de cáncer de testículo SSX2 o la región N-terminal de ERC/mesotelina (NERCMSL),
- proteínas antitumorales, por ejemplo FasL, o las heparán-sulfato 6-O-endosulfatasas (hSULF),
- polipéptidos microbianos, víricos y/o parasitarios,
- cualesquiera otras proteínas útiles (por ejemplo contactinas).

La proteína FasL es una proteína proapoptótica que puede utilizarse como agente antitumoral. Se encuentra codificada, por ejemplo, por la SEC ID n° 88.

Las proteínas hSulf (o hSULF) son heparán-sulfato 6-O-endosulfatasas que regulan la estructura de la heparina sulfato y presentan un impacto drástico sobre el crecimiento y la progresión de las células malignas in vivo (Dai *et al.*, 2005). En el contexto de la invención, preferentemente es la proteína hSulf2, y más preferentemente hSulf-2<sup>ΔTMD</sup>, en las que el dominio transmembrana (TMD) ha sido eliminado para incrementar su solubilidad, presentando este mutante la secuencia de aminoácidos SEC ID n° 95 y estando codificada, por ejemplo, en la SEC ID n° 94.

El vector de la invención también puede utilizarse para expresar y purificar péptidos y/o polipéptidos de interés. En el contexto de la presente invención, los términos "péptidos" y "polipéptidos" pretenden ser sinónimos, refiriéndose a polímeros cortos de monómeros aminoácidos (también denominados "residuos") unidos mediante enlaces peptídicos. Dichos polímeros preferentemente contienen menos de 100 residuos, más preferentemente menos de 50 residuos.

En particular, el vector de la invención puede utilizarse para expresar y purificar polipéptidos microbianos diagnósticos, tales como polipéptidos bacterianos, víricos o parasitarios. Son ejemplos de dichos polipéptidos los péptidos antigénicos, mucinas y/o toxinas secretadas o expresadas por bacterias, virus o parásitos. Preferentemente, dicho péptido antigénico es expresado por el virus Influenza, el virus de la hepatitis A, el virus de la hepatitis B, el virus de la hepatitis C, el virus de la hepatitis G, el virus VIH, el virus de la fiebre amarilla, el virus Dengue, el virus de la encefalitis japonesa, el virus de la encefalitis transmitida por garrapatas, los virus Usutu o del Nilo Occidental, el virus de la fiebre del valle del Rift o el virus Toscana, el virus Chikungunya, el virus sincitial respiratorio, el virus Rocio, el virus de la encefalitis de Murray, el virus Wesselbron, el virus Zika, el virus de la coreomeningitis linfocítica, un paravovirus humano, un papilomavirus humano, el citomegalovirus humano o cualquier virus identificado. Preferentemente, dicho péptido antigénico es expresado por protozoos parasitarios (tal como *Entamoeba histolytica* o *Giardia lamblia*), anélidos (tales como nemátodos, céstodos o tremátodos) o artrópodos (tales como crustáceos, insectos o arácnidos). Preferentemente dicho péptido antigénico es

expresado por bacterias infecciosas, por ejemplo de los géneros *Streptococcus* o *Staphylococcus* y *E. coli*. Las toxinas infecciosas son bien conocidas de la técnica. Pueden citarse a título de ejemplos las neurotoxinas botulínicas, la toxina epsilon de *Clostridium perfringens*, la ricina, la saxitoxina, la shigatoxina, la tetrodotoxina, las enterotoxinas estafilocócicas, etc. Las mucinas también son bien conocidas de la técnica. Son ejemplos de las mismas MUC5AC, MUC5B y MUC2. Dichos ejemplos no son limitativos y puede expresarse cualquier péptido/polipéptido mediante el método de la invención.

Las contactinas son un subgrupo de moléculas pertenecientes a la superfamilia de las inmunoglobulinas que se expresan exclusivamente en el sistema nervioso (ver la revisión de Shimoda y Watanabe, 2009). Participan en trastornos psiquiátricos, en particular en el autismo. Las contactinas preferentes que debe producir el sistema de la invención son las contactinas 2 y 4. La contactina 4 (CNTN4) está codificada por, por ejemplo, la SEC ID nº 91 (correspondiente a los aminoácidos 19 a 990 de la proteína completa NP\_783200.1).

Es conocido que numerosos antígenos del cáncer resultan eficientes dianas de vacuna para el tratamiento del cáncer. La producción en cantidad elevada de dichos polipéptidos (ver las listas en Cheever *et al.*, 2009) aparentemente resulta muy importante para obtener una vacuna del cáncer eficiente. Resulta interesante que los vectores de la invención permiten obtener un nivel elevado de antígenos recombinantes del cáncer que pueden utilizarse en inmunoterapia, o para producir anticuerpos o en métodos diagnósticos del cáncer.

SSX2 y NERCMSL son dos ejemplos de antígenos del cáncer. El antígeno de cáncer SSX2 está codificado por el ADN que presenta la SEC ID nº 76 (GenBank: NM\_175698). La región N-terminal de ERC/mesotelina (NERCMSL) está codificada por la SEC ID nº 83. Este antígeno se utiliza comúnmente como antígeno de detección en pacientes que sufren de mesotelioma maligno.

Mediante los métodos de la invención puede producirse cualquier proteína.

Sin embargo, las proteínas preferentes son las proteínas terapéuticas, tales como la insulina, IFN, FasL, la mesotelina, hSULF o las contactinas.

Más generalmente, las proteínas preferentes son aquellas que ha resultado difícil producir en cantidades elevadas hasta el momento. Dichas proteínas son, por ejemplo, FasL, granzima M, hSULF, mesotelina y contactinas.

La secuencia de ADN codificante del polipéptido de fusión que comprende dicha señal peptídica, dicho enzima MGMT, o mutante, y dicha proteína de interés recombinante, puede asociarse operativamente con un promotor inducible que resulta funcional en las mismas células hospedadoras en las que es funcional la señal peptídica.

Más preferentemente, en el vector de la invención, dicho marco de lectura abierto se encuentra asociado operativamente a un promotor inducible que es funcional en la misma célula hospedadora en la que es funcional la señal peptídica.

Una secuencia codificante se encuentra "operativamente asociada con" una secuencia de control de la expresión (es decir, secuencias de control transcripcional y traduccional) en una célula, en el caso de que la ARN polimerasa transcriba la secuencia codificante en ARN, que seguidamente es procesado en trans-ARN (en caso de contener intrones) y, en el caso de que la secuencia codifique una proteína, se traduce en dicha proteína.

Un "promotor" es una secuencia de nucleótidos a partir de la que puede iniciarse la transcripción del ADN unido operablemente cadena abajo (es decir, en la dirección 3' en la cadena de sentido del ADN de doble cadena). Dentro de la secuencia del promotor se encuentra un sitio de inicio de la transcripción (que se encuentra convenientemente, por ejemplo, mediante el mapeado con nucleasa S1), así como dominios de unión de proteína (secuencias de consenso) responsables de la unión de la ARN polimerasa.

Los promotores que pueden utilizarse para controlar la expresión génica en el contexto de la presente invención son, por ejemplo, los funcionales en las células no de vertebrado o en las células de vertebrado. Por ejemplo, para las células no de vertebrado, pueden utilizarse las secuencias reguladoras del gen de metalotioneína (Brinster *et al.*, Nature 296:39-42, 1982).

Preferentemente, el promotor inducible que se encuentra presente en el vector de la invención presenta una actividad de promotor en una célula de insecto, y más preferentemente en una célula de *Drosophila*. Es, por ejemplo, el promotor de metalotioneína de *Drosophila* (Lastowski-Perry *et al.*, J. Biol. Chem. 260:1527, 1985) que dirige un nivel elevado de transcripción del gen en presencia de metales, por ejemplo CuSO<sub>4</sub>. Alternativamente, puede utilizarse el promotor del gen actina 5C de *Drosophila*, que es un promotor constitutivo y no requiere la adición de un metal (B.J. Bond *et al.*, Mol. Cell. Biol. 6:2080, 1986). Entre los ejemplos de otros promotores de *Drosophila* conocidos se incluyen, por ejemplo, los promotores inducibles de choque térmico (Hsp70) y COPIA LTR. El promotor temprano del SV40 proporciona un nivel más bajo de expresión que el de metalotioneína de *Drosophila*.

Preferentemente, el promotor inducible que se encuentra presente en el vector de la invención presenta una actividad de promotor en una célula de *Drosophila melanogaster*, preferentemente en células S2 de *Drosophila*. Es, por ejemplo, el promotor de metalotioneína que se describe en detalle en Lastowski-Perry *et al.*, J. Biol. Chem. 260:1527, 1985.

Entre los promotores de la exposición adecuados para la expresión constitutiva en células de mamífero se incluyen, el promotor temprano inmediato del citomegalovirus (CMV), el promotor tardío mayor de adenovirus, el promotor fosfogliceroquinas (PGK) y el promotor timidina cinasa (TK) del virus herpes simplex (VSH)-1. Entre los promotores eucarióticos inducibles regulados por compuestos suministrados exógenamente se incluyen, aunque sin limitación, el promotor de metalotioneína (MT) inducible con cinc, el promotor del virus del tumor mamario de ratón (MMTV) inducible por dexametasona (Dex), el sistema del promotor de polimerasa de T7 (documento nº WO 98/10088), el promotor ecdisona de insecto, el promotor reprimible por tetraciclina, el promotor inducible por tetraciclina, el promotor inducible por RU486 y el promotor inducible por rapamicina.

Preferentemente, el promotor que se encuentra presente en el vector de la exposición presenta una actividad de promotor en una célula de mamífero, preferentemente en células HeLa. Es, por ejemplo, el promotor del SV 40.

Tal como se da a conocer en la presente memoria, se encuentra disponible un abanico de promotores de levadura para la expresión de proteínas en las células hospedadoras de levadura. Algunos, tales como ADH2 y SUC2, son inducible, mientras que otros, como GAPDH, presentan una expresión constitutiva. Entre otros promotores adecuados para la expresión en levaduras se incluyen los promotores TEF, PGK, MF alfa, CYC-1, GAL-1, GAL4, GAL10, PHO5, gliceraldehído-3-fosfato deshidrogenasa (GAP o GAPDH) y alcohol deshidrogenasa (ADH).

Tal como se da a conocer en la presente memoria, para la utilización en células vegetales, el promotor utilizado más comúnmente es el promotor 35S del virus del mosaico de la coliflor (CaMV) o su versión potenciada, aunque puede utilizarse uno de entre varios promotores alternativos, tales como el promotor híbrido *(ocs)3mas* o el promotor ubiquitina del maíz y de *A. Thaliana*. En contraste con dichos promotores constitutivos, el promotor  $\alpha$ -amilasa RAm3D del arroz es inducido por la privación de azúcar (Hellwig S. *et al.*, 2004).

Tal como se da a conocer en la presente memoria, entre los promotores adecuados para la expresión en una célula hospedadora de *E. coli* se incluyen, aunque sin limitación, el promotor pL del bacteriófago lambda, el promotor lac, y los promotores pTAC inducibles por TRP e IPTG.

Resulta preferido que la señal de secreción peptídica y el promotor inducible son funcionales en la misma célula hospedadora.

Más preferentemente, la señal de secreción peptídica y el promotor inducible son funcionales tanto en células S2 de *Drosophila* como en células de vertebrado.

El término "inducible" tal como se aplica a un promotor es bien conocido por el experto en la materia. En esencia, la expresión bajo el control de un promotor inducible resulta "activado" o incrementado en respuesta a la aplicación de un estímulo. La naturaleza del estímulo varía entre promotores. Algunos promotores inducibles provocan niveles bajos o indetectables de expresión (o una expresión nula) en ausencia del estímulo apropiado. Otros promotores inducibles causan una expresión constitutiva detectable en ausencia del estímulo. Sea cual sea el nivel de expresión en ausencia del estímulo, la expresión a partir de cualquier promotor inducible se encuentra incrementada en presencia del estímulo correcto.

Tras construir un vector apropiado y transfectarlo en la célula hospedadora seleccionada, preferentemente una línea celular de *Drosophila*, se induce la expresión de una proteína heteróloga mediante la adición de un agente inductor apropiado para el promotor inducible. Por ejemplo, el cadmio o el cobre son agentes inductores para el promotor Hsp70. Para los promotores constitutivos, tales como el promotor actina 5C, no se requiere ningún agente inductor para la expresión.

El enzima MGMT humano de la invención preferentemente se encuentra codificado por la secuencia del gen de MGMT humano de referencia NM\_002412.3, el gen ID 4255 (SEC ID nº 3) o por la secuencia optimizada SEC ID nº 68 (que comprende únicamente 50% de G/C). Sin embargo, puede utilizarse cualquier secuencia homóloga de la misma en el contexto de la invención, con la condición de que codifique un enzima MGMT funcional, o mutante del mismo, preferentemente la SEC ID nº 4 o la SEC ID nº 2. Las secuencias de ADN preferentes codificantes de dicho mutante de MGMT son las secuencias de ADN de SNAP SEC ID nº 1 o las secuencias de ADN SEC ID nº 47 o SEC ID nº 67 codificantes de la SEC ID nº 2 pero que presentan un contenido de G/C de 51%.

En otra realización de la exposición, el vector de nucleótidos de la invención codifica por lo menos un fragmento del enzima MGMT (por ejemplo un fragmento de SEC ID nº 4) o un fragmento de un homólogo del mismo (por ejemplo un fragmento del mutante de MGMT de secuencia SEC ID nº 2) que conserva la actividad biológica de

incremento de la expresión de la proteína de interés en un factor de por lo menos 0,5 veces el nivel obtenido con el enzima de longitud completa del que es un fragmento. A título de ejemplo, en el caso de que el nivel de producción sea de 100 mg/l con el enzima de longitud completa de SEC ID nº 4, se encuentran comprendidos dentro de la presente invención cualesquiera fragmentos de SEC ID nº 4 que presenta un nivel de producción de por lo menos 50 mg/l (bajo las mismas condiciones experimentales que para el enzima de longitud completa de SEC ID nº 4).

En otra realización de la invención, el vector de expresión de nucleótidos codifica adicionalmente por lo menos un sitio de corte peptídico que se encuentra situado preferentemente entre el enzima MGMT y la proteína recombinante de interés.

Un sitio de corte peptídico (también denominado "sitio de corte del péptido") es una secuencia de aminoácidos que resulta reconocida por como mínimo un enzima proteasa (por ejemplo una serina proteasa o una cisteína proteasa, entre otros). Un ejemplo de un sitio de corte peptídico es el sitio de corte de enterocinasa de SEC ID nº 62 (AspAspAspAspLys/Asp), por ejemplo codificado por la secuencia de ADN SEC ID nº 12. La enterocinasa es un enzima serina proteasa (EC 3.4.21.9) que es conocido que convierte el tripsinógeno inactivo en tripsina activa mediante el corte en el extremo C-terminal de la secuencia: Val--(Asp)<sub>4</sub>--Lys--Ile--Val~ (tripsinógeno) → Val--(Asp)<sub>4</sub>--Lys (hexapéptido) + Ile--Val~ (tripsina). La enterocinasa corta después de la lisina en el caso de que Lys se encuentre precedido por cuatro Asp y no seguido por un residuo de prolina.

Otro sitio de corte peptídico útil es el sitio de corte de la denominada "proteasa del TEV", que presenta la secuencia de aminoácidos SEC ID nº 53 o SEC ID nº 65 (Glu Asn Leu Tyr Phe Gln Gly o Ser), y que se encuentra, por ejemplo, codificado por la secuencia de ADN SEC ID nº 52 o SEC ID nº 66. La proteasa del TEV es el nombre común para el dominio catalítico de 27 kDa de la inclusión nuclear de una proteína codificada por el virus del grabado del tabaco. Se encuentra disponible comercialmente (Invitrogen).

El sitio de corte del precursor membranal prM del virus Dengue serotipo 1 (SEC ID nº 61) también puede utilizarse en el vector de la invención.

En otra realización, el vector de expresión de nucleótidos de la invención codifica además un marcaje, preferentemente situado en el extremo C-terminal de la proteína recombinante en el polipéptido de fusión de la invención (que comprende la señal peptídica, la proteína MGMT u homólogo de la misma, y la proteína recombinante).

En el contexto de la invención, un "marcaje" está destinado a facilitar la recuperación del polipéptido a partir del lisado en bruto de células hospedadoras y preferentemente se selecciona entre el grupo que comprende: proteínas fluorescentes, etiquetas poli-histidina (poli-his) o poli-histidina-glicina (poli-his-gly), etiquetas HA de la gripe, etiqueta c-myc, etiquetas de glucoproteína D (gD) del virus del herpes simplex, péptidos Flag, epítomos de alfa-tubulina o etiquetas peptídicas de proteína de gen 10 de T7. Sin embargo, podría resultar de utilidad cualquier otro marcaje. En una forma de realización preferida de la invención, los vectores comprenden el ADN codificante de una etiqueta hexa-histidina que presenta la SEC ID nº 14.

En otra forma de realización, el vector de expresión de nucleótidos de la invención codifica además una o más secuencias espaciadoras, situadas preferentemente entre el enzima MGMT y la proteína recombinante de interés y/o entre la proteína recombinante de interés y el marcaje.

En el contexto de la invención, una secuencia espaciadora es una secuencia de aminoácidos que comprende por lo menos tres aminoácidos, destinada a separar espacialmente dos polipéptidos unidos (uniendo seguidamente estos polipéptidos de manera indirecta). Dicho espaciador puede ser, por ejemplo, la secuencia de aminoácidos Glicina-Glicina-Glicina-Serina (GGGS, SEC ID nº 63) y la secuencia espaciadora de ADN codificante de la misma puede ser SEC ID nº 13. En el contexto de la presente invención, dicha secuencia de ADN se denomina en lo sucesivo "secuencia espaciadora de ADN" y se encuentra situada entre el ADN codificante de MGMT o el dominio catalítico del mismo y la secuencia de ADN recombinante, preferentemente cadena arriba de la secuencia de ADN codificante del sitio de corte peptídico.

El vector de expresión de nucleótidos que se da a conocer en la presente invención puede presentar la secuencia SEC ID nº 9, la secuencia SEC ID nº 10 o la secuencia SEC ID nº 64 (correspondiente a vectores vacíos sin gen recombinante de interés insertado en los sitios de clonación). En una forma de realización particular, el vector de la invención puede codificar:

- una señal peptídica BiP de insecto (que preferentemente es funcional en las células S2 de *Drosophila*) o una señal de tipo BiP tal como se ha definido anteriormente,
- una proteína MGMT de SEC ID nº 4 o una proteína SNAP de SEC ID nº 2,

- una proteína recombinante de interés,
- un sitio de corte peptídico de enterocinasa o un sitio de corte de pro-TEV tal como se ha definido anteriormente,
- un marcaje poli-histidina, y
- dos secuencias espaciadoras que presentan la secuencia de aminoácidos Glicina-Glicina-Glicina-Serina (GGGS, SEC ID nº 63).

En una forma de realización más preferente, el vector de expresión de la invención codifica:

- una señal peptídica de BiP de insecto de SEC ID nº 48,
- una proteína MGMT de SEC ID nº 4 o una proteína SNAP de SEC ID nº 2,
- una proteína recombinante de interés,
- un sitio de corte peptídico de enterocinasa de SEC ID nº 62,
- un marcaje poli-histidina, y
- dos secuencias espaciadoras que presentan la secuencia de aminoácidos Glicina-Glicina-Glicina-Serina (GGGS).

En otra forma de realización preferida, el vector de expresión de la invención codifica:

- una señal peptídica de tipo BiP de SEC ID nº 51,
- una proteína MGMT de SEC ID nº 4 o una proteína SNAP de SEC ID nº 2,
- una proteína recombinante de interés,
- un sitio de corte peptídico de pro-TEV de SEC ID nº 53,
- un marcaje poli-histidina, y
- dos secuencias espaciadoras que presentan la secuencia de aminoácidos Glicina-Glicina-Glicina-Serina (GGGS).

Dichos vectores pueden comprender, por ejemplo, la secuencia SEC ID nº 19 (en el caso de que la proteína de interés sea la nucleoproteína N del virus la fiebre del valle del Rift (FVR), SEC ID nº 20 (en el caso de que la proteína de interés sea la nucleoproteína N del virus del Nilo Occidental), SEC ID nº 21 o nº 57 o nº 72 o nº 74 (en el caso de que la proteína de interés sea IFN $\alpha$ ), SEC ID nº 77, nº 79 o nº 81 (en el caso de que la proteína de interés sea el antígeno del cáncer SSX2), SEC ID nº 55 (en el caso de que la proteína de interés sea granzima M), SEC ID nº 89 (en el caso de que la proteína de interés sea FasL), SEC ID nº 84 o nº 86 (en el caso de que la proteína de interés sea el antígeno del cáncer NERCMSL) o SEC ID nº 92 (en el caso de que la proteína de interés sea la contactina CNTN4).

En un segundo aspecto, la presente invención da a conocer además un vector para la expresión de proteínas recombinantes en células hospedadoras, que comprende una secuencia de nucleótidos codificante en un único marco de lectura abierta, de 5' a 3':

- a) una señal de secreción peptídica,
- b) una proteína MGMT de SEC ID nº 4 o una proteína SNAP de SEC ID nº 2,
- c) por lo menos un sitio de corte peptídico,
- d) un marcaje poli-histidina, y
- e) por lo menos una secuencia espaciadora.

En una forma de realización preferida, dicha señal de secreción peptídica es la señal peptídica de tipo BiP de SEC ID nº 50.

En una forma de realización más preferente, dicho vector comprende dos sitios de corte peptídico de pro-TEV de

SEC ID nº 52 y/o dos secuencias espaciadoras que presentan la secuencia de aminoácidos SEC ID nº 63.

En una forma de realización particularmente preferente, dicho vector comprende la secuencia SEC ID nº 59 o SEC ID nº 69, haciendo referencia a dichas secuencias en la presente solicitud como el casete DeSNAP universal "DeSNAP Univ" y el casete DeMGMT "DeMGMT Univ", respectivamente.

Estas secuencias "DeSNAP Univ" (SEC ID nº 59) y "DeMGMT Univ" (SEC ID nº 69) se consideran secuencias "universales" ya que pueden insertarse en cualquier tipo de vector destinado a la transfección de células hospedadoras con el fin de producir proteínas heterólogas, es decir vectores de vertebrado (tales como el vector pcDNA3 o el vector pCI-neo), así como vectores no de vertebrado (tales como pMT/BiP/V5-HisA, que resulta útil en el sistema DES; ver los ejemplos, posteriormente).

Son ejemplos de plásmido que comprende dichas secuencias universales, SEC ID nº 64 (pUC57 que comprende DeSNAP Univ) y SEC ID nº 71 (pUC57 con DeMGMT Univ).

Tras clonar en la presente invención la secuencia heteróloga de una proteína de interés, dicho vector puede transfectarse ventajosamente en células hospedadoras de vertebrado o de no vertebrado, en la presente invención células de insecto, de manera que se produzca la proteína de interés en cantidades elevadas.

En un tercer aspecto, la presente exposición se refiere además a una célula recombinante que es transfectada establemente por un plásmido que comprende cualquier vector de expresión dado a conocer en la presente memoria.

Por ejemplo, la presente exposición presenta como diana la célula recombinante, que resulta establemente transfectada por dicho vector DeSNAP Univ o DeMGMT Univ, es decir, por el vector de expresión que comprende una secuencia de nucleótidos codificante en un único marco de lectura abierta, de 5' a 3':

- a) una señal de secreción peptídica,
- b) una proteína MGMT de SEC ID nº 4 o una proteína SNAP de SEC ID nº 2,
- c) por lo menos un sitio de corte peptídico,
- d) un marcaje poli-histidina, y
- e) por lo menos una secuencia espaciadora,

siendo cada componente tal como se ha definido anteriormente.

Preferentemente comprende los plásmidos de SEC ID nº 64 (pUC57 que comprende DeSNAP Univ) o SEC ID nº 71 (pUC57 con DeMGMT Univ) o por lo menos la secuencia de nucleótidos SEC ID nº 59 (DeSNAP Univ) o SEC ID nº 69 (DeMGMT Univ).

La célula recombinante también puede ser una bacteria, una célula de insecto, más preferentemente células S2 de *Drosophila* o derivados de las mismas, células de mamífero, más preferentemente células CHO, células YB2/0, células COS, células HEK, células NIH3T3, células HeLa, células PER.C6 o derivados de las mismas, y células aviares tales como células EBx, preferentemente células EB66 o derivados de las mismas.

Preferentemente, en dicho aspecto de la exposición, dicha célula recombinante es una célula de *E. coli*.

Dicha célula recombinante se utiliza con el fin de amplificar y purificar los vectores de expresión de la invención, preferentemente aquellos que comprenden DeSNAP Univ de SEC ID nº 59 (tal como SEC ID nº 64) o DeMGMT Univ de SEC ID nº 69 (tal como SEC ID nº 71).

Por lo tanto, la presente exposición también presenta como diana la utilización de dicha célula recombinante para la producción de cualquier vector de expresión de la invención (siendo dichos vectores tal como se ha definido anteriormente).

Los vectores de expresión de nucleótidos de la invención y de la exposición pueden comprender además un gen codificante de un marcador de selección y/o una secuencia de terminador.

Los genes marcadores de selección que pueden incluirse en el constructo típicamente son los que confieren fenotipos seleccionables, tales como resistencia a antibióticos (por ejemplo blasticidina, ampicilina, canamicina, higromicina, puromicina y cloranfenicol).

En un cuarto aspecto, la presente exposición se refiere a un polipéptido de fusión que comprende una señal de secreción peptídica que es funcional en células hospedadoras, preferentemente en células no de vertebrado o de vertebrado, más preferentemente en células de insecto y el enzima 6-metilguanina-ADN-metiltransferasa (MGMT) (EC 2.1.1.63), mutante o dominio catalítico del mismo tal como se ha definido anteriormente.



Más exactamente, la presente invención se refiere a un polipéptido de fusión tal como se define en la reivindicación 13.

5 En el polipéptido de fusión de la exposición, dicho enzima MGMT preferentemente es la proteína de SEC ID nº 4 o un homólogo de la misma.

10 El polipéptido de fusión de la invención o exposición preferentemente comprende además una proteína recombinante de interés tal como se ha definido anteriormente, preferentemente situada en el extremo C-terminal del enzima MGMT o dominio catalítico del mismo, y/o un marcaje, tal como se ha definido anteriormente. Dicho marcaje preferentemente es un marcaje polihistidina y preferentemente se localiza en el extremo C-terminal de la proteína recombinante de interés.

15 El polipéptido de fusión de la invención puede ser la secuencia de aminoácidos de SEC ID nº 33 a nº 43, SEC ID nº 56 o SEC ID nº 58 (en el caso de que la proteína recombinante de interés sea GrM), SEC ID nº 73 o nº 75 (en el caso de que la proteína recombinante de interés sea IFN $\alpha$ ), SEC ID nº 78 o nº 80 o nº 82 (en el caso de que la proteína recombinante de interés sea el antígeno del cáncer SSX2), SEC ID nº 85 o nº 87 (en el caso de que la proteína recombinante de interés sea NERCMSL), SEC ID nº 90 (en el caso de que la proteína recombinante de interés sea FasL) o SEC ID nº 93 (en el caso de que la proteína recombinante de interés sea CNTN4).

20 Resulta interesante que las proteínas de fusión de la invención pueden almacenarse a 4°C durante varios meses sin degradación. Este efecto de estabilización in vitro durante el almacenamiento podría ser el resultado de las propiedades de andamiaje de la proteína MGMT y/o de la elevada concentración que se obtiene gracias a la presencia de la proteína MGMT (típicamente de por lo menos 40 mg/ml).

25 Más importante, la asociación con MGMT estabiliza las proteínas recombinantes durante el procedimiento de purificación de las proteínas secretadas. De esta manera, podría utilizarse para estabilizar proteínas recombinantes in vivo tras la administración en un sujeto que lo necesita. El acoplamiento con MGMT sería un medio para incrementar el periodo de vida de dichas proteínas recombinantes in vivo. Este efecto de estabilización in vivo se está investigando en la actualidad.

30 En un quinto aspecto, la presente exposición se refiere a una célula hospedadora recombinante de no vertebrado que comprende el vector de expresión de la exposición.

35 Más exactamente, la presente invención se refiere a una célula de insecto recombinante tal como se define en cualquiera de las reivindicaciones 10 a 12.

Las células de no vertebrado pueden ser cualesquiera células de las divisiones Insectos, Arácnidos, Crustáceos, Moluscos, Anélidos, Cirripedios, Radiados, Celenterados e Infusorios.

40 En el contexto de la invención, se utilizan células de no vertebrado y son células de insecto, tales como células de *Drosophila* o de *Mosquito*. Más preferentemente son células S2 de *Drosophila*.

45 Las células S2 de *Drosophila* han sido ampliamente descritas. Resultan especialmente adecuadas para la producción de alto rendimiento de proteínas debido a que pueden mantenerse en cultivos en suspensión a temperatura ambiente ( $24 \pm 1^\circ\text{C}$ ). El medio de cultivo era M<sub>3</sub> complementado con 5% a 10% (v/v) de suero de feto bovino (SFB) inactivado por calor. En la puesta en práctica preferida de la invención el medio de cultivo contiene SFB al 5%. Tras la inducción las células se cultivan en medio sin suero. En dicho medio, las células S2 pueden cultivarse en cultivos en suspensión, por ejemplo en matraces de agitación de 250 a 2.000 ml bajo agitación a 50-60 rpm. Las densidades celulares típicamente se mantienen entre  $10^6$  y  $10^7$  células por ml.

50 La presente invención también presenta como diana células S2 recombinantes de *Drosophila* que comprenden los vectores de expresión de la invención, comprendiendo dichos vectores de expresión preferentemente la secuencia de nucleótidos seleccionada de entre el grupo que consiste de:

55 – la secuencia plasmídica SEC ID nº 64 (pUC57 con DeSNAP Univ) o la secuencia de nucleótidos clonada en la célula que ha sido depositada según el Tratado de Budapest en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 9 de diciembre de 2011, con el número CNCM I-4581,

60 – el vector que comprende la SEC ID nº 19 o la secuencia de nucleótidos clonada en la célula que ha sido depositada según el Tratado de Budapest en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 19 de agosto de 2010, con el número CNCM I-4357,

65 – el vector de la invención que comprende SEC ID nº 22 o la secuencia de nucleótidos clonada en la célula

que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 27 de octubre de 2010 con el número CNCM I-4381,

- 5 – el vector de la invención que comprende SEC ID nº 21 o la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 27 de octubre de 2010 con el número CNCM I-4382,
- 10 – el vector de la invención que comprende SEC ID nº 9 o la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 29 de septiembre de 2010 con el número CNCM I-4368, y
- 15 – el vector de la invención que comprende la SEC ID nº 20 la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 29 de septiembre de 2010 con el número CNCM I-4369,
- 20 – el vector de la SEC ID nº 71,
- el vector de la invención que comprende la SEC ID nº 57 o nº 72 o nº 74 (en el que la proteína de interés es IFN $\alpha$ ), SEC ID nº 77, nº 79 o nº 81 (en el caso de que la proteína de interés sea el antígeno del cáncer SSX2), SEC ID nº 55 (en el caso de que la proteína de interés sea granzima M), SEC ID nº 89 (en el caso de que la proteína de interés sea FasL), SEC ID nº 84 o nº 86 (en el caso de que la proteína de interés sea el antígeno del cáncer NERCMSL) o SEC ID nº 92 (en el caso de que la proteína de interés sea la contactina CNTN4) o SEC ID nº 96 (en el caso de que la proteína de interés sea hSULF-2 <sup>$\Delta$ TMD</sup>).
- 25 Las células S2 transfectadas establemente de la invención también pueden seleccionarse de entre el grupo que consiste de:
  - 30 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 19 de agosto de 2010, con el número CNCM I-4357,
  - 35 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 27 de octubre de 2010, con el número CNCM I-4381,
  - 40 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 27 de octubre de 2010, con el número CNCM I-4382,
  - 45 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 29 de septiembre de 2010, con el número CNCM I-4368,
  - 50 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 29 de septiembre de 2010, con el número CNCM I-4369,
  - 55 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 5 de diciembre de 2011, con el número CNCM I-4565,
  - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 5 de diciembre de 2011, con el número CNCM I-4566,
  - 60 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 5 de diciembre de 2011, con el número CNCM I-4567,
  - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 5 de diciembre de 2011, con el número CNCM I-4568,
  - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut

- Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 5 de diciembre de 2011, con el número CNCM I-4569,
- 5 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 5 de diciembre de 2011, con el número CNCM I-4570,
- 10 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 5 de diciembre de 2011, con el número CNCM I-4571,
- 15 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 5 de diciembre de 2011, con el número CNCM I-4572,
- 20 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 8 de diciembre de 2011, con el número CNCM I-4576,
- 25 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 8 de diciembre de 2011, con el número CNCM I-4577,
- 30 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 8 de diciembre de 2011, con el número CNCM I-4578,
- 35 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 8 de diciembre de 2011, con el número CNCM I-4579,
- 40 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 8 de diciembre de 2011, con el número CNCM I-4580,
- 45 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 9 de diciembre de 2011, con el número CNCM I-4582,
- 50 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 9 de diciembre de 2011, con el número CNCM I-4583,
- 55 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 9 de diciembre de 2011, con el número CNCM I-4584,
- 60 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 9 de diciembre de 2011, con el número CNCM I-4585, y
- 65 – la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, 25, rue du Docteur Roux, 75724 Paris Cedex 15, Francia, el 9 de diciembre de 2011, con el número CNCM I-4586.
- La célula recombinante depositada con el número CNCM I-4357 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende el vector plasmídico de SEC ID nº 19 (pMT/BiP/SNAP-FVR.N/His<sub>etiqueta</sub>), en el que FVR.N es el antígeno N del virus de la fiebre del valle del Rift (FVR) (ver Brehin *et al.*, *Virology* 371:185, 2008).
- La célula recombinante depositada con el número CNCM I-4381 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/V5-His<sub>etiqueta</sub> en el que la SEC ID nº 22 (SNAP/WN.EDIII) ha sido insertado después de la secuencia de BiP, en la que NO.EDIII es el dominio III de la glucoproteína E del virus del Nilo Occidental.
- La célula recombinante depositada con el número CNCM I-4382 es la línea celular S2 estable de macrófagos de

*Drosophila* que comprende el vector plasmídico pMT/V5-His<sub>etiqueta</sub> en el que ha sido insertada la SEC ID nº 21 (BiP/SNAP/IFN $\alpha$ 1). IFN $\alpha$ 1 es el interferón alfa 1 humano de SEC ID nº 32 (Mokkim *et al.*, Protein expression purif. 63:140, 2009).

- 5 La célula recombinante depositada con el número CNCM I-4369 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/V5-His<sub>etiqueta</sub> en el que la SEC ID nº 20 (NO.sE/SNAP/His<sub>etiqueta</sub>), en el que NO.sE es la forma soluble de la proteína de cubierta E del virus del Nilo Occidental.
- 10 La célula recombinante depositada con el número CNCM I-4369 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/V5-His<sub>etiqueta</sub> en el que la SEC ID nº 20 (NO.sE/SNAP/His<sub>etiqueta</sub>), en el que NO.sE es la forma soluble de la proteína de cubierta E del virus del Nilo Occidental.
- 15 La célula recombinante depositada con el número CNCM I-4565 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+DV1.EDIII/His<sub>etiqueta</sub> en el que DV1.EDIII codifica la proteína EDIII del virus Dengue 1 y presenta la secuencia SEC ID nº 27.
- 20 La célula recombinante depositada con el número CNCM I-4566 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+DV2.EDIII/His<sub>etiqueta</sub> en el que DV2.EDIII codifica la proteína EDIII del virus Dengue 2 y presenta la secuencia SEC ID nº 28.
- 25 La célula recombinante depositada con el número CNCM I-4567 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+DV3.EDIII/His<sub>etiqueta</sub> en el que DV3.EDIII codifica la proteína EDIII del virus Dengue 3 y presenta la secuencia SEC ID nº 29.
- 30 La célula recombinante depositada con el número CNCM I-4568 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+DV4.EDIII/His<sub>etiqueta</sub> en el que DV4.EDIII codifica la proteína EDIII del virus Dengue 4 y presenta la secuencia SEC ID nº 30.
- 35 La célula recombinante depositada con el número CNCM I-4569 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+YF.EDIII/His<sub>etiqueta</sub> en el que YF.EDIII codifica la proteína EDIII del virus de la fiebre amarilla y presenta la secuencia SEC ID nº 31.
- 40 La célula recombinante depositada con el número CNCM I-4570 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+JE.EDIII/His<sub>etiqueta</sub> en el que JE.EDIII codifica la proteína EDIII del virus de la encefalitis japonesa y presenta la secuencia SEC ID nº 25.
- 45 La célula recombinante depositada con el número CNCM I-4571 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+USU.EDIII/His<sub>etiqueta</sub> en el que USU.EDIII codifica la proteína EDIII del virus Usutu y presenta la secuencia SEC ID nº 24.
- 50 La célula recombinante depositada con el número CNCM I-4572 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+TBE.EDIII/His<sub>etiqueta</sub> en el que TBE.EDIII codifica la proteína EDIII del virus de la encefalitis transmitida por garrapatas y presenta la secuencia SEC ID nº 26.
- 55 La célula recombinante depositada con el número CNCM I-4576 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+MVE.EDIII/His<sub>etiqueta</sub> en el que EVM.EDIII codifica la proteína EDIII del virus de la encefalitis del valle de Murray.
- 60 La célula recombinante depositada con el número CNCM I-4577 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+Rocio.EDIII/His<sub>etiqueta</sub> en el que Rocio.EDIII codifica la proteína EDIII del virus Rocio.
- 65 La célula recombinante depositada con el número CNCM I-4578 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+SLE.EDIII/His<sub>etiqueta</sub> en el que ESL.EDIII codifica la proteína EDIII del virus de la encefalitis de Saint Louis.
- 60 La célula recombinante depositada con el número CNCM I-4579 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+WSL.EDIII/His<sub>etiqueta</sub> en el que WSL.EDIII codifica la proteína EDIII del virus Wesselbron.
- 65 La célula recombinante depositada con el número CNCM I-4580 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+Zika.EDIII/His<sub>etiqueta</sub> en el que Zika.EDIII codifica la proteína EDIII del virus Zika.

La célula recombinante depositada con el número CNCM I-4583 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+SSX2/His<sub>etiqueta</sub> en el que SSX2 es de SEC ID nº 76.

5 La célula recombinante depositada con el número CNCM I-4584 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+NERCMSL/His<sub>etiqueta</sub> en el que NERCMSL es de SEC ID nº 83.

10 La célula recombinante depositada con el número CNCM I-4585 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/SNAP+GrM/His<sub>etiqueta</sub> en el que GrM es de SEC ID nº 54.

15 La célula recombinante depositada con el número CNCM I-4586 es la línea celular S2 estable de macrófagos de *Drosophila* que comprende un vector plasmídico pMT/BiP/ProTEV/His<sub>etiqueta</sub> en el que proTEV es de SEC ID nº 52.

20 En un sexto aspecto, la presente exposición también presenta como diana una célula hospedadora recombinante de vertebrado que se encuentra establemente transfectada con el vector de expresión de la exposición.

Preferentemente, dicha célula recombinante de vertebrado es una célula de mamífero, preferentemente una célula CHO, YB2/O, COS, HEK, NIH3T3, HeLa o derivados de las mismas. Más preferentemente, en este caso, el vector de la expresión de la invención comprende la SEC ID nº 57 o nº 72 o nº 74 (en el caso de que la proteína de interés sea IFN $\alpha$ ), SEC ID nº 77, nº 79 o nº 81 (en el caso de que la proteína de interés sea el antígeno del cáncer SSX2), SEC ID nº 55 (en el caso de que la proteína de interés sea granzima M), SEC ID nº 89 (en el caso de que la proteína de interés sea FasL), SEC ID nº 84 o nº 86 (en el caso de que la proteína de interés sea el antígeno del cáncer NERCMSL) o SEC ID nº 92 (en el caso de que la proteína de interés sea la contactina CNTN4) o SEC ID nº 96 (en el caso de que la proteína de interés sea hSULF-2<sup>ΔTMD</sup>). En un séptimo aspecto, la presente invención se refiere a un método de incremento de la expresión de la proteína o proteínas recombinantes que comprende coexpresar dicha proteína o proteínas con una señal de secreción peptídica, conjuntamente con el enzima 6-metilguanina-ADN-metiltransferasa (MGMT) (EC 2.1.1.63), un mutante o un dominio catalítico del mismo. Dicha coexpresión se lleva a cabo preferentemente en células de no vertebrado y, más preferentemente, células de insecto.

35 Más preferentemente, en dicho método, el enzima MGMT es la proteína de SEC ID nº 4 o un homólogo de la misma, por ejemplo la proteína SNAP de SEC ID nº 2 o un homólogo de la misma.

40 En el contexto de la invención, la expresión "que potencia la expresión" de una proteína heteróloga se refiere a que la expresión de dicha proteína en el sobrenadante de las células recombinantes o dentro de las células mismas mejora en un factor de por lo menos 2 veces, preferentemente de 5 veces, más preferentemente de 10 veces, y todavía más preferentemente de 20 veces, en comparación con la expresión y/o la secreción de dicha proteína obtenida con un vector recombinante de la técnica anterior, es decir, que no coexpresa la proteína con una proteína MGMT o una proteína SNAP. En una forma de realización preferida, la expresión "que potencia la expresión" se refiere a que resulta posible recuperar del sobrenadante de las células hospedadoras que han sido transfectadas con el vector de la invención por lo menos 40 mg/l, preferentemente por lo menos 50 mg/l, más preferentemente por lo menos 60 mg/l de una proteína de interés.

50 El término "coexpresantes" se refiere a que las secuencias de ADN codificantes de i) la proteína recombinante, ii) el enzima MGMT, o un mutante del mismo, e iii) la señal de secreción peptídica, se encuentran operablemente ligadas y reguladas por la misma secuencia de control de la expresión (es decir, secuencias de control transcripcional y traduccional). La traducción de las secuencias de ADN codificantes de la señal de secreción peptídica, la proteína heteróloga de interés y el enzima MGMT conduce por lo tanto a la formación de un polipéptido de fusión, en el que las proteínas pueden separarse con una secuencia espaciadora y/o un sitio de corte de enzima tal como se ha definido anteriormente.

55 La "señal de secreción peptídica" del polipéptido de fusión de la exposición es una señal de secreción que preferentemente es funcional en células no de vertebrado o en células de vertebrado, o ambas, y más preferentemente en células de insecto, todavía más preferentemente en células S2 de *Drosophila*. Más exactamente, la "señal de secreción peptídica" del polipéptido de fusión de la invención es una señal de secreción que es funcional en las células de insecto, todavía más preferentemente en células S2 de *Drosophila*.

60 Son ejemplos de señales de secreción peptídica que son funcionales en las células de insecto: ssBiP de insecto de SEC ID nº 48, la señal de tipo BiP de SEC ID nº 51 y cualquier señal peptídica presente en un arbovirus, por ejemplo la proteína E de cubierta del virus del Nilo Occidental (SEC ID nº 15).

65 Un ejemplo de una señal de secreción peptídica que es funcional en células tanto de vertebrado como de no

vertebrado es la señal de tipo BiP de SEC ID nº 51.

5 En un octavo aspecto, la presente exposición se refiere a un método para mejorar la producción de una proteína recombinante de interés o para producir las proteínas recombinantes en cultivo celular, comprendiendo la utilización del vector tal como se ha indicado anteriormente, o las células hospedadoras recombinantes tal como se ha indicado anteriormente.

10 Más exactamente, dicho método para mejorar la producción de una proteína recombinante de interés o para producir proteínas recombinantes en cultivo celular comprende las etapas de:

- 15 a) proporcionar el vector de expresión de nucleótidos tal como se ha indicado anteriormente, codificante de dicha proteína de interés,
- b) introducir dicho vector de expresión en células hospedadoras, preferentemente células de no vertebrado o de vertebrado,
- c) proporcionar la expresión del nucleótido introducido en dichas células hospedadoras para producir dicha proteína recombinante de interés.

20 Preferentemente, dichas células hospedadoras de no vertebrado son células de insecto, por ejemplo células S2 de *Drosophila*.

25 Preferentemente, dichas células hospedadoras de vertebrado son células de mamífero, por ejemplo células CHO, YB2/O, COS, HEK, NIH3T3, HeLa o derivados de las mismas.

Mediante la utilización de dicho método se expresa la proteína recombinante de interés a razón de por lo menos 40 mg/ml del sobrenadante de cultivo celular recuperado o superior.

30 La utilización de la línea celular S2 de *Drosophila* que secreta el producto génico directamente al medio es una forma de realización preferida de la presente invención (la secreción directa al medio permite la utilización de un sistema de purificación eficiente en una etapa).

35 En una novena forma de realización, la presente exposición se refiere a la utilización del enzima 6-metilguanina-ADN-metiltransferasa (MGMT) (EC 2.1.1.63), mutante o dominio catalítico del mismo, para potenciar el nivel de producción de una o más proteínas recombinantes, preferentemente en células hospedadoras de no vertebrado y/o de vertebrado, más preferentemente en células de insecto o células de mamífero, infectadas con vectores replicativos o defectuosos.

40 Más exactamente, la presente invención se refiere a la utilización tal como se define en cualquiera de las reivindicaciones 1 a 6.

45 El enzima MGMT puede ser la MGMT humano (de referencia NP\_002403.2) de secuencia SEC ID nº 4, la MGMT de ratón identificada como NP\_032624.1 (SEC ID nº 45), la MGMT de rata identificada como NP\_036993.1 (SEC ID nº 46), una secuencia homóloga de la misma, o subfragmentos de la misma.

Preferentemente, el enzima MGMT mutante es la proteína SNAP de SEC ID nº 2 o es homóloga de la misma, es decir, es por lo menos idéntica a un nivel superior al 80%, preferentemente 85%, más preferentemente 90% respecto a la proteína SNAP de secuencia SEC ID nº 2.

50 Dichas células hospedadoras de no vertebrado preferentemente son células de insecto, por ejemplo células S2 de *Drosophila*.

55 En una forma de realización preferida, la presente exposición se refiere a la utilización del enzima 6-metilguanina-ADN-metiltransferasa (MGMT) (EC 2.1.1.63), mutante o dominio catalítico del mismo, para potenciar el nivel de producción de una o más proteínas recombinantes en células de vertebrado, por ejemplo en células de mamífero, infectadas con vectores replicativos o defectuosos.

Dichas células de vertebrado preferentemente son células EBX, CHO, YB2/O, COS, HEK, NIH3T3 o derivados de las mismas.

60 Además, la presente exposición se refiere a la utilización de una secuencia de ADN codificante de un enzima MGMT, mutante o dominio catalítico del mismo, para mejorar el nivel de producción de una o más proteínas de interés en células recombinantes.

65 La presente exposición se refiere además a la utilización de una secuencia de ADN codificante de un enzima MGMT, mutante o dominio catalítico del mismo, para i) estabilizar una o más proteínas recombinantes de interés

in vitro e in vivo, y, de esta manera, ii) potenciar el periodo de vida in vitro e in vivo.

Dicha secuencia de ADN es, por ejemplo, la secuencia del gen de MGMT humano NM\_002412.3, gen ID 4255 (SEC ID nº 3) o cualquier secuencia homóloga de la misma codificante de un enzima MGMT funcional, un mutante o un dominio catalítico del mismo (preferentemente SEC ID nº 1, SEC ID nº 47, SEC ID nº 67 ó SEC ID nº 68).

En particular, la presente exposición se refiere a la utilización del enzima 6-metilguanina-ADN-metiltransferasa (MGMT, EC 2.1.1.63), mutantes o dominio catalítico del mismo como polipéptido protector fusionado o unido a proteínas recombinantes para mejorar la semivida de la proteína recombinante en medio de almacenamiento, en plasma o en tampón, para mejorar la semivida de la proteína recombinante utilizada como medicamento o vacuna, o para mejorar la semivida de la proteína recombinante utilizada en kits diagnósticos.

Más exactamente, la presente invención se refiere a la utilización según se define en la reivindicación 14.

En el contexto de la invención, la expresión "mejorar el nivel de producción" o "potenciar el nivel de producción" de una proteína heteróloga se refiere a que la expresión de dicha proteína en el sobrenadante de dichas células o dentro de las células mejora en un factor de por lo menos 2 veces, preferentemente de 5 veces, más preferentemente de 10 veces, y todavía más preferentemente de 20 veces, en comparación con la expresión y/o la secreción de dicha proteína obtenida con un vector recombinante de la técnica anterior, es decir, que no comprende el vector de la invención. En una forma de realización preferida, la expresión "que mejora la producción" se refiere a que resulta posible recuperar del sobrenadante de las células hospedadoras que han sido transfectadas con el vector de la invención por lo menos 40 mg/l, preferentemente por lo menos 50 mg/l, más preferentemente por lo menos 60 mg/l de una proteína de interés.

En una forma de realización preferida, dicha proteína recombinante se selecciona de entre: insulina, IFN, SSX2, granzima M, FasL, mesotelina (NERMCSL), endosulfatasa (hSUFL) o contactinas.

En una forma de realización particular, dicha utilización es un método para la producción de una proteína recombinante de interés, comprendiendo el método las etapas de:

- (a) obtener una secuencia de ADN heterólogo codificante de una proteína recombinante de interés,
- (b) insertar dicha secuencia de ADN heterólogo en el vector de expresión de nucleótidos de la invención, presentando dicho vector, por ejemplo, la secuencia de ADN SEC ID nº 9, SEC ID nº 10, SEC ID nº 64 o SEC ID nº 71,
- (c) transfectar una célula hospedadora (preferentemente una célula de insecto o una célula de mamífero) con el polinucleótido obtenido en la etapa (b),
- (d) proporcionar la expresión de dicho polinucleótido obtenido en la etapa (c) para producir la proteína de interés,
- (e) opcionalmente, cortar el polipéptido MGMT,
- (f) recuperar la proteína de interés,
- (g) opcionalmente, recuperar la proteína de interés.

Para llevar a cabo las diferentes etapas de dicho método pueden utilizarse técnicas convencionales de biología molecular, microbiología y de ADN recombinante comprendidas dentro de conocimiento del experto en la materia. Dichas técnicas se explican en detalle en la literatura. Ver, por ejemplo, Sambrook, Fitch & Maniatis, *Molecular Cloning: A Laboratory Manual*, segunda edición, 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (en la presente memoria, "Sambrook *et al.*, 1989"); *DNA Cloning: A Practical Approach*, volúmenes I y II (D. N. Glover ed. 1985); *Oligonucleotide Synthesis* (M. J. Gait ed. 1984); *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins, eds. 1984); *Animal Cell Culture* (R. I. Freshney, ed. 1986); *Immobilized Cells and Enzymes* (IRL Press, 1986); B. E. Perbal, *A Practical Guide to Molecular Cloning*, 1984; F. M. Ausubel *et al.* (editores), *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc., 1994.

El término "transfección" se refiere a la introducción de un ácido nucleico foráneo en una célula hospedadora eucariótica de manera que la célula hospedadora exprese el gen o secuencia introducido para producir una sustancia deseada, en la presente invención una proteína codificada por el gen o secuencia introducido. Una célula hospedadora que recibe y expresa el ADN o ARN introducido ha sido "transfectada" y es un "transfectante" o un "clon". El ADN o ARN introducido en una célula hospedadora puede proceder de cualquier fuente, incluyendo células del mismo género o especie que la célula o células hospedadoras de un género o especie diferente.

En el contexto de la invención, la transfección de las células hospedadoras con los polinucleótidos puede llevarse a cabo mediante un método clásico de la técnica, por ejemplo mediante transfección, infección o electroporación. En otra forma de realización, el vector de la invención puede introducirse in vivo mediante lipofección (en forma de ADN desnudo) o con otros agentes facilitadores de la transfección (péptidos, polímeros, etc.). Los lípidos catiónicos sintéticos pueden utilizarse para preparar liposomas para la transfección in vivo de un gen codificante de un marcador (Felgner *et al.*, Proc. Natl. Acad. Sci. U.S.A. 84:7413-7417, 1987). Los compuestos y composiciones lipídicas útiles para la transferencia de ácidos nucleicos se describen en los documentos nº WO 95/18863 y nº WO 96/17823 y en la patente US nº 5.459.127. Los lípidos pueden acoplarse químicamente a otras moléculas con el propósito de la administración dirigida (ver Mackey *et al.*, Proc. Natl. Acad. Sci. U.S.A. 85:8027-8031, 1988). Los péptidos dirigidos, tales como hormonas o neurotransmisores, y proteínas tales como anticuerpos o las moléculas no peptídicas pueden acoplarse químicamente a liposomas. Otras moléculas también resultan útiles para facilitar la transfección de un ácido nucleico in vivo, tal como los oligopéptidos catiónicos (ver el documento nº WO 95/21931), péptidos derivados de proteínas de unión a ADN (ver el documento nº WO 96/25508) o polímeros catiónicos (ver el documento nº WO 95/21931). También resulta posible introducir el vector in vivo en forma de un plásmido de ADN desnudo. Los vectores de ADN desnudo para la terapia génica pueden introducirse en las células hospedadoras deseadas mediante métodos conocidos de la técnica, tales como la electroporación, la microinyección, la fusión celular, DEAE-dextrano, la precipitación con fosfato de calcio, la utilización de una pistola génica o la utilización de un transportador de vector de ADN (ver Wu *et al.*, J. Biol. Chem. 267:963-967, 1992; Wu y Wu, J. Biol. Chem. 263:14621-14624, 1988; Williams *et al.*, Proc. Natl. Acad. Sci. U.S.A. 88:2726-2730, 1991).

La expresión "que proporciona la expresión" de un polinucleótido en la presente memoria se refiere a que el estímulo de las secuencias reguladoras presentes en el vector (por ejemplo el estímulo que activa el promotor inducible) y todos los componentes necesarios se encuentran presentes en cantidad suficiente para que se produzca la traducción del polinucleótido.

En caso necesario, el corte del polipéptido MGMT/SNAP de la proteína de fusión producido se obtiene mediante la adición de una proteasa que presenta un sitio de corte definido en el sobrenadante o en el interior de las células recombinantes. Por ejemplo, el corte del sitio de corte de enterocinasa DDDK/D se obtiene mediante la adición de un enzima enterocinasa al sobrenadante de las células recombinantes. Alternativamente, el polipéptido MGMT/SNAP puede mantenerse de manera que se incremente el periodo de vida de las proteínas recombinantes.

Además, el experto en la materia apreciará que puede detectarse una proteína o polipéptido expresado o secretado al medio de cultivo utilizado para mantener o cultivar las células hospedadoras presentes. El medio de cultivo puede separarse de las células hospedadoras mediante procedimientos conocidos, tales como la centrifugación o la filtración. A continuación, la proteína o polipéptido puede detectarse en el medio de cultivo sin células aprovechando las propiedades conocidas características de la proteína o polipéptido. Entre dichas propiedades pueden incluirse las propiedades inmunológicas, enzimáticas o físicas diferenciadas de la proteína o polipéptido. Por ejemplo, en el caso de que la proteína o polipéptido presente una actividad enzimática única, puede llevarse a cabo un ensayo para dicha actividad en el medio de cultivo utilizado por las células hospedadoras. Además, en el caso de que se encuentren disponibles anticuerpos reactivos con una proteína o polipéptido dado, dichos anticuerpos pueden ser utilizados para detectar la proteína o polipéptido en cualquier ensayo inmunológico conocido (por ejemplo tal como en Harlowe *et al.*, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1988).

La recuperación de la proteína de interés está mediada por medios bien conocidos de la técnica, entre ellos, aunque sin limitación, la electroforesis preparativa de disco en gel, el enfoque isoeléctrico, la HPLC, la HPLC de fase inversa, la filtración en gel, el intercambio iónico y la cromatografía de repartición, la precipitación y la cromatografía de precipitación en concentración salina, la extracción y la distribución contracorriente, y similares. Debido a que resulta preferible producir la proteína de interés en el sistema recombinante de la invención unida a un marcaje, dicho marcaje facilita la recuperación del polipéptido a partir del lisado en bruto de la célula hospedadora mediante cromatografía en una matriz de fase sólida apropiada. Alternativamente, pueden utilizarse como reactivos de recuperación anticuerpos producidos contra la proteína o contra péptidos derivados de la misma.

Puede llevarse a cabo una etapa adicional (g) de purificación aunque resulta interesante que no resulta necesaria.

Un material purificado puede contener menos de aproximadamente 50%, preferentemente menos de aproximadamente 75% y todavía más preferentemente menos de aproximadamente 90% de los componentes celulares con los que se encontraba originalmente asociado. La expresión "sustancialmente puro" indica el grado más alto de pureza que puede conseguirse utilizando técnicas de purificación convencionales conocidas de la técnica.



En una forma de realización de la invención, los métodos de la invención permiten obtener por lo menos 40 mg/l, preferentemente por lo menos 50 mg/l, más preferentemente por lo menos 60 mg/l de la proteína sustancialmente pura de interés en el sobrenadante de cultivo celular recuperado.

5 Las proteínas recombinantes de interés y las proteínas de fusión de la invención (es decir, las proteínas recombinantes acopladas con el polipéptido MGMT/SNAP, que son más estables que las proteínas recombinantes solas) pueden resultar útiles en una diversidad de productos. Por ejemplo, dichas proteínas recombinantes y/o de fusión pueden utilizarse en composiciones farmacéuticas, por ejemplo para el tratamiento de infecciones víricas.

10 En una forma de realización preferida, dicha proteína recombinante se selecciona de entre: insulina, IFN, FasL, granzima M, SSX2, mesotelina (NERMCSL), endosulfatasa (hSUFL) o contactinas.

15 En otra forma de realización, la presente exposición proporciona partículas víricas infecciosas que comprenden los vectores de ácidos nucleicos anteriormente indicados: Típicamente, dichas partículas víricas se producen mediante un procedimiento que comprende las etapas de:

(a) introducir el vector vírico de la invención en una línea celular adecuada,

20 (b) cultivar dicha línea celular bajo condiciones adecuadas de manera que se permita la producción de dicha partícula vírica infecciosa,

(c) recuperar la partícula vírica infecciosa producida a partir del cultivo de dicha línea celular, y

25 (d) opcionalmente purificar dicha partícula vírica infecciosa recuperada.

En el caso de que el vector vírico sea defectuoso, las partículas infecciosas habitualmente se producen en una línea celular complementaria o mediante la utilización de un virus ayudante, que suministra *in trans* los genes víricos no funcionales. Por ejemplo, entre las líneas celulares adecuadas para complementar los vectores adenovíricos con delección de E1 se incluyen las células 293, así como las células PER-C6. Las partículas víricas infecciosas pueden recuperarse a partir del sobrenadante de cultivo o de las células después de la lisis. Pueden purificarse adicionalmente según técnicas estándares (cromatografía, ultracentrifugación en un gradiente de cloruro de cesio tal como se indica en, por ejemplo, los documentos nº WO96/27677, nº WO98/00524, nº WO98/22588, nº WO98/26048, nº WO00/40702, nº EP 1 016 700 y nº WO00/50573).

35 De esta manera, la presente exposición se refiere a composiciones farmacéuticas que comprenden el vector de expresión, las proteínas recombinantes, las proteínas de fusión, las células hospedadoras o las partículas víricas de la invención, o cualquier combinación de los mismos. Dichas composiciones farmacéuticas comprenden una cantidad terapéutica del vector, partículas, células o proteínas obtenidas mediante el método indicado en la presente memoria mezclado con un portador farmacéuticamente aceptable.

40 La composición puede administrarse sistemáticamente por vía parenteral, intravenosa o subcutánea. En el caso de que se administre sistemáticamente, la composición terapéutica para la utilización en la presente invención se encuentra en forma de una solución de proteína libre de pirógenos parenteralmente aceptable. La preparación de dicha solución de proteína parenteralmente aceptable, considerando debidamente el pH, la isotonicidad, la estabilidad y similares, se encuentra dentro de los conocimientos del experto en la materia.

50 El régimen de administración será determinado por el médico responsable a partir de la consideración de diversos factores que modifican la acción de los fármacos, por ejemplo, la condición, peso corporal y dieta del paciente, la gravedad de la infección, el tiempo de administración y otros factores clínicos. El portador farmacéutico y otros componentes de una composición farmacéutica son seleccionados por el experto en la materia.

55 Adicionalmente, las proteínas de fusión y recombinantes de la presente invención pueden utilizarse como componentes de vacunas para la inoculación de sujetos humanos contra la infección vírica, por ejemplo. Estas proteínas pueden utilizarse solas o con otras proteínas recombinantes o agentes vacunales terapéuticos. Los componentes de dicha vacuna serán determinados por un experto en la materia.

60 La presente exposición comprende además la utilización de las proteínas de fusión, los vectores de expresión, las partículas víricas infecciosas, las células hospedadoras o las composiciones farmacéuticas de la invención para la preparación de un medicamento, en particular una vacuna. Más exactamente, la presente invención se refiere a un polipéptido de fusión o a un homólogo del mismo tal como se define en la reivindicación 15.

65 La presente exposición proporciona además un método para el tratamiento o la prevención de un organismo humano o animal, que comprende administrar en dicho organismo una cantidad terapéuticamente eficaz de las proteínas de fusión, los vectores de fusión, las partículas víricas infecciosas, las células hospedadoras o las

composiciones de la invención.

Finalmente, las proteínas de la presente invención, y especialmente la fusión de SNAP-proteína de interés, pueden resultar útiles como agentes diagnósticos para la detección de la presencia de cáncer, infección vírica o anticuerpos contra proteínas víricas en líquidos biológicos, tales como sangre, suero, saliva y similares. Estas proteínas también pueden utilizarse en métodos para identificar y/o aislar proteínas víricas en líquidos y tejidos biológicos. De esta manera, las proteínas pueden ser componentes en kits para llevar a cabo dichos métodos.

De esta manera, en otro aspecto, la presente exposición se refiere además a la utilización de proteínas recombinantes o proteínas recombinantes etiquetadas con MGMT o SNAP de microorganismos patógenos o no patógenos obtenidos mediante cualquier método descrito en la presente memoria para la identificación de la presencia de dichos microorganismos patógenos o no patógenos en una muestra biológica. En una forma de realización preferida, dicho microorganismo patógeno es un virus, y la proteína etiquetada con MGMT o SNAP es una proteína vírica, tal como EDIII de los virus Chikungunya, Dengue, de la encefalitis japonesa (EJ), de la encefalitis transmitida por garrapatas (ETG), de la fiebre amarilla (FA), Usutu (USU) o del Nilo Occidental, o la nucleoproteína N de los virus de la fiebre del valle del Rift o del virus Toscana.

En el contexto de la exposición, dicha muestra biológica se refiere a una muestra de sangre, una muestra de orina o cualquier muestra biológica que posiblemente se encuentre infectada por virus.

## Ejemplos

### 1. Construcción de uno o más plásmidos

1.1. Se utilizó el plásmido pMT/BiP/V5-His A. Contiene 3.642 nucleótidos y los elementos siguientes:

- promotor metalotioneína: bases 412 a 778
- Inicio de transcripción: base 778
- sitio de cebador directo MT: bases 814 a 831
- Secuencia de señal de BiP: bases 851 a 904 (SEC ID nº 11).
- Sitio de clonación múltiple: bases 906 a 999
- etiqueta epítipo V5: bases 1.003 a 1.044
- región polihistidina: bases 1.054 a 1.074
- sitio de cebador inverso BGH: bases 1.094 a 1.111
- señal de poliadenilación tardía del SV40: bases 1.267 a 1.272
- origen de pUC: bases 1.765 a 2.438 (cadena complementaria)
- promotor *bla*: bases 3.444 a 3.542 (cadena complementaria)
- ORF de gen de resistencia a la ampicilina (*bla*) bases 2.583 a 3.443 (cadena complementaria)

El vector pUC57 Amp también puede utilizarse para los fines de la invención. Dicho vector comprende:

- el sitio de clonación única *EcoRI*
- el promotor metalotioneína,
- la región no codificante 5' del ARN genómico del virus del Nilo Occidental cepa IS-98-ST1,
- un codón de inicio de traducción (ATG),
- el péptido de señal de la proteína E de cubierta del virus del Nilo Occidental cepa IS-98-ST1 (SEC ID nº 15),
- La región no codificante 3' del ARN genómico del virus del Nilo Occidental cepa IS-98-ST1 en el que han sido eliminadas dos secuencias repetidas y el tallo-bucle del extremo 3',
- el motivo de señal poliA de SV40,
- un único sitio de clonación *Apal*.

### 1.2. Clonación de SNAP

La amplificación del ADN codificante de la secuencia SEC ID nº 2 de la proteína SNAP se llevó a cabo en el molde pMT/BiP/CHIK.se2+etiqueta SNAP mediante PCR utilizando el par de cebadores 5'-SNAP y 3'-MCS tal

como se indica posteriormente.

Cebador 5'-SNAP: 5'-aaaaagatctgacaagactgcgaaatg-3' (SEC ID nº 7)

Cebador 3'-MCS: 5'-gaggagagggtaggataggctacc-3' (SEC ID nº 8)

5 A continuación, se digirió el producto de PCR con *Bgl*II y *Not*I y se insertó entre los sitios únicos *Bgl*II (en el extremo 5' de MCS) y *Not*I (en el extremo 3' de MCS) del plásmido linearizado p/MT/BiP/V5-A en el sistema DES.

10 El plásmido resultante es el vector pMT/BiP/SNAP-His<sub>6</sub>etiqueta de SEC ID nº 9, que comprende:

- 10 – la secuencia de BiP<sub>mc</sub> de insecto de SEC ID nº 11,
- la secuencia de ADN de SNAP de SEC ID nº 1,
- 15 – el sitio de corte de enterocinasa de SEC ID nº 12,
- un sitio de restricción *EcoRV-Sma*I,
- el ADN codificante de una etiqueta His<sub>6</sub> (SEC ID nº 14) situada cadena abajo del sitio de restricción, y
- 20 – dos secuencias espaciadoras de ADN de SEC ID nº 13 situadas i) entre la secuencia de intensificador y el sitio de restricción *EcoRV-Sma*I, e ii) entre el sitio de restricción *EcoRV-Sma*I y el ADN codificante de una etiqueta His<sub>6</sub>.

25 También puede obtenerse un vector pMT/BiP/SNAP-His<sub>6</sub>etiqueta a partir de un esqueleto de pUC57 y se obtiene un vector que presenta la secuencia SEC ID nº 10. Dicho vector comprende:

- el sitio de clonación única *EcoR*I
- 30 – promotor metalotioneína
- la región no codificante 5' del ARN genómico del virus del Nilo Occidental cepa IS-98-ST1,
- un codón de inicio de traducción,
- 35 – el péptido de señal de la proteína E de cubierta del virus del Nilo Occidental cepa IS-98-ST1 de SEC ID nº 15
- la secuencia de ADN de SNAP de SEC ID nº 47,
- 40 – sitios de clonación únicos *EcoRV* y *Sma*I/*Xma*I para la inserción dentro del marco de la secuencia foránea,
- el sitio de corte de enterocinasa de SEC ID nº 12, situado entre el ADN intensificador de SNAP y los sitios de clonación,
- 45 – un ADN codificante de una secuencia de etiqueta hexa-histidina (SEC ID nº 14),
- dos secuencias espaciadoras de ADN de SEC ID nº 13 situadas i) entre la secuencia de intensificador y el sitio de restricción *EcoRV-Sma*I, e ii) entre el sitio de restricción *EcoRV-Sma*I y el ADN codificante de una etiqueta His<sub>6</sub>.
- 50 – dos codones de parada de traducción,
- La región no codificante 3' del ARN genómico del virus del Nilo Occidental cepa IS-98-ST1 en el que han sido deletadas dos secuencias repetidas y el tallo-bucle del extremo 3',
- 55 – motivos de señal poliA de S40 y
- 60 – un único sitio de clonación *Ap*aI.

También puede obtenerse un vector pMT/tipo BiP/SNAP-His<sub>6</sub>etiqueta a partir de un esqueleto de pUC57 en el que se ha insertado la SEC ID nº 59 (ver también la figura 8) entre los sitios únicos *EcoRV* e *Hind*III. Dicho vector

presenta la secuencia SEC ID nº 64. Comprende:

- el sitio de clonación única *EcoRI*
- 5 – promotor metalotioneína
- la región no codificante 5' del ARN genómico del virus del Nilo Occidental cepa IS-98-ST1,
- 10 – un codón de inicio de traducción,
- El péptido de señal de la proteína E de cubierta del virus del Nilo Occidental cepa IS-98-ST1 de SEC ID nº 15,
- 15 – la secuencia de ADN de SNAP de SEC ID nº 47,
- sitios de clonación únicos *BamHI*, *EcoRV*, *Apal* y *XmaI* para la inserción en el marco de la secuencia foránea,
- 20 – dos sitios de corte de pro-TEV de SEC ID nº 52, situados entre el ADN intensificador de SNAP y la etiqueta de His,
- un ADN codificante de una secuencia de etiqueta hexa-histidina (SEC ID nº 14),
- 25 – dos secuencias espaciadoras de ADN de SEC ID nº 13 situadas i) entre la secuencia de intensificador y el sitio de restricción *EcoRV-SmaI*, e ii) entre el sitio de restricción *Apal* y el ADN codificante de una etiqueta His<sub>6</sub>.
- dos codones de parada de traducción,
- 30 – La región no codificante 3' del ARN genómico del virus del Nilo Occidental cepa IS-98-ST1 en el que han sido eliminadas dos secuencias repetidas y el tallo-bucle del extremo 3',
- motivos de señal poliA de S40 y
- 35 – Sitio de clonación único *Apal*.

### 1.3. Clonación de un gen de interés

#### 1.3.1. Nucleoproteína N del virus de la fiebre del valle del Rift (FVR-N)

40 Se llevó a cabo la mutagénesis dirigida en un ADNc codificante de la secuencia proteica de FVR-N mediante PCR utilizando el par de cebadores 5'-N y 3'-N tal como se lista posteriormente.

Cebador 5'-N:

45 5'- aaaaaggcgccaggggggtggcgatctgacaactatcaagagctcgagtcagttgctgctc- 3' (SEC ID nº 17)

Cebador 3'-N:

50 5'- aaaaaaccggcaatgatgatgatgatgacttcaccgccggctgctgtcttgaagcctgagcgg- 3' (SEC ID nº 18)

#### 1.3.2. Proteína no vírica, por ejemplo IFN $\alpha$ 1 o granzima M

55 También puede utilizarse la secuencia de la proteína IFN $\alpha$ 1 humana de SEC ID nº 32.

También puede utilizarse la secuencia de la proteína granzima M humana de SEC ID nº 54. La granzima M es una serina proteasa de tipo quimiotripsina que corta preferentemente sus sustratos después de Met o Leu. Se expresa constitutivamente en células asesinas naturales efectoras innatas activadas. Esta proteasa también presenta propiedades antivíricas y antitumorales (van Domselaar R. et al, The Journal of Immunology, 2010; Hu D. et al., The Journal of Biological Chemistry, 2010).

60 1.4. Inserción de un gen codificante de proteína en el vector pMT/BiP/SNAP-His<sub>etiqueta</sub> o pMT/tipo BiP/SNAP-His<sub>etiqueta</sub>, de manera que se obtiene el vector pMT/BiP/PROTEÍNA SNAP-His<sub>etiqueta</sub> o el vector pMT/tipo BiP/PROTEÍNA SNAP-His<sub>etiqueta</sub>

## 1.4.1. FVR.N

5 Los productos de PCR obtenidos en el punto 1.3.1 se digirieron con *BssHII* y *AgeI* y se insertaron entre los sitios únicos *BssHII* (en el extremo 3' del gen SNAP) y *AgeI* (en el extremo 3' del MCS del vector lanzadera) del plásmido linearizado pMT/BiP/SNAP-His<sub>etiqueta</sub> obtenido en el punto 1.2.

El plásmido resultante es, por ejemplo, pMT/BiP/SNAP-FVR.N/His<sub>etiqueta</sub> (SEC ID nº 19).

## 10 1.4.2. Proteínas ED III de diferentes flavivirus

15 Para potenciar la especificidad de los ensayos ELISA o de inmunotransferencia basados en antígenos flavivíricos recombinantes, el dominio antigénico III de la proteína E (ED III) aparentemente es un enfoque prometedor (Ludolfs *et al.*, 2007). La producción de EDIII recombinante de los virus del Nilo Occidental (NO), Usutu (USU), de la encefalitis japonesa (EJ), de la encefalitis transmitida por garrapatas (ETG), dengue serotipos 1 a 4 (DEN-1, -2, -3 y -4) y de la fiebre amarilla (FA) han sido sometida a ensayo con el método de la invención.

20 Los virus NO, USU y EJ pertenecen al serocomplejo EJ. El sistema de expresión inducible de S2 de *Drosophila* (DES, Invitrogen) se ha seleccionado para la producción en masa de EDIII individual de flavivirus en células no de vertebrado. Los genes sintéticos codificantes del dominio III de longitud completa de las proteína E de los flavivirus NO, USU, EJ, ETG, DEN-1 a DEN-4 y FA se listan en SEC ID nº 23 a SEC ID nº 31. Las secuencias de ED III se fusionaron dentro del marco con el extremo C-terminal de la proteína SNAP en el plásmido pMT/BiP/SNAP-His<sub>etiqueta</sub> obtenido en 1.2 para generar las proteínas de fusión SNAP-EDIII.

25 1.4.3. IFN $\alpha$ 

Se obtuvieron los plásmidos pMT/BiP/SNAP-IFN-His<sub>etiqueta</sub> (ver la figura 4) y pMT/tipo BiP/SNAP-IFN-His<sub>etiqueta</sub> (ver datos en la figura 8).

## 30 1.4.4. Granzima M

Se obtuvo el plásmido pMT/BiP/SNAP-GrM-His<sub>etiqueta</sub> (ver datos en la figura 6).

35 **2. Transfección en células hospedadoras**

## 2.1. Transfección en células S2

40 Los plásmidos resultantes pMT/BiP/PROTEÍNA SNAP-His<sub>etiqueta</sub> que permiten la producción de proteínas etiquetadas con SNAP como proteínas de fusión secretadas terminadas en His<sub>etiqueta</sub> se cotransfectaron con el marcador de selección pCo-Blast en células S2 para generar la línea celular estable S2/sPROTEÍNA-SNAP-His<sub>etiqueta</sub> que mostraba resistencia a la ampicilina.

45 Las líneas celulares S2 estables cultivadas en matraz de agitación (1.000 ml) se estimularon durante 10 días con metal pesado cadmio (Cd<sup>2+</sup>) y las proteínas del medio extracelular se concentraron y se purificaron.

Se observó la acumulación de proteína etiquetada con SNAP secretada en los sobrenadantes de células S2/sPROTEÍNA-SNAP-His<sub>etiqueta</sub> tras 10 días de inducción con metal pesado cadmio.

50 Los ensayos de inmunotransferencia permitieron detectar las proteínas etiquetadas con SNAP extracelulares utilizando anticuerpo sérico de cabra anti-His<sub>etiqueta</sub> (ver las figuras 2B, 3B, 4B y 6C).

## 2.2. Transfección en células HeLa

55 Se transfectó el plásmido pMT/tipo BiP/SNAP-IFN-His<sub>etiqueta</sub> en células HeLa.

Los ensayos de inmunotransferencia permiten detectar IFN extracelular etiquetado con SNAP utilizando anticuerpos anti-SNAP (ver la figura 7B).

60 **3. Purificación de proteínas recombinantes**

Se purificaron proteínas etiquetadas con His extracelular y etiquetadas con SNAP utilizando resina quelante de metales y el método de HPLC.

## 65 3.1. FVR-N y TOS-N

FVR-N

En colaboración con Plate-Forme 5 Production de Protéines recombinantes et d'Anticorps (Institut Pasteur), se han obtenido hasta 97 mg de proteína FVR-N etiquetada con SNAP altamente purificada a partir de 2 litros de sobrenadantes de células S2/sSNAP-FVRV.N-His<sub>etiqueta</sub> 10 días después de la estimulación con Cd<sup>2+</sup>.

5 TOS-N

En colaboración con Plate-Forme 5 Production de Protéines recombinantes et d'Anticorps (Institut Pasteur), se han obtenido hasta 41 mg de proteína FVR.N etiquetada con SNAP altamente purificada a partir de 2 litros de sobrenadantes de células S2/sSNAP-TOS.N-His<sub>etiqueta</sub> 10 días después de la estimulación con Cd<sup>2+</sup>.

10

Resumen de los niveles de producción:

Antígenos víricos	Plásmidos	Líneas celulares estables	Proteínas purificadas	Concentración (/2L)
Gen N de FVR	pMT/BiP/SNAP+FVR.N	S2/SNAP+FVR.N	SNAP-FVR.N	97 mg
Gen N de TOS	pMT/BiP/SNAP+TOS.N	S2/SNAP+TOS.N	SNAP-TOS.N	41 mg

15 3.2. IFN $\alpha$ 1 soluble

Las proteínas IFN $\alpha$ 1 solubles se liberaron de la etiqueta SNAP mediante el corte con el enzima enterocinasa (kit de Novagen).

20 3.3. Antígenos de diferentes flavivirus

Reservas de proteínas etiquetadas con SNAP secretadas utilizando el sistema de expresión de *Drosophila*

Antígenos víricos	Plásmidos	Proteínas purificadas	Producción por cada litro de cultivo celular	Concentración de proteínas purificadas
EDIII de DEN-1	pMT/BiP/SNAP+DV1.EDIII	sSNAP-DV1.EDIII	132 mg	4,93 mg/ml
EDIII de DEN-2	pMT/BiP/SNAP+DV2.EDIII	sSNAP-DV2.EDIII	59 mg	2,67 mg/ml
EDIII de DEN-3	pMT/BiP/SNAP+DV3.EDIII	sSNAP-DV3.EDIII	124 mg	3,5 4mg/ml
EDIII de DEN-4	pMT/BiP/SNAP+DV4.EDIII	sSNAP-DV4.EDIII	43 mg	1,67 mg/ml
EDIII de WN	pMT/BiP/SNAP+WN.EDIII	sSNAP-WN.EDIII	176 mg	4,2 mg/ml
EDIII de JE	pMT/BiP/SNAP+EJ.EDIII	sSNAP-EJ.EDIII	223 mg	6,38 mg/ml
EDIII de USU	pMT/BiP/SNAP+USU.EDIII	sSNAP-USU.EDIII	182 mg	4,8 mg/ml
EDIII de TBE	pMT/BiP/SNAP+ETG.EDIII	sSNAP-ETG.EDIII	180 mg	5,13 mg/ml
EDIII de YF	pMT/BiP/SNAP+FA.EDIII	sSNAP-FA.EDIII	120 mg	3,46 mg/ml
EDIII de MVE	pMT/BiP/SNAP+EVM.EDIII	sSNAP-EVM.EDIII	87 mg	
EDIII de ROCIO	pMT/BiP/SNAP+Rocio.EDIII	sSNAP-Rocio.EDIII	79 mg	
EDIII de WSL	pMT/BiP/SNAP+WSL.EDIII	sSNAP-WSL.EDIII	63 mg	2 mg/ml
EDIII de ZIKA	pMT/BiP/SNAP+Zika.EDIII	sSNAP-ZIKA.EDIII	152 mg	3,8 mg/ml
SNAP-DV1ectoM	pMT/BiP/SNAP-DV1ectoM	SNAP-DV1ectoM	49 mg	1,4 mg/ml
Gen N de FVR	pMT/BiP/SNAP+N.FVR	sSNAP-N.FVR	97 mg	1,3 mg/ml
Gen N de TOS	pMT/BiP/SNAP+N.TOS	sSNAP-N.TOS	41 mg	1,65 mg/ml
SNAP	pMT/BiP/SNAP	SNAP	13 mg	1 mg/ml
sE de NO	pMT/BiP/WN.sE+SNAP	WN.sE+SNAP	40 mg	2,3 mg/ml
sE2 de CHIK	pMT/BiP/CHIK.sE2+SNAP	CHIK.sE2-etiqueta SNAP	90 mg	1,2 mg/ml
SNAP-EKS-IFNA1	pMT/BiP/SNAP-EKS-IFNA1	SNAP-EKS-IFNA1	49 mg	3,5 mg/ml

25 • EDIII : dominio antigénico III de proteínas E de flavivirus (Dengue [DEN], Nilo Occidental [NO], encefalitis japonesa [EJ], Usutu [USU], encefalitis transmitida por garrapatas [ETG], fiebre amarilla [FA], encefalitis del valle de Murray [EVM], Wesselbron [WSL], Rocio, Zika)

30 • ectoM : ectodominio de la proteína M del virus Dengue de tipo 1

• Gen N de FVR: nucleoproteína N del virus de la fiebre del valle del Rift (antígeno vírico mayor)

• Gen N de TOS: nucleoproteína N del virus Toscana (antígeno vírico mayor)

35 • sE de WN : forma soluble de la proteína E de cubierta del virus del Nilo Occidental

- sE2 de CHIK: forma soluble de la proteína E2 de cubierta del virus Chikungunya
- SNAP-IFN $\alpha$ 1: interferón-alfa 1 en fusión con SNAP.

#### 5 3.4. Producción de granzima M

Se recuperaron 10 mg de proteína SNAP-GrM por cada litro de sobrenadante de cultivo en 7 días.

10 Tras las etapas de purificación, se detectaron tres formas de SNAP-GrM (ver la figura 6C), que correspondían al corte de la proteína SNAP por el enzima GrM acoplado.

Lo anterior significa claramente que la proteasa humana se encuentra activa tras su producción mediante el método de la invención (ver posteriormente).

#### 15 4. Control de las proteínas etiquetadas con SNAP

Los ensayos de inmunotransferencia utilizando anticuerpos específicos (que reconocían la proteína de interés y/o el marcaje de etiqueta histidina) detectaron una producción sustancial de proteínas extracelulares etiquetadas con SNAP:

20 el ensayo de inmunotransferencia detectó proteína extracelular FVR.N etiquetada con SNAP utilizando anticuerpo sérico de cabra anti-His<sub>etiqueta</sub> (figura 2B). Los sueros inmunológicos humano y de ratón contra FVR.N reconocen específicamente proteína FVR.N recombinante etiquetada con SNAP.

25 Los ensayos de inmunotransferencia no mostraron reactividad cruzada entre ED III de NO y USU recombinantes utilizando sueros policlonales de ratón específicos a pesar del elevado nivel de similitud de las secuencias. De esta manera, la proteína SNAP-EDIII soluble secretada de los virus NO, EJ y USU resultan adecuados como antígenos víricos recombinantes para el diagnóstico específico de miembros del serocomplejo EJ ya que el virus USU y, en menor medida, los virus EJ han sido identificados recientemente como arbovirus potencialmente  
30 emergentes en Europa.

#### 5. Actividad de las proteínas recombinantes etiquetadas con SNAP

35 • La proteína SNAP-IFN $\alpha$ 1 recombinante soluble secretada a partir de células S2/SNAP-IFN $\alpha$ 1 inducidas muestra un potente efecto antivírico contra CHIKV.

Se recolectaron sobrenadantes (5 ml) de células S2/SNAP-IFN $\alpha$ 1 ( $5 \times 10^6$  células/ml) estimuladas con Cd<sup>2+</sup> diez días después de la inducción. Se observó acumulación de la proteína SNAP-IFN $\alpha$ 1 soluble en el sobrenadante celular mediante inmunotransferencia utilizando anticuerpo anti-His<sub>etiqueta</sub> (ver posteriormente). Se evaluó la actividad antivírica de SNAP-IFN $\alpha$ 1 en células HeLa infectadas por virus Chikungunya que expresaba el gen luciferasa (Luc). Se determinó la actividad de Luc 6 h después de la infección. Se utilizó IFN alfacon1 (Infergen) como un ensayo interno al conocer su potente efecto antivírico contra CHIKV en células HeLa. El sobrenadante de células S2/SNAP-Tos.N (la proteína N del virus Toscana) estimuladas con Cd<sup>2+</sup> sirvió de control negativo. El gráfico ilustrado en la figura 4C demuestra que 1  $\mu$ l de SNAP-IFN $\alpha$ 1 secretado o 0,1  $\mu$ g de Infergen podrían suprimir la replicación del CHIKV en el interior de las células hospedadoras infectadas. En el gráfico se muestra un efecto dependiente de la dosis de SNAP-IFN $\alpha$ . Todavía se observó un veinte por ciento de actividad de Luc con 0,1  $\mu$ l de SNAP-IFN $\alpha$ 1 soluble o 0,01  $\mu$ g de Infergen. No se observó ningún efecto antivírico utilizando SNAP-TOS-N a la dosis más alta sometida a ensayo.

50 • El granzima M es activo tras ser producido en el sobrenadante de las células S2

Tal como se ha indicado anteriormente, se han detectado tres formas de SNAP-GrM en el sobrenadante de las células S2 transfectadas con el vector pMT/BiP/SNAP-GrM-His<sub>etiqueta</sub> (ver la figura 6C).

55 Dichas tres formas corresponden al corte de la proteína SNAP por el enzima GrM acoplado. SNAP en efecto contiene tres sitios potenciales de corte de GrM (ver la figura 6B). Los inmunoensayos con anticuerpos anti-His o anti-SNAP han revelado que dichas tres formas en efecto son fragmentos de la proteína de fusión secretada SNAP-GrM.

60 La forma más pequeña (35 kDa) corresponde a GrM, que ha sido delecionada con la mayor parte de SNAP durante el procedimiento de purificación.

Estos resultados demuestran claramente que la proteasa GrM que ha sido producida por el sistema de la invención se encuentra activo, aunque acoplado con la proteína SNAP.

65

Lo anterior resulta realmente interesante debido a que es conocido que las proteasas humanas activas resultan muy difíciles de producir recombinantemente.

- hSULF-2<sup>ΔTMD</sup> se encuentra activa tras ser producida en el sobrenadante de las células HEK 293. hSULF-2<sup>ΔTMD</sup> fue expresado y purificado a partir de células HEK-293 transfectadas con un plásmido recombinante pcDNA3/De SNAPuniv-hSULF-2<sup>ΔTMD</sup> (ver la figura 13A).

La actividad enzimática del polipéptido hSULF-2<sup>ΔTMD</sup> obtenido en el sobrenadante se evaluó de la manera siguiente:

se transfectaron transitoriamente células HEK 293 con pcDNA3/DeSNAP-hSULF2ΔTMDs. Tras dos días, se incubó una alícuota del sobrenadante celular con el pseudosustrato no fluorescente 4-metil-umbeliferona (4-MUS) a una concentración de 20 mM en Tris 50 mM, pH 7,5, con MgCl<sub>2</sub> 20 mM, y se incubó (1:1, v/v) con el enzima (en medio condicionado) durante 2 a 4 horas a 37°C en una placa de 96 pocillos. Se detuvo la reacción enzimática mediante la adición (1:1 v/v) de NaHCO<sub>3</sub>/Na<sub>2</sub>CO<sub>3</sub> 1,5 M, pH 10,5 y se monitorizó la generación del producto fluorescente 4-metil-umbeliferona mediante fluorimetría (excitación: 360 nm). Se midieron los valores de la actividad de SULF en el sobrenadante celular a partir de la densidad óptica (DO) a 460 nm.

Resulta interesante que la proteína secretada (acoplada a SNAP?) se encuentra activa, tal como se muestra en la figura 13B.

#### 6. Estabilidad de las proteínas recombinantes etiquetadas con SNAP

Inesperadamente se ha observado que las proteínas de fusión que comprenden el péptido SNAP son mucho más estables in vitro que en su ausencia.

Se incubaron las proteínas CHIK.sE2-SNAP, SNAP-WN.EDIII y SNAP-IFNAI altamente purificadas a las concentraciones no saturantes de 0,1 mg/ml (vol.: 0,1 ml) en PBS estéril durante 4 días a -80°C, a 4°C, a 25°C o a 37°C, o durante dos meses a la misma temperatura.

Se separaron muestras de proteína (1 µg) en SDS-PAGE al 4-12% y se visualizaron con pigmento azul brillante de Coomassie G-250 utilizando la solución de tinción de proteínas PageBlue (Fermentas).

Las figuras 10A y B dan a conocer los resultados obtenidos mediante la comparación de la estabilidad de tres proteínas de fusión diferentes in vitro.

Resulta importante que todas las proteínas de fusión aparentemente se encontraban intactas tras dos meses a 4°C y también tras cuatro días de incubación a temperatura ambiente (25°C) o a 37°C.

En particular, IFN no resulta afectado tras 4 días a la temperatura corporal (37°C) y todavía se observa tras dos meses a 37°C, de manera que la estabilidad in vivo es probable que resulte altamente incrementada mediante su acoplamiento con SNAP.

#### 7. Detección in vitro de SNAP-FVR.N producido por células S2

Se utilizaron proteínas de fusión SNAP-FVR.N que habían sido producidas siguiendo los protocolos anteriormente indicados como herramienta diagnóstica para la detección de anticuerpos anti-FVR.N en el suero de ovinos infectados.

Dichas proteínas de fusión se sometieron a ensayo y se compararon con kits comerciales de detección (kit de detección de IgG de FVR de BDSL y FVR multiespecie de IdVet).

Los ensayos se llevaron a cabo con 46 sueros de oveja, estando las ovejas inmunizadas con vacunas de FVR. Se utilizaron las proteínas de fusión SNAP-FVR.N para recubrir directamente el fondo de los pocillos o se biotinilaron y se añadieron a pocillos con recubrimiento de estreptavidina.

Se detectaron anticuerpos anti-FVR mediante el método de ELISA indirecto utilizando placas de microtitulación de 96 pocillos directamente recubiertos con 0,2 µg de antígeno SNAP-FVR.N recombinante altamente purificado en BS (concentración: 2 µg de proteína/ml) durante la noche a 4°C. Tras la saturación, los sueros diluidos se incubaron con SNAP-FVR.N. Como anticuerpo secundario se utilizó anticuerpo de cabra anti-IgG conjugado con peroxidasa. Se llevó a cabo el ELISA con el sistema de sustrato de peroxidasa y se midió la densidad óptica (DO) a 450 nm. Los sueros de muestra se consideraron positivos en el caso de que la OD fuese el doble de la DO de los sueros no inmunológicos.

Resulta interesante que los resultados demuestran que las proteínas de fusión SNAP-FVR.N proporcionan la



misma sensibilidad y especificidad que las proteínas comerciales al utilizarse para recubrir directamente los pocillos (no mostrado). Los resultados son menos reproducibles en el caso de que las proteínas se encuentren biotiniladas.

- 5 Se obtuvieron los mismos resultados con sueros obtenidos de bovinos inmunizados naturalmente (datos no mostrados).

Estos resultados demuestran que las proteínas de fusión de la invención pueden ser utilizadas como herramientas diagnósticas para la identificación de infecciones víricas o bacterianas en muestras biológicas.

10

### 8. Inmunoensayo basado en perlas multiplex

En el contexto de la invención, se desarrolló un inmunoensayo basado en perlas multiplex para la detección rápida y simultánea de anticuerpos contra arbovirus en líquidos biológicos.

15

El sistema se basa en la tecnología xMAP (Luminex Corporation) y utiliza una mezcla de microesferas recubiertas con antígenos como reactivos de captura para inmunoglobulinas humanas específicas. Se acoplaron grupos diferentes de microesferas (Magplex, Luminex Corporation) con proteínas de fusión de MGMT purificadas, es decir, las proteínas recombinantes víricas etiquetadas con SNAP indicadas en la sección 3.3: sSNAP-DV1.EDIII, sSNAP-DV2.EDIII, sSNAP-DV3.EDIII, sSNAP-DV4.EDIII, sSNAP-WN.EDIII, sSNAP-EJ.EDIII, sSNAP-USU.EDIII, sSNAP-ETG.EDIII, sSNAP-FA.EDIII, sSNAP-EVM.EDIII, sSNAP-Rocio.EDIII, sSNAP-WSL.EDIII, sSNAP-ZIKA.EDIII, SNAP-DVlectoM, sSNAP-N.FVR, sSNAP-N.TOS y CHIK.sE2-SNAP. Los antígenos recombinantes se acoplaron covalentemente con la superficie carboxilo de las microesferas utilizando un sustrato de la proteína MGMT como conector (BG-PEG-NH<sub>2</sub>, New England Biolabs), incrementando de esta manera la eficiencia de captura de los anticuerpos en comparación con los procedimientos estándares de acoplamiento de amina.

20

La validación técnica utilizando anticuerpos anti-etiqueta SNAP y anticuerpos monoclonales de ratón específicos confirmó la eficiencia de acoplamiento y demostró la estabilidad a largo plazo de los antígenos (hasta seis meses). La presente solicitud no se limita a antígenos víricos, ya que puede utilizarse cualquier péptido o polipéptido para el recubrimiento de perlas y la posterior captura de anticuerpos.

25

### Referencias bibliográficas

35

Ausubel F. M. *et al.* (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, Inc. (1994).

Bond B.J. *et al.*, Mol. Cell. Biol. 6:2080 (1986)

Brehin *et al.*, Virology 371:185, 2008

40

Brinster *et al.*, Nature, 296:39-42, 1982

Cheever *et al.*, Clinical Cancer Research, 2009, 15:5323-5337

45

Dai *et al.*, the Journal of Biological Chemistry, 2005, vol.280. n°48, pp.40066-40073

Daniels D.S. *et al.*, EMBO J. 19: 1719-1730, 2000

Felgner *et al.*, Proc. Natl. Acad. Sci. U.S.A., 84:7413-7417, 1987

50

Harlowe, *et al.*, 1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press

Hellwig S *et al.*, Nature Biotechnology, n°11, vol.22, 2004

55

Hu D. *et al.*, 2010 The Journal of Biological Chemistry, vol.285, n°24, pp18326-18335

Juillerat A. *et al.*, Chemistry & Biology, vol.10, 313-317, 2003

Lastowski-Perry *et al.*, J.Biol. Chem. 260:1527 (1985)

60

Lim A. *et al.*, EMBO J. 15: 4050-4060, 1996

Ludolfs *et al.*, 2007 Eur J Clin Microbiol Infect Dis. 2007 Jul;26(7):467-73

Ma JKC *et al.*, Nature Genetics, Review 2004, 794-800

65

Mackey *et al.*, Proc. Natl. Acad. Sci. U.S.A., 85:8027-8031, 1988

Miller and Rosman, *BioTechniques*, 7:980-990, 1992

Mokkim *et al.* *protein expression purif.* 63:140, 2009.

5 Pegg A.E. *et al.*, *Mutat.Res.* 2000; 462, 82-100

Perbal B. E., *A Practical Guide to Molecular Cloning* (1984);

10 Sambrook, Fitch & Maniatis, *Molecular Cloning: A Laboratory Manual, Second Edition* (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (referred to herein as "*Sambrook et al.*, 1989");

Shimoda Y. and Watanabe K. , *Cell adhesion and migration*, 2009, 3:1, 64-70

15 van Domselaar R. *et al.*, *The Journal of Immunology*, 2010

Wibley *et al.*, *Nucleic acid research* 2000

Williams *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 88:2726-2730, 1991

20 Wu and Wu, *J. Biol. Chem.*, 263:14621-14624,1988;

Wu *et al.*, *J. Biol. Chem.*, 267:963-967, 1992;

25 DNA Cloning: A Practical Approach, Volumes I and II (D. N. Glover ed. 1985);

Oligonucleotide Synthesis (M. J. Gait ed. 1984);

Nucleic Acid Hybridization (B. D. Hames & S. J. Higgins, eds. 1984);

30 Animal Cell Culture (R. I. Freshney, ed. 1986);

Immobilized Cells and Enzymes (IRL Press, 1986);

35 **Listado de secuencias**

<110> INSTITUT PASTEUR

<120> Método basado en MGMT para obtener un rendimiento elevado de expresión de proteínas recombinantes

40 <130> 359961D29054

<150> EP 10306389.7

<151> 2010-12-09

45 <150> US 61/505.694

<151> 2011-07-08

<160> 97

50 <170> PatentIn versión 3.5

<210> 1

<211> 579

55 <212> ADN

<213> Secuencia artificial

<220>

<223> ADN de SNAP (mutante de MGTh)

60 <400> 1

ES 2 627 117 T3

```

agatctgaca aagactgcga aatgaagcgc accaccctgg atagccctct gggcaagctg      60
gaactgtctg ggtgcgaaca gggcctgcac gagatcaagc tgctgggcaa aggaacatct      120
gccgccgacg ccgtggaagt gctgccccca gccgccgtgc tgggaggacc agagccactg      180
atgcaggcca ccgcctggct caacgcctac tttcaccagc ctgaggccat cgaggagttc      240
cctgtgccag ccctgcacca cccagtgttc cagcaggaga gctttaccog ccaggtgctg      300
tggaaactgc tgaaagtggg gaagttcggg gaggtcatca gctaccagca gctggccgcc      360
ctggccggca atcccgccgc caccgccgcc gtgaaaaccg ccctgagcgg aatcccctg      420
ccattctga tcccctgcca ccgggtgggt tctagctctg gcgccgtggg gggctacgag      480
ggcgggctcg ccgtgaaaga gtggctgctg gccacagagg gccacagact gggcaagcct      540
gggctgggtc ctgcaggtat aggcgcgcca ggtcccta                                579

```

<210> 2  
 <211> 193  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Secuencia de aminoácidos de SNAP

10

<400> 2

```

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

```

ES 2 627 117 T3

Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His Glu Ile  
 20 25 30

Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu Val Pro  
 35 40 45

Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln Ala Thr  
 50 55 60

Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu Glu Phe  
 65 70 75 80

Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser Phe Thr  
 85 90 95

Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly Glu Val  
 100 105 110

Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala Ala Thr  
 115 120 125

Ala Ala Val Lys Thr Ala Leu Ser Gly Asn Pro Val Pro Ile Leu Ile  
 130 135 140

Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly Tyr Glu  
 145 150 155 160

Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly His Arg  
 165 170 175

Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro Gly Ser  
 180 185 190

Leu

<210> 3  
 <211> 1265  
 <212> ADN  
 <213> homo sapiens

<400> 3

ctcggccccg ccccgcgcc cgggatatgc tgggacagcc cgcgccccta gaacgctttg 60  
 cgtcccgcagc cccgcaggtc ctcgcggtgc gcaccgtttg cgacttggtta cttgaaaaa 120  
 tggacaagga ttgtgaaatg aaacgcacca cactggacag ccctttgggg aagctggagc 180  
 tgtctggttg tgagcaggtt ctgcacgaaa taaagctcct gggcaagggg acgtctgcag 240  
 ctgatgccgt ggagggccca gcccccgctg cggttctcgg aggtccggag cccctgatgc 300

10

ES 2 627 117 T3

agtgcacagc ctggtgaat gcctatttcc accagcccga ggctatcgaa gagttccccg 360  
 tgccggctct tcaccatccc gttttccagc aagagtcggt caccagacag gtggtatgga 420  
 agctgctgaa ggttgtgaaa ttcggagaag tgatttctta ccagcaatta gcagccctgg 480  
 caggcaacc caaagccgcg cgagcagtgg gaggagcaat gagaggcaat cctgtcccca 540  
 tcctcatccc gtgccacaga gtggtctgca gcagcggagc cgtgggcaac tactccggag 600  
 gactggccgt gaaggaatgg cttctggccc atgaaggcca ccggttgggg aagccaggct 660  
 tgggaggggag ctgaggtctg gcaggggcct ggctcaaggg agcgggagct acctcgggct 720  
 ccccgcctgc tggccgaaac tgagtatgtg cagtaggatg gatgtttgag cgacacacac 780  
 gtgtaacact gcatcggatg cggggcgtgg aggcaccgct gtattaaagg aagtggcagt 840  
 gtcctgggaa caagcgtgtc tgccctttct gtttccatat tttacagcag gatgagttca 900  
 gacgcccggg gtcctgcaca catttgtttc cttctctaac gctgcccttg ctctattttt 960  
 catgtccatt aaaacaggcc aagtgagtgt ggaaggcctg gctcatgttg ggccacagcc 1020  
 caggatgggg cagtctggca ccctcaggcc acagacggct gccatagccg ctgtccaggg 1080  
 ccagctaagg cccatcccag gccgtccaca ctagaagct ggccctgcc catccccacc 1140  
 atgcctccct tcctggctgt gtccatggct gtgatggcat tctccactca gcagttccta 1200  
 gcatcccaca ccaggtctc actgaaagaa aggggaacag gccatggcag tcagtgtta 1260  
 cagag 1265

<210> 4

<211> 238

5 <212> PRT

<213> homo sapiens

<400> 4

Met Leu Gly Gln Pro Ala Pro Leu Glu Arg Phe Ala Ser Arg Arg Pro  
 1 5 10 15

Gln Val Leu Ala Val Arg Thr Val Cys Asp Leu Val Leu Gly Lys Met  
 20 25 30

Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp Ser Pro Leu Gly  
 35 40 45

Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His Glu Ile Lys Leu  
 50 55 60

Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu Val Pro Ala Pro  
 65 70 75 80

ES 2 627 117 T3

Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln Cys Thr Ala Trp  
85 90 95

Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu Glu Phe Pro Val  
100 105 110

Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser Phe Thr Arg Gln  
115 120 125

Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly Glu Val Ile Ser  
130 135 140

Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Lys Ala Ala Arg Ala  
145 150 155 160

Val Gly Gly Ala Met Arg Gly Asn Pro Val Pro Ile Leu Ile Pro Cys  
165 170 175

His Arg Val Val Cys Ser Ser Gly Ala Val Gly Asn Tyr Ser Gly Gly  
180 185 190

Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly His Arg Leu Gly  
195 200 205

Lys Pro Gly Leu Gly Gly Ser Ser Gly Leu Ala Gly Ala Trp Leu Lys  
210 215 220

Gly Ala Gly Ala Thr Ser Gly Ser Pro Pro Ala Gly Arg Asn  
225 230 235

<210> 5  
<211> 713  
5 <212> ADN  
<213> Secuencia artificial

<220>  
10 <223> Secuencia de ADN de BiP + SNAP+ sitio de enterocinasa + EcoRV/XmaI+Etiqueta de His

<400> 5  
atgaagtat gcatattact ggccgtcgtg gcctttgttg gcctctcgct cgggagatct 60  
gacaaagact gcgaaatgaa gcgcaccacc ctggatagcc ctctgggcaa gctggaactg 120  
tctgggtgcg aacagggcct gcacgagatc aagctgctgg gcaaaggaac atctgccgcc 180  
gacgccgtgg aagtgcctgc cccagccgcc gtgctgggcg gaccagagcc actgatgcag 240  
gccaccgcct ggtcaacgc ctactttcac cagcctgagg ccatcgagga gttccctgtg 300  
ccagccctgc accaccagt gttccagcag gagagcttta cccgccaggt gctgtgaaaa 360  
ctgctgaaaag tggatgaagt cggagaggtc atcagctacc agcagctggc cgccctggcc 420  
ggcaatcccg ccgccaccgc cgccgtgaaa accgccctga gcggaatcc cgtgcccatt 480  
ctgatcccct gccaccgggt ggtgtctagc tctggcgccg tggggggcta cgagggcggg 540  
ctcgccgtga aagagtggct gctggcccac gagggccaca gactgggcaa gcctgggctg 600  
ggtcctgcag gtatagggcg gccagggtcc ctaggtggcg gatctgatga cgatgataaa 660  
gatataaaaa acccgggccc tggaagtcac catcatcatc atcattgacc ggt 713

ES 2 627 117 T3

<210> 6  
 <211> 614  
 <212> ADN  
 <213> Secuencia artificial  
 5  
 <220>  
 <223> Secuencia de ADN de SNAP+Etiqueta de His  
 <400> 6  
 agatctgaca aagactgcga aatgaagcgc accaccctgg atagccctct gggcaagctg 60  
 gaactgtctg ggtgcgaaca gggcctgcac gagatcaagc tgctgggcaa aggaacatct 120  
 gccgccgacg ccgtggaagt gcctgccccca gccgccgtgc tggggggacc agagccactg 180  
 atgcaggcca ccgcctggct caacgcctac ttccaccagc ctgaggccat cgaggagtcc 240  
 cctgtgccag ccctgcacca cccagtgttc cagcaggaga gctttaccgc ccaggtgctg 300  
 tggaaactgc tgaagtgggt gaagttcgga gaggtcatca gctaccagca gctggccgcc 360  
 ctggccggca atcccgccgc caccgccgcc gtgaaaaccg ccctgagcgg aaatcccgtg 420  
 cccattctga tcccctgccca ccgggtggtg tctagctctg gcgccgtggg gggctacgag 480  
 ggcgggctcg ccgtgaaaga gtggctgctg gcccaagagg gccacagact gggcaagcct 540  
 gggctgggtc ctgcaggat aggcgcgccca gggccctgg agcatcatca tcatcatcat 600  
 10 tgatgagcgg ccgc 614  
 <210> 7  
 <211> 29  
 <212> ADN  
 15 <213> Secuencia artificial  
 <220>  
 <223> Cebador 5'-SNAP  
 20 <400> 7  
 aaaaaagatc tgacaaagac tgcgaaatg 29  
 <210> 8  
 <211> 27  
 25 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 <223> Cebador 3'-MCS (SNAP)  
 30 <400> 8  
 gaggagaggg ttaggatag gcttacc 27  
 <210> 9  
 <211> 4181  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 40 <223> Secuencia de ADN del vector de la invención (DeSNAP1)  
 <400> 9

ES 2 627 117 T3

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat	300
tacgccagct ggcgaaaggg ggatgtgctg caagggcatt aagtgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccag tgaattttaa cgttgcagga	420
caggatgtgg tgcccgatgt gactagctct ttgctgcagg ccgtcctatc ctctggtcc	480
gataagagac ccagaactcc ggccccccac cgcccaccgc ccccccata catatgtggt	540
acgcaagtaa gactgcctgc gcatgcccc tgtgccccac caagagtttt gcatccata	600
caagtcccca aagtggagaa ccgaaccaat tcttcgctgg cagaacaaaa gcttctgcac	660
acgtctccac tcgaatttgg agccggccgg cgtgtgcaaa agaggtgaat cgaacgaaag	720
acccgtgtgt aaagccgctg ttccaaaatg tataaaaccg agagcatctg gccaatgtgc	780
atcagttgtg gtcagcagca aatcaagtg aatcatctca gtgcaactaa aggggggatc	840
cgatctcaat atgaagttat gcatattact ggccgtcgtg gcctttgttg gcctctcgt	900
cgggagatct gacaaagact gcgaaatgaa gcgcaccacc ctggatagcc ctctgggcaa	960
gctggaactg tctgggtgcg aacagggcct gcacgagatc aagctgctgg gcaaaggaac	1020
atctgccgcc gacgccgtgg aagtgcctgc ccagccgcc gtgctgggcg gaccagagcc	1080
actgatgcag gccaccgctt ggctcaacgc ctactttcac cagcctgagg ccatcgagga	1140
gttccctgtg ccagccctgc accaccaggt gttccagcag gagagcttta cccgccaggt	1200
gctgtggaaa ctgctgaaag tggatgaagt cggagaggtc atcagctacc agcagctggc	1260
cgccctggcc ggcaatcccg ccgccaccgc cgccgtgaaa accgccctga gcggaaatcc	1320
cgtgccatt ctgatccct gccaccgggt ggtgtctagc tctggcgccg tggggggcta	1380
cgagggcggg ctgcgcgtga aagagtggct gctggccac gagggccaca gactgggcaa	1440
gcctgggctg ggtcctgcag gtataggcgc gccaggtcc ctggagcatc atcatcatca	1500



ES 2 627 117 T3

tcattgatga gcggccgctc gagtctagag ggcccttcga aggtaagcct atccctaacc 1560  
ctctcctcgg tctcgtattct acgcgtaccg gtcacatca ccatcacat tgagtttaaa 1620  
cccgtgatc agcctcgact gtgccttcta aggcctgagc tcgctgatca gcctcgatcg 1680  
aggatccaga catgataaga tacattgatg agtttgaca aaccacaact agaatgcagt 1740  
gaaaaaaaaatg ctttatttgt gaaatttgtg atgctattgc tttatttga accattataa 1800  
gctgcaataa acaagttaac aacaacaatt gcattcattt tatgtttcag gttcaggggg 1860  
agggtgggga ggttttttaa agcaagtaa acctctaca atgtggtatg gctgattatg 1920  
atcagtcgac ctgcaggcat gcaagcttgg cgtaatcatg gtcatactg tttcctgtgt 1980  
gaaattgtta tccgctcaca attccacaca acatacagc cggaagcata aagtgtaaag 2040  
cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgtt 2100  
tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 2160  
gcggtttgcg tattgggocg tcttccgctt cctcgtcac tgactcgtg cgctcgtcg 2220  
ttcggctgcg gcgagcggta tcagctcact caaagcgggt aatacggta tccacagaat 2280  
caggggataa cgcaggaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 2340  
aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag catcacaataa 2400  
atcagcgtc aagtcagagg tggcgaacc cgacaggact ataaagatac caggcgtttc 2460  
ccctggaag ctccctcgtg cgctctcctg ttcgaccct gccgcttacc ggatacctgt 2520  
ccgcctttct ccttcggga agcgtggocg tttctcatag ctacgctgt aggtatctca 2580  
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg 2640  
accgctgocg ottatccggg aactatcgtc ttgagtccaa ccggtaaga cacgacttat 2700  
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta gccggtgcta 2760  
cagagttctt gaagtgggtg cctaactacg gctacactag aaggacagta ttgggtatct 2820  
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 2880  
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaa 2940  
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 3000  
actcagtta agggattttg gtcacatgat tatcaaaaag gatcttcacc tagatccttt 3060  
taaattaaaa atgaagtttt aatcaatct aaagtatata tgagtaaac tggctgaca 3120  
gttaccaatg cttaatcagt gaggcacctt tctcagcgt ctgtctatct cgttcacca 3180  
tagttgctg actccccgtc gtgtagataa ctacgatac ggagggctta ccatctggcc 3240  
ccagtgctgc aatgataccg cgagaccac gtcaccggc tccagattta tcagcaataa 3300  
accagccagc cgggaagggc gagcgcagaa gtggctcctg aactttatcc gcctccatcc 3360  
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcca 3420

ES 2 627 117 T3

acgttgttgc cattgctaca ggcacgtgg tgtcacgctc gtcgtttgg atggcttcat 3480  
 tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatggtg tgcaaaaaag 3540  
 cggtagctc cttcggctct ccgatcgttgc tcagaagtaa gttggccgca gtgttatcac 3600  
 tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 3660  
 ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt 3720  
 gctcttgccc ggcgtcaata cgggataata ccgcccaca tagcagaact ttaaaagtgc 3780  
 tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 3840  
 ccagttcgat gtaaccact cgtgcacca actgatcttc agcatctttt actttcacca 3900  
 gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaggga ataaggcgca 3960  
 cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 4020  
 gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 4080  
 ttccgcgcac atttccccga aaagtccac ctgacgtcta agaaaccatt attatcatga 4140  
 cattaaccta taaaaatagg cgtatcacga ggcccttctg t 4181

<210> 10  
 <211> 1720  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Secuencia de ADN del vector DeSNAP2 de la invención

<400> 10  
 gaattcgttg caggacagga tgtgggtgcc gatgtgacta gctctttgct gcaggccgctc 60  
 ctatcctctg gttccgataa gagaccaga actccggccc cccaccgcc accgccaccc 120  
 ccatacatat gtggtacgca agtaagagt cctgcgcgatg ccccatgtgc cccaccaaga 180  
 gttttgcac ccatacaagt ccccaaagt gagaaccgaa ccaattctc gcgggcagaa 240  
 caaaagcttc tgcacacgct tccactcgaa tttggagccg gccggcgtgt gcaaaagagg 300  
 tgaatcgaac gaaagaccg tgtgtaaagc cgcgtttcca aaatctataa aaccgagagc 360  
 atctggacca tgtgcatcag ttgtggtcag cagcaaaatc aagtgaatca tctcagtgca 420  
 actaaagggg ggtaccgatc tcaatgcgag ctgtttctta gcacgaagat ctcgatgtct 480  
 aagaaaccag gagggccggg caagagccgg gctgtcaata ccatggttgt gtttgcgtg 540  
 ctattgcttt tggtgcccc agcttacagc cttgatattg aatttacaga caaagactgc 600  
 gaaatgaaaa gaactacatt ggattcacca cttgggaagt tggaaactgag tggatgcgag 660  
 caaggattgc atgaaattaa gcttctggga aaaggaactt ctgcagctga tgcagttgaa 720  
 gttccagcac cagcagctgt tcttgagggt cctgagcccc tcatgcaagc cacagcctgg 780

ES 2 627 117 T3

cttaacgcat atttccacca gcctgaggcc attgaggaat ttccagtccc cgcccttcac 840  
 catcctgtgt ttcagcagga aagcttcacc cgccagggtcc tgtggaaatt gctgaaggtg 900  
 gtcaagtttg gtgaagtgat ttcatatcag caacttgctg cattggccgg taaccccgca 960  
 gctacagctg ccgtgaaaac tgctctcagc ggaaatcctg tgcccatcct gatcccttgt 1020  
 cacagagtgc tttcatcttc cggagctgta ggtggctatg aaggaggact ggcagttaag 1080  
 gagtggctgc tggctcatga aggtcataga cttgaaaaac ctggtttggg aggtggcggg 1140  
 tctgatgacg atgataaaga tatcatatac ccgggcggtg gaagtcatca tcaccatcac 1200  
 cactgataaa tatttaatca attgtaaata gacaataata gtatgcataa aagtgtagtt 1260  
 ttatagtagt atttagtggg gttagtgtaa atagttaaga aaatthttgag gagaaagtca 1320  
 ggccgggaag ttcccgccac cggagttga gtagacggtg ctgcctgcga ctcaacccca 1380  
 ggaggactgg gtgaacaaag ccgcgaagtg atccatgtaa gccctcagaa ccgtctcgga 1440  
 aggaggacct cacatgttgt aacttcaaag cccaatgtca gaccacgcta cggcgtgcta 1500  
 ctctgcggag agtgcagtct gcgatagtgt aacaaaggca aatcaacgcc ccacgcggcc 1560  
 ctagccccgg taatggtggt aaccagggcg aaaggactag aggttagagg agaccccccg 1620  
 gtttaaagtg cacggcccag cctggctgaa gctgtaggtc aggggtggag accccgtgcc 1680  
 acaaaacacc acaacaaaac agcataaata aacagggcc 1720

<210> 11

<211> 54

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> Secuencia de ADN codificante de la secuencia de BiPmc

<400> 11

atgaagttat gcatattact gccctcgtg gcctttgtg gcctctgct cggg 54

<210> 12

15 <211> 18

<212> ADN

<213> Secuencia artificial

<220>

20 <223> Secuencia de ADN codificante del sitio de corte de enterocinasa

<400> 12

gatgacgatg ataaagat 18

<210> 13

25 <211> 12

<212> ADN

<213> Secuencia artificial

<220>

30 <223> Secuencia espaciadora de ADN

<400> 13

ggtggcggat ct 12

35

<210> 14

<211> 18

<212> ADN

# ES 2 627 117 T3

<213> Secuencia artificial

<220>

<223> Secuencia de ADN codificante de la etiqueta de His

5 <400> 14  
catcatcatc atcatcat 18

10 <210> 15  
<211> 17  
<212> PRT  
<213> Cepa de virus del Nilo Occidental IS-98-ST1

<400> 15  
Met Val Val Phe Val Val Leu Leu Leu Leu Val Ala Pro Ala Tyr Ser  
1 5 10 15

15 Leu

<210> 16  
<211> 1361  
<212> ADN

20 <213> Secuencia artificial

<220>

<223> Secuencia de ADN de SNAP/FVR.N/Etiqueta de His

25 <400> 16

agatctgaca	aagactgcga	aatgaagcgc	accaccctgg	atagccctct	gggcaagctg	60
gaactgtctg	ggtgcgaaca	gggcctgcac	gagatcaagc	tgctgggcaa	aggaacatct	120
gccgccgacg	ccgtggaagt	gcctgccccca	gccgccgtgc	tgggaggacc	agagccactg	180
atgcaggcca	ccgcctggct	caacgcctac	tttcaccagc	ctgaggccat	cgaggagttc	240
cctgtgccag	ccctgcacca	cccagtgttc	cagcaggaga	gctttacccg	ccaggtgctg	300
tggaaactgc	tgaaagtggg	gaagttcgga	gaggtcatca	gctaccagca	gctggccgcc	360
ctggccggca	atcccgccgc	caccgccgcc	gtgaaaaccg	ccctgagcgg	aaatcccgtg	420
cccattctga	tcccctgcca	ccgggtggg	tctagctctg	gcgccgtggg	gggctacgag	480
ggcgggctcg	ccgtgaaaga	gtggctgctg	gccacagagg	gccacagact	gggcaagcct	540
gggctgggtc	ctgcaggat	aggcgcgccca	gggtccctag	gtggcggatc	tgacaactat	600
caagagcttc	gagtccagtt	tgctgctcaa	gcagtggacc	gcaatgagat	tgaacagtgg	660

ES 2 627 117 T3

	gtccgagagt ttgcttatca agggtttgat gcccgtagag ttatcgaact cttaaagcag	720
	tatggtgggg ctgactggga gaaggatgcc aagaaaatga ttgttctggc tctaactcgt	780
	ggcaacaagc ccaggaggat gatgatgaaa atgtcgaaag aaggcaaagc aactgtggag	840
	gctctcatca acaagtataa gctaaaggaa gggaatcctt cccgggatga gttgactcta	900
	tcacgagttg ctgccgcctt ggctggctgg acatgccagg ctttggctcgt cttgagtgag	960
	tggcttctctg tcaactgggac taccatggac ggcctatccc ctgcataccc gaggcataatg	1020
	atgcacccca gctttgctgg catggtggat ccttctctac caggagacta tctaagggca	1080
	atattagatg ctcaactctct gtatctgctg cagttctccc gggatcatcaa cccaaacctc	1140
	cgaggtagaa caaaagagga ggttgctgca acgttcacgc agccaatgaa tgcagcagtg	1200
	aatagcaact ttataagcca tgagaagagg agagaattct tgaaagcctt tggacttgtg	1260
	gattccaatg ggaagccgtc agctgctgtc atggcagccg ctcaggctta caagacagca	1320
	gccggcgggtg gaagtcatca tcatcatcat cattgaccgg t	1361
	<210> 17	
	<211> 66	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
10	<223> Cebador 5'N (para la clonación de FVR-N)	
	<400> 17	
	aaaaaggcgc gccaggggggt ggcggatctg acaactatca agagcttcga gtccagtttg	60
	ctgctc	66
	<210> 18	
15	<211> 69	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
20	<223> Cebador 3'N (para la clonación de FVR-N)	
	<400> 18	
	aaaaaacccg tcaatgatga tgatgatgat gacttcacc gccggctgct gtcttgtaag	60
	cctgagcgg	69
25	<210> 19	
	<211> 4854	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Secuencia de ADN de pMT/BiP/SNAP-FVR.N/Etiqueta de His	
	<400> 19	

ES 2 627 117 T3

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60  
 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
 ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240  
 attcgccatt caggctgcmc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300  
 tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360  
 tttcccagtc acgacgttgt aaaacgacgg ccagtgccag tgaattttaa cgttcgagga 420  
 caggatgtgg tgcccgatgt gactagctct ttgctgcagg ccgtcctatc ctctggttcc 480  
 gataagagac ccagaactcc ggccccccac cgcccaccgc ccccccata catatgtggt 540  
 acgcaagtaa gagtgcctgc gcatgcccc a tgtgccccac caagagtttt gcatcccata 600  
 caagtcccc aagtgagaa ccgaaccaat tcttcgctgg cagaacaaaa gcttctgcac 660  
 acgtctccac tcgaatttgg agccggccgg cgtgtgcaaa agaggtgaat cgaacgaaag 720  
 acccgtgtgt aaagccgctt tccaaaatg tataaaaccg agagcatctg gccaatgtgc 780  
 atcagttgtg gtcagcagca aatcaagtg aatcatctca gtgcaactaa aggggggatc 840  
 cgatctcaat atgaagttat gcatattact ggccgtcgtg gcctttgttg gcctctcgt 900  
 cgggagatct gacaaagact gcgaaatgaa gcgcaccacc ctggatagcc ctctgggcaa 960  
 gctggaactg tctgggtgcg aacagggcct gcacgagatc aagctgctgg gcaaaggaac 1020  
 atctgccgcc gacgccgtgg aagtgcctgc ccagccgcc gtgctgggcg gaccagagcc 1080  
 actgatgcag gccaccgctt ggctcaacgc ctactttcac cagcctgagg ccatcgagga 1140  
 gttccctgtg ccagccctgc accaccaggt gttccagcag gagagcttta cccgccaggt 1200  
 gctgtggaaa ctgctgaaag tgggtgaagt cggagaggtc atcagctacc agcagctggc 1260  
 cgcctggcc ggcaatccc cgcaccgc cgccgtgaaa accgccctga gcggaaatcc 1320  
 cgtgccatt ctgatccct gccaccgggt ggtgtctagc tctggcgcg tgggggcta 1380  
 cgagggcggg ctcccgctga aagagtggct gctggccac gagggccaca gactgggcaa 1440  
 gcctgggctg ggtcctgcag gtataggcgc gccagggtcc ctagggtggc gatctgacaa 1500  
 ctatcaagag cttcgagtcc agtttctgc tcaagcagtg gaccgcaatg agattgaaca 1560  
 gtgggtccga gagtttgcct atcaagggtt tgatgccctg agagttatcg aactctaaa 1620  
 gcagtatggt ggggctgact gggagaagga tgccaagaaa atgattgttc tggctctaac 1680  
 tcgtggcaac aagcccagga ggatgatgat gaaaatgtcg aaagaaggca aagcaactgt 1740  
 ggaggtctc atcaacaagt ataagctaaa ggaagggaat ccttcccggg atgagttgac 1800  
 tetatcacga gttgctgcc ccttggctgg ctggacatgc caggtttgg tcgtcttgag 1860  
 tgagtggctt cctgtcactg ggactaccat ggacggccta tcccctgcat acccgaggca 1920

ES 2 627 117 T3

tatgatgcac	cccagctttg	ctggcatggt	ggatccttct	ctaccaggag	actatetaag	1980
ggcaatatta	gatgctcact	ctctgtatct	gctgcagttc	tcccgggtca	tcaacccaaa	2040
cctccgaggt	agaacaaaaag	aggaggttgc	tgcaacgttc	acgcagccaa	tgaatgcagc	2100
agtgaatagc	aactttataa	gccatgagaa	gaggagagaa	ttcttgaaag	cctttggact	2160
tgtggattcc	aatgggaagc	cgtcagctgc	tgtcatggca	gccgctcagg	cttacaagac	2220
agcagccggc	ggtggaagtc	atcatcatca	tcatcattga	ccggtcatca	tcaccatcac	2280
cattgagttt	aaaccgctg	atcagcctcg	actgtgcctt	ctaaggcctg	agctcgctga	2340
tcagcctoga	tcgaggatcc	agacatgata	agatacattg	atgagtttgg	acaaaccaca	2400
actagaatgc	agtgaaaaaa	atgctttatt	tgtgaaatth	gtgatgctat	tgctttatth	2460
gtaaccatta	taagctgcaa	taaacaagtt	aacaacaaca	attgcattca	ttttatgtht	2520
caggttcagg	gggaggtgtg	ggaggtthtt	taaagcaagt	aaaacctcta	caaatgtggt	2580
atggtgatt	atgatcagtc	gacctgcagg	catgcaagct	tggcgtaatc	atggtcatag	2640
ctgthtctcg	tgtgaaattg	ttatccgctc	acaattccac	acaacatacg	agccggaagc	2700
ataaagtgt	aagcctgggg	tgctaatga	gtgagctaac	tcacattaat	tgcgthtgcg	2760
tcactgccc	ctthccagtc	gggaaacctg	tcgtgccagc	tgcattaatg	aatcgcccaa	2820
cgcgcgggga	gaggcgtht	gcgtattggg	cgctcttcog	cttcctcgct	cactgactcg	2880
ctgcgctcgg	tcgttcggct	gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	2940
ttatccacag	aatcagggga	taacgcagga	aagaacatgt	gagcaaaaag	ccagcaaaaag	3000
gccaggaacc	gtaaaaaggc	cgcgttgctg	gcgthtttcc	ataggctccg	ccccctgac	3060
gagcatcaca	aaaatcgacg	ctcaagtcag	agggtggcga	accgcagcag	actataaaga	3120
taccagcgct	ttccccctgg	aagctccctc	gtgcgctctc	ctgthtccgac	cctgcgctth	3180
accggatacc	tgtccgcctt	tctcccttcg	ggaagcgtgg	cgctthtctca	tagctcacgc	3240
tgtaggtatc	tcagthtgggt	gtaggtcgth	cgctccaagc	tgggctgtgt	gcacgaaacc	3300
cccgttcagc	ccgaccgctg	cgccttatcc	ggtaactatc	gtcttgagtc	caaccggta	3360
agacacgact	tatcgccact	ggcagcagcc	actggtaaca	ggattagcag	agcgaggtat	3420
gtaggcggtg	ctacagagth	cttgaagtgg	tggcctaact	acggctacac	tagaaggaca	3480
gtatthtggta	tctgcgctct	gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	3540
tgatccggca	aacaaaccac	cgctggtagc	ggthgtthtt	thgtthtcaa	gcagcagatt	3600
acgcgcagaa	aaaaaggatc	tcaagaagat	cctthgatct	thtctacggg	gtctgacgct	3660
cagtggaacg	aaaactcacg	ttaagggtat	thggtcatga	gattatcaaa	aaggatcttc	3720
acctagatcc	thttaaatta	aaaatgaagt	thtaaatcaa	tctaaagtat	atatgagtaa	3780

ES 2 627 117 T3

acttggctctg acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta 3840  
 tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc 3900  
 ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat 3960  
 ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtgggcc tgcaacttta 4020  
 tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt 4080  
 aatagtttgc gcaacggtgt tgccattgct acaggcatcg tgggtgtcacg ctcgctgttt 4140  
 ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg atccccatg 4200  
 ttgtgcaaaa aagcggttag ctcttccggt cctccgatcg ttgtcagaag taagttggcc 4260  
 gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc 4320  
 gtaagatgct tttctgtgac tgggtgagtac tcaaccaagt cattctgaga atagtgtatg 4380  
 cggcgaccga gttgctcttg cccggcgta atacgggata ataccgccc acatagcaga 4440  
 actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta 4500  
 ccgctgttga gatccagttc gatgtaaccc actcgtgac ccaactgatc ttcagcatct 4560  
 tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag 4620  
 ggaataaggg cgacacggaa atggtgaata ctcatactct tcctttttca atattattga 4680  
 agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 4740  
 aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 4800  
 attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgt 4854

<210> 20  
 <211> 1957  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 10 <223> Secuencia de ADN de WNsE/SNAP/Etiqueta de His

<400> 20  
 agatctttca actgccttgg aatgagcaac agagacttct tggaggaggt gtctggagca 60  
 acatgggtgg atttggttct cgaaggcgac agctgctgta ctatcatgtc taaggacaag 120  
 cctaccatcg atgtgaagat gatgaatatg gaggcggcca acctggcaga ggtccgcagt 180  
 tattgtatt tggctaccgt cagcgatctc tccaccaaag ctgctgccc gaccatggga 240  
 gaagctcaca atgacaaaacg tgctgacca gcttttgtgt gcagacaagg agtagtggac 300  
 aggggctggg gcaacggctg cggactatth ggcaaaaggaa gcattgacac atgcgcaaaa 360  
 tttgcctgct ctaccaaggc aataggaaga accatcttga aagagaatat caagtacgaa 420  
 gtggccattt ttgtccatgg accaactact gtggagtcgc acggaaacta ctccacacag 480  
 gttggagcca ctcaggcagg gagattcagc atcactcctg cggcgccttc atacacacta 540



ES 2 627 117 T3

aagcttggag aatatggaga ggtgacagtg gactgtgaac cacggtcagg gattgacacc 600  
aatgcatact acgtgatgac tgttggaaaca aagacgttct tggccatcg tgagtggttc 660  
atggacctca acctcccttg gagcagtgct ggaagtactg tgtggaggaa cagagagacg 720  
ttaatggagt ttgaggaacc acacgccacg aagcagtctg tgatagcatt gggctcacia 780  
gagggagctc tgcatacaagc tttggctgga gccattcctg tggaaatttc aagcaacact 840  
gtcaagtga cgtcgggtca tttgaagtgt agagtgaaga tggaaaaatt gcagttgaag 900  
ggaacaacct atggcgtctg ttcaaaggct ttcaagtttc ttgggactcc cgcagacaca 960  
ggtcacggca ctgtggtggt ggaattgcag taaactggca cggatggacc ttgcaaagtt 1020  
cctatctcgt cagtggcttc attgaacgac ctaacgccag tgggcagatt ggtcactgtc 1080  
aaccttttg tttcagtggc cacggccaac gctaaggctc tgattgaatt ggaaccaccc 1140  
tttgagact catacatagt ggtgggcaga ggagaacaac agattaatca ccattggcac 1200  
aagtctggaa gcagcattgg caaagccttt acaaccaccc tcaaaggagc gcagagacta 1260  
gccgctctag gagacacagc ttgggacttt ggatcagttg gaggggtgtt cacctcagtt 1320  
gggaaggctg tgcggccgct gggcggaggt agcaaagact gcgaaatgaa gcgcaccacc 1380  
ctggatagcc ctctgggcaa gctggaactg tctggtgctg aacagggcct gcacgagatc 1440  
aagctgctgg gcaaaggaac atctgccgcc gacgccgtgg aagtgcctgc cccagccgcc 1500  
gtgctgggcg gaccagagcc actgatgcag gccaccgctt ggctcaacgc ctactttcac 1560  
cagcctgagg ccatacgagga gttccctgtg ccagccctgc accaccaggt gttccagcag 1620  
gagagcttta cccgccaggt gctgtggaag ctgctgaaag tgggtgaagt cggagaggtc 1680  
atcagctacc agcagctggc cgccctggcc ggcaatcccg ccgccaccgc cgccgtgaaa 1740  
accgccctga gcgaaatcc cgtgccatt ctgatccct gccaccgggt ggtgtctagc 1800  
tctggcggcg tggggggcta cgagggcggg ctcccgctga aagagtggct gctggcccac 1860  
gagggccaca gactgggcaa gcctgggctg ggtcctgcag gtataggcgc gccaggtcc 1920  
ctggagcatc atcatcatca tcattgatga cgggcc 1957

<210> 21  
<211> 960  
5 <212> ADN  
<213> Secuencia artificial

<220>  
10 <223> Secuencia de ADN de BiP/SNAP-IFN/Etiqueta de His

<400> 21  
atgaagtat gcatattact ggccgtcgtg gcctttgttg gcctctcgt cgggagatct 60  
gacaaagact gcgaaatgaa gcgcaccacc ctggatagcc ctctgggcaa gctggaactg 120

ES 2 627 117 T3

tctgggtgcg aacagggcct gcacgagatc aagctgctgg gcaaaggaac atctgccgcc 180  
gacgccgtgg aagtgcctgc cccagccgcc gtgctgggcg gaccagagcc actgatgcag 240  
gccaccgcct ggctcaacgc ctactttcac cagcctgagg ccatcgagga gttccctgtg 300  
ccagccctgc accaccagc gttccagcag gagagcttta cccgccaggt gctgtggaaa 360  
ctgctgaaag tggatgaagt cggagaggtc atcagctacc agcagctggc cgccctggcc 420  
ggcaatcccc cggccaccgc cgccgtgaaa accgccctga gcggaaatcc cgtgcccatt 480  
ctgatcccct gccaccgggt ggtgtctagc tctggcgccg tggggggcta cgagggcggg 540  
ctcgccgtga aagagtggct gctggcccac gagggccaca gactgggcaa gcctgggctg 600  
ggtcctgcag gtatagggcg gccagggtcc ctagggtggc gatctgatga cgatgataaa 660  
gatatctgtg atctccctga gaccacagc ctggataaca ggaggacctt gatgctcctg 720  
gcacaaaatga gcagaatctc tccttcctcc tgtctgatgg acagacatga ctttgattt 780  
cccagggagg agtttgatgg caaccagttc cagaaggtc cagccatctc tgtcctccat 840  
gagctgatcc agcagatctt caacctctt accacaaaag attcatctgc tgcttgggat 900  
gaggacctcc tagacaaatt ctgcaccgaa ctctaccagc agctgaatga cttggaagcc 960

<210> 22  
<211> 957  
5 <212> ADN  
<213> Secuencia artificial

<220>  
10 <223> Secuencia de ADN de SNAP-EDIIWN/Etiqueta de His

<400> 22  
agatctgaca aagactgcga aatgaagcgc accaccctgg atagccctct gggcaagctg 60  
gaactgtctg ggtgcgaaca gggcctgcac gagatcaagc tgctgggcaa aggaacatct 120  
gccgccgacg cgtggaagt gcctgcccc aaccgcctgc tggcgggacc agagccactg 180  
atgcaggcca cgcctggct caacgcctac ttccaccagc ctgaggccat cgaggagtcc 240  
cctgtgccag cctgcacca cccagtgtc cagcaggaga gctttaccg ccaggtgctg 300  
tgaaactgc tgaagtggg gaagttcga gaggtcatca gctaccagca gctggccgcc 360  
ctggccggca atcccgcgc caccgccgc gtgaaaaccg ccctgagcgg aaatcccgtg 420  
ccattctga tcccctgcca ccgggtggtg tctagctctg gcgccgtgg gggctacgag 480  
ggcgggctcg cgtgaaaga gtggctgctg gccacagag gccacagact gggcaagcct 540  
gggctgggtc ctgcaggat aggcgcgcca ggaggtggc ggtctcagtt gaagggaaca 600  
acatggcg tctgttcaaa ggctttcaag tttcttggga ctccgcaga cacaggtcac 660  
ggcactgtgg tgttgaatt gcagtacact ggcacggatg gaccttgcaa agttcctatc 720  
tcgtcagtg cttcattgaa cgacctaac ccagtgggca gattgggtcac tgtcaaccct 780  
tttgttctag tggccacggc caacgctaag gtcctgattg aattggaacc acccttggg 840  
gactcataca tagtgggtgg cagaggagaa caacagatca atcaccattg gcacaagtct 900  
ggaagcagca ttggcaaagg aggtggccat caccatcacc atcactgatg accggtt 957

15 <210> 23  
<211> 398

ES 2 627 117 T3

<212> ADN  
 <213> virus del Nilo Occidental

<220>  
 5 <221> misc\_feature  
 <223> Secuencia de ADN de la proteína ED III del virus WN

<400> 23  
 taggcgcgcc aggaggtggc gggctctcagt tgaaggaac aacctatggc gtctgttcaa 60  
 aggctttcaa gtttcttggg actcccgag acacaggtca cggcactgtg gtggtggaat 120  
 tgcagtacac tggcacggat ggaccttgca aagtctctat ctcgtcagtg gcttcattga 180  
 acgacctaac gccagtgggc agattggtca ctgtcaaccc ttttgtttca gtggccacgg 240  
 ccaacgctaa ggtcctgatt gaattggaac cacccttgg agactcatac atagtgtgg 300  
 gcagaggaga acaacagatc aatcaccatt ggcacaagtc tggagcagc attggcaaa 360  
 gaggtggcca tcaccatcac catcactgat gaccggtt 398

10 <210> 24  
 <211> 398  
 <212> ADN  
 <213> Aislado USU629-05 de virus Usutu

15 <220>  
 <221> misc\_feature  
 <223> Secuencia de ADN de la proteína ED III del virus USU

20 <400> 24  
 taggcgcgcc aggaggtggc gggctctacac taaaaggcac cacctacggc atgtgcacgg 60  
 aaaagttttc ttttgcaaaa aatccggctg acacgggtca cggcactgtg gtccttgaac 120  
 tgcagtacac gggatctgac ggaccttgca aaatcccaat ttccattgtg gcatcacttt 180  
 ccgatctcac cccattggt agaatggtta cagcaaacc ttatgtggct tcatccgaag 240  
 ccaacgcgaa agtgttgggt gagatggaac caccatttgg agattcatac attgtggttg 300  
 gaagagggga taagcagata aaccatcact ggcacaaagc aggaagtcc attggaaaag 360  
 gtggaggcca ccatcacat caccattgat gaccggtt 398

25 <210> 25  
 <211> 410  
 <212> ADN  
 <213> cepa GP05 de virus de la encefalitis japonesa

30 <220>  
 <221> misc\_feature  
 <223> Secuencia de ADN de la proteína ED III del virus EJ

<400> 25  
 taggcgcgcc agggtccctg gagggaggtg gcgggtctgc tctgaaaggc acaacctatg 60  
 gcatgtgtac agaaaaattc tcgttcgcga aaaatccggc ggacactggt cacggaacag 120  
 ttgtcattga actctctac tctgggagt atggcccctg caaaattccg attgtctccg 180  
 tcgcgagcct caatgacatg actcctgttg ggcggctggt gacagtgaac ccctttgtcg 240  
 cggttccag tgccaactca aaggtgctgg tcgagatgga accccccttc ggagactcct 300  
 atatcgtggt tggagggga gacaagcaga tcaaccacca ttggcacaga gctggaagca 360  
 cgctgggcaa gggaggtggc catcaccatc accatcactg atgaccggtt 410

ES 2 627 117 T3

<210> 26  
 <211> 398  
 <212> ADN  
 5 <213> Cepa A 52 del virus Kumlinge de la encefalitis transmitida por garrapatas  
  
 <220>  
 <221> misc\_feature  
 <223> Secuencia de ADN de la proteína ED III del virus ETG  
 10  
 <400> 26  
 taggcgcgcc agggctccctg gagggaggtg gcgggtctct ggaaaaactg aagatgaaag 60  
 gtcttacgta cacaatgtgt gacaaaacaa agttcacatg gaagagagct ccaacagata 120  
 gtgggcatga cacagtggtc atggaagtca cattctctgg aacaaagccc tgtaggatcc 180  
 cagtcagggc agtggcacat ggatctccag atgtgaacgt ggccatgctg ataacgccaa 240  
 accctacaat tgaaaacaat ggaggtggct tcatagagat gcagctgccc ccaggggata 300  
 acatcatcta tgttggggaa ctgagtcacc aatggtcca aaaagggagc agcattggag 360  
 gaggtggcca tcaccatcac catcactgat gaccggtt 398  
  
 <210> 27  
 15 <211> 404  
 <212> ADN  
 <213> Cepa FGA/NA d1d del virus Dengue tipo 1  
  
 <220>  
 20 <221> misc\_feature  
 <223> Secuencia de ADN de la proteína ED III del virus DEN-1  
  
 <400> 27  
 taggcgcgcc agggctccctg gagggaggtg gcgggtctct gactttaaaa gggatgtcat 60  
 atgtgatgtg cacaggctca ttttagctag agaaggaagt ggctgagacc cagcatggaa 120  
 ctgtcctagt gcaggttaaa tacgaaggaa cagatgcgcc atgcaagatc cccttttcga 180  
 cccaagatga gaaaggagtg acccagaatg ggagattgat aacagccaat cccatagtta 240  
 ctgacaaaaga aaaaccaatc aacattgaga cagaaccacc ttttggtgag agctacatca 300  
 tagtaggggc aggtgaaaaa gctttgaaac taagctgggt caagaaagga agcagcatag 360  
 25 ggaaaggagg tggccatcac catcaccatc actgatgacc gggt 404  
  
 <210> 28  
 <211> 404  
 <212> ADN  
 30 <213> virus Dengue-2  
  
 <220>  
 <221> misc\_feature  
 <223> Secuencia de ADN de la proteína ED III de virus DEN-2  
 35 <400> 28

ES 2 627 117 T3

	taggcgcgcc aggttcctg gagggaggtg gcgggtctct acagctcaa ggaatgcat	60
	attctatgtg tacaggaaag tttaaagttg tgaaggaaat agcagaaaca caacatggaa	120
	caatagttct cagagtacaa tatgaagggg acggttctcc gtgcaagatc cttttgaaa	180
	taatggattt ggaaaaaaga catgtcttag gtcgcttgat cacagtcaac ccaattgtta	240
	cagaaataga cagcccagtc aacatagaag cagaacctcc attcggagac agctacatca	300
	ttataggagt agaaccggga caactgaagc tcagctggtt taagaaagga agttccattg	360
	gccaaaggagg tggccatcac catcaccatc actgatgacc ggtt	404
5	<210> 29 <211> 401 <212> ADN <213> virus Dengue-3	
10	<220> <221> misc_feature <223> Secuencia de ADN de la proteína ED III del virus DEN-3	
	<400> 29	
	taggcgcgcc aggttcctg gagggaggtg gcgggtctga actcaagggg atgagctatg	60
	caatgtgctt gaataccttt gtggtgaaga aagaagtctc cgaaacgcag catgggacaa	120
	tactcattaa ggttgagtac aaaggggaag atgcaccttg caagattcct ttctccacag	180
	aggatggaca agggaaagcc cacaatggta gactgatcac agccaacca gtggttacta	240
	agaagagga gcctgtcaac attgaggctg aacctccttt tggggaaagc aacatagtga	300
	ttggagttgg agacaaagcc ttgaaaatta actggtacaa gaagggaaagc tcgattggga	360
	agggaggtgg ccatcaccat caccatcact gatgaccggt t	401
15	<210> 30 <211> 401 <212> ADN <213> virus Dengue-4	
20	<220> <221> misc_feature <223> Secuencia de ADN de la proteína ED III del virus DEN-4	
	<400> 30	
	taggcgcgcc aggttcctg gagggaggtg gcgggtctag aatcaagga atgtcataca	60
	cgatgtgctc aggaaagttc tcaattgaca aagagatggc agaaacacag catgggacaa	120
	cagtggtgaa agtcaagtat gaaggtgctg gagctccgtg taaagtcccc atagagatac	180
	gagatgtaaa taaggaaaaa gtggttgggc gtgtcatctc atccaccctc ctagctgaga	240
	ataccaacag tgtgaccaac atagaactgg aacccccctt tggggacagt tacatagtca	300
	taggtgttgg gaacagtgca ttgacactcc attggttcag gaaaggaagt tctattggca	360
25	agggaggtgg ccatcaccat caccatcact gatgaccggt t	401
30	<210> 31 <211> 410 <212> ADN <213> Cepa ASIBI del virus de la fiebre amarilla	
	<220> <221> misc_feature	

ES 2 627 117 T3

<223> Secuencia de ADN de la proteína ED III del virus Asibi de la FA

<400> 31

```

taggcgcgcc agggtcacct gagggagggtg gcggtcttc agctttgaca ctcaagggga      60
catcctacaa aatgtgacct gacaaaatgt cttttgtcaa gaaccact gacactggcc      120
atggcactgt tgtgatgcag gtgaaagtgc caaaaggagc cccctgcaag attccagtga      180
tagtagctga tgatcttaca gcggaatca ataaaggcat tttggttaca gttaacccca      240
tcgcctcaac caatgatgat gaagtgtgta ttgagtgtaa cccacctttt ggagacagct      300
acattatcgt tgggacagga gattcacgtc tcacttacca gtggcacaaa gagggaaagct      360
caataggaaa gggagggtggc catcaccatc accatcactg atgaccggtt      410

```

5

<210> 32

<211> 595

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Secuencia de ADN del gen sintético IFNAI

<400> 32

```

gcgcgccagg gtccttaggt ggcggatctg atgacgatga taaagatata tgtgatctcc      60
ctgagaccca cagcctggat aacaggagga ccttgatgct cctggcacia atgagcagaa      120
tctctccttc ctctgtctg atggacagac atgactttgg atttcccag gaggagtttg      180
atggcaacca gttccagaag gctccagcca tctctgtcct ccatgagctg atccagcaga      240
tcttcaacct ctttaccaca aaagattcat ctgtgtcttg ggatgaggac ctctagaca      300
aattctgcac cgaactctac cagcagctga atgacttggg agcctgtgtg atgcaggagg      360
agagggtggg agaaactccc ctgatgaatg cggactocat cttggctgtg aagaaatact      420
tccgaagaat cactctctat ctgacagaga agaaatacag cccttgtgcc tgggaggttg      480
tcagagcaga aatcatgaga tcctctctt tatcaacaaa cttgcaagaa agattaagga      540
ggaaggaagg caagtggggc ggtggaagtc atcatcatca tcatcattga ccggt      595

```

15

<210> 33

<211> 469

<212> PRT

<213> Secuencia artificial

20

<220>

<223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-FVR-N + espaciador-etiqueta (His)<sub>6</sub>

25

<400> 33

ES 2 627 117 T3

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
 20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
 50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
 65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
 85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
 100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
 115 120 125

ES 2 627 117 T3

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
 130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
 165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
 180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
 195 200 205

Gly Ser Leu Gly Gly Gly Ser Asp Asn Tyr Gln Glu Leu Arg Val Gln  
 210 215 220

Phe Ala Ala Gln Ala Val Asp Arg Asn Glu Ile Glu Gln Trp Val Arg  
 225 230 235 240

Glu Phe Ala Tyr Gln Gly Phe Asp Ala Arg Arg Val Ile Glu Leu Leu  
 245 250 255

Lys Gln Tyr Gly Gly Ala Asp Trp Glu Lys Asp Ala Lys Lys Met Ile  
 260 265 270

Val Leu Ala Leu Thr Arg Gly Asn Lys Pro Arg Arg Met Met Met Lys  
 275 280 285

Met Ser Lys Glu Gly Lys Ala Thr Val Glu Ala Leu Ile Asn Lys Tyr  
 290 295 300

Lys Leu Lys Glu Gly Asn Pro Ser Arg Asp Glu Leu Thr Leu Ser Arg  
 305 310 315 320

Val Ala Ala Ala Leu Ala Gly Trp Thr Cys Gln Ala Leu Val Val Leu  
 325 330 335

Ser Glu Trp Leu Pro Val Thr Gly Thr Thr Met Asp Gly Leu Ser Pro  
 340 345 350

Ala Tyr Pro Arg His Met Met His Pro Ser Phe Ala Gly Met Val Asp  
 355 360 365

Pro Ser Leu Pro Gly Asp Tyr Leu Arg Ala Ile Leu Asp Ala His Ser  
 370 375 380



ES 2 627 117 T3

Leu Tyr Leu Leu Gln Phe Ser Arg Val Ile Asn Pro Asn Leu Arg Gly  
385 390 395 400

Arg Thr Lys Glu Glu Val Ala Ala Thr Phe Thr Gln Pro Met Asn Ala  
405 410 415

Ala Val Asn Ser Asn Phe Ile Ser His Glu Lys Arg Arg Glu Phe Leu  
420 425 430

Lys Ala Phe Gly Leu Val Asp Ser Asn Gly Lys Pro Ser Ala Ala Val  
435 440 445

Met Ala Ala Ala Gln Ala Tyr Lys Thr Ala Ala Gly Gly Gly Ser His  
450 455 460

His His His His His  
465

<210> 34

<211> 398

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador- IFNa + espaciador-etiqueta (His)<sub>6</sub>

<400> 34

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
50 55 60

Val Pro Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln Ala  
65 70 75 80

Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu Glu  
85 90 95

Phe Pro Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser Phe Thr  
100 105 110

ES 2 627 117 T3

Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly Glu Val  
 115 120 125

Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala Ala Thr  
 130 135 140

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

Pro Cys His Arg Val Ser Ser Ser Gly Ala Val Gly Gly Tyr Glu Gly  
 165 170 175

Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly His Arg Leu  
 180 185 190

Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro Gly Ser Leu  
 195 200 205

Gly Gly Gly Ser Asp Asp Asp Asp Lys Asp Ile Cys Asp Leu Pro Glu  
 210 215 220

Thr His Ser Leu Asp Asn Arg Arg Thr Leu Met Leu Leu Ala Gln Met  
 225 230 235 240

Ser Arg Ile Ser Pro Ser Ser Cys Leu Met Asp Arg His Asp Phe Gly  
 245 250 255

Phe Pro Gln Glu Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala Pro Ala  
 260 265 270

Ile Ser Val Leu His Glu Leu Ile Gln Gln Ile Phe Asn Leu Phe Thr  
 275 280 285

Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Asp Leu Leu Asp Lys Phe  
 290 295 300

Cys Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Met  
 305 310 315 320

Gln Glu Glu Arg Val Gly Glu Thr Pro Leu Met Asn Ala Asp Ser Ile  
 325 330 335

Leu Ala Val Lys Lys Tyr Phe Arg Arg Ile Thr Leu Tyr Leu Thr Glu  
 340 345 350

Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met  
 355 360 365

Arg Ser Leu Ser Leu Ser Thr Asn Leu Gln Glu Arg Leu Arg Arg Lys  
 370 375 380

Glu Gly Lys Trp Gly Gly Gly Ser His His His His His His  
 385 390 395

<210> 35

<211> 333

<212> PRT

<213> Secuencia artificial

5

ES 2 627 117 T3

<220>

<223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-EDIII WN+ espaciador- etiqueta (His)<sub>6</sub>

5

<400> 35

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
115 120 125

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
165 170 175

ES 2 627 117 T3

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
 180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
 195 200 205

Gly Gly Gly Gly Ser Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys Ser  
 210 215 220

Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly Thr  
 225 230 235 240

Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys Val  
 245 250 255

Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly Arg  
 260 265 270

Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala Lys  
 275 280 285

Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Val  
 290 295 300

Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly Ser  
 305 310 315 320

Ser Ile Gly Lys Gly Gly Gly His His His His His His  
 325 330

<210> 36  
 <211> 335  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 10 <223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-EDIII DEN-1+espaciador-etiqueta (His)<sub>6</sub>

<400> 36  
 Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
 20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45

ES 2 627 117 T3

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

<210> 37

5 <211> 335

ES 2 627 117 T3

<212> PRT  
 <213> Secuencia artificial

<220>  
 5 <223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-EDIII DEN-2+espaciador-etiqueta (His)<sub>6</sub>

<400> 37  
 Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15  
 Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
 20 25 30  
 Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45  
 Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
 50 55 60  
 Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
 65 70 75 80  
 Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
 85 90 95  
 Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
 100 105 110  
 Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
 115 120 125  
 Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
 130 135 140  
 Ala Thr Ala Ala Val Lys Thr Ala Leu Ser Gly Asn Pro Val Pro Ile  
 145 150 155 160  
 Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly

ES 2 627 117 T3

				165						170										175
Tyr	Glu	Gly	Gly	Leu	Ala	Val	Lys	Glu	Trp	Leu	Leu	Ala	His	Glu	Gly					
			180					185					190							
His	Arg	Leu	Gly	Lys	Pro	Gly	Leu	Gly	Pro	Ala	Gly	Ile	Gly	Ala	Pro					
		195					200					205								
Gly	Ser	Leu	Glu	Gly	Gly	Gly	Gly	Ser	Leu	Gln	Leu	Lys	Gly	Met	Ser					
	210					215					220									
Tyr	Ser	Met	Cys	Thr	Gly	Lys	Phe	Lys	Val	Val	Lys	Glu	Ile	Ala	Glu					
225					230					235					240					
Thr	Gln	His	Gly	Thr	Ile	Val	Leu	Arg	Val	Gln	Tyr	Glu	Gly	Asp	Gly					
				245					250					255						
Ser	Pro	Cys	Lys	Ile	Pro	Phe	Glu	Ile	Met	Asp	Leu	Glu	Lys	Arg	His					
			260					265					270							
Val	Leu	Gly	Arg	Leu	Ile	Thr	Val	Asn	Pro	Ile	Val	Thr	Glu	Ile	Asp					
		275					280					285								
Ser	Pro	Val	Asn	Ile	Glu	Ala	Glu	Pro	Pro	Phe	Gly	Asp	Ser	Tyr	Ile					
		290				295					300									
Ile	Ile	Gly	Val	Glu	Pro	Gly	Gln	Leu	Lys	Leu	Ser	Trp	Phe	Lys	Lys					
305					310					315					320					
Gly	Ser	Ser	Ile	Gly	Gln	Gly	Gly	Gly	His	His	His	His	His	His	His					
				325					330						335					

<210> 38  
 <211> 334  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 10 <223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-EDIII DEN-3+espaciador-etiqueta (His)<sub>6</sub>

Met	Lys	Leu	Cys	Ile	Leu	Leu	Ala	Val	Val	Ala	Phe	Val	Gly	Leu	Ser					
1				5					10					15						
Leu	Gly	Arg	Ser	Asp	Lys	Asp	Cys	Glu	Met	Lys	Arg	Thr	Thr	Leu	Asp					
			20					25					30							
Ser	Pro	Leu	Gly	Lys	Leu	Glu	Leu	Ser	Gly	Cys	Glu	Gln	Gly	Leu	His					

ES 2 627 117 T3

35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
115 120 125

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
195 200 205

Gly Ser Leu Glu Gly Gly Gly Ser Glu Leu Lys Gly Met Ser Tyr  
210 215 220

Ala Met Cys Leu Asn Thr Phe Val Leu Lys Lys Glu Val Ser Glu Thr  
225 230 235 240

Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr Lys Gly Glu Asp Ala  
245 250 255

Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly Gln Gly Lys Ala His  
260 265 270

Asn Gly Arg Leu Ile Thr Ala Asn Pro Val Val Thr Lys Lys Glu Glu  
275 280 285

Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile Val  
290 295 300

Ile Gly Val Gly Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys Gly  
305 310 315 320

Ser Ser Ile Gly Lys Gly Gly Gly His His His His His His  
325 330

<210> 39



ES 2 627 117 T3

<211> 334  
 <212> PRT  
 <213> Secuencia artificial

5 <220>  
 <223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-EDIII DEN-4+espaciador-etiqueta (His)<sub>6</sub>

<400> 39  
 Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15  
 Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
 20 25 30  
 Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45  
 Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
 50 55 60  
 Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
 65 70 75 80  
 Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
 85 90 95  
 Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
 100 105 110  
 Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
 115 120 125  
 Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
 130 135 140  
 Ala Thr Ala Ala Val Lys Thr Ala Leu Ser Gly Asn Pro Val Pro Ile  
 145 150 155 160

10

ES 2 627 117 T3

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
 165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
 180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
 195 200 205

Gly Ser Leu Glu Gly Gly Gly Gly Ser Arg Ile Lys Gly Met Ser Tyr  
 210 215 220

Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met Ala Glu Thr  
 225 230 235 240

Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly Ala Gly Ala  
 245 250 255

Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys Glu Lys Val  
 260 265 270

Val Gly Arg Val Ile Ser Ser Thr Pro Leu Ala Glu Asn Thr Asn Ser  
 275 280 285

Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val  
 290 295 300

Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe Arg Lys Gly  
 305 310 315 320

Ser Ser Ile Gly Lys Gly Gly Gly His His His His His His  
 325 330

<210> 40

<211> 338

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-EDIII EJ+espaciador- etiqueta (His)<sub>6</sub>

<400> 40

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
 20 25 30

ES 2 627 117 T3

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
 50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
 65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
 85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
 100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
 115 120 125

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
 130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
 165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
 180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
 195 200 205

Gly Ser Leu Glu Gly Gly Gly Gly Ser Ser Ala Leu Lys Gly Thr Thr  
 210 215 220

Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys Asn Pro Ala Asp  
 225 230 235 240

Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr Ser Gly Ser Asp  
 245 250 255

Gly Pro Cys Lys Ile Pro Ile Val Ser Val Ala Ser Leu Asn Asp Met  
 260 265 270

Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ala Ala Ser  
 275 280 285

ES 2 627 117 T3

Ser Ala Asn Ser Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly Asp  
 290 295 300

Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn His His Trp  
 305 310 315 320

His Arg Ala Gly Ser Thr Leu Gly Lys Gly Gly Gly His His His His  
 325 330 335

His His

<210> 41

<211> 333

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-EDIII TBE + espaciador-etiqueta (His)<sub>6</sub>

<400> 41

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
 20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
 50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
 65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
 85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
 100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
 115 120 125

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
 130 135 140

ES 2 627 117 T3

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
195 200 205

Gly Ser Leu Glu Gly Gly Gly Gly Ser Leu Glu Lys Leu Lys Met Lys  
210 215 220

Gly Leu Thr Tyr Thr Met Cys Asp Lys Thr Lys Phe Thr Trp Lys Arg  
225 230 235 240

Ala Pro Thr Asp Ser Gly His Asp Thr Val Val Met Glu Val Thr Phe  
245 250 255

Ser Gly Thr Lys Pro Cys Arg Ile Pro Val Arg Ala Val Ala His Gly  
260 265 270

Ser Pro Asp Val Asn Val Ala Met Leu Ile Thr Pro Asn Pro Thr Ile  
275 280 285

Glu Asn Asn Gly Gly Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp  
290 295 300

Asn Ile Ile Tyr Val Gly Glu Leu Ser His Gln Trp Phe Gln Lys Gly  
305 310 315 320

Ser Ser Ile Gly Gly Gly Gly His His His His His His  
325 330

<210> 42

<211> 337

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de la proteína de fusiónBIP-SNAP-espaciador-EDIII USU+espaciador-etiqueta (His)<sub>6</sub>

<400> 42

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

ES 2 627 117 T3

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

ES 2 627 117 T3

Pro Ile Gly Arg Met Val Thr Ala Asn Pro Tyr Val Ala Ser Ser Glu  
 275 280 285

Ala Asn Ala Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly Asp Ser  
 290 295 300

Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn His His Trp His  
 305 310 315 320

Lys Ala Gly Ser Ser Ile Gly Lys Gly Gly His His His His His  
 325 330 335

His

<210> 43

<211> 337

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de la proteína de fusión BiP-SNAP-espaciador-EDIII FA+espaciador- etiqueta (His)<sub>6</sub>

<400> 43

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
 20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
 50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
 65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
 85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
 100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
 115 120 125

ES 2 627 117 T3

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
 130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
 165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
 180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
 195 200 205

Gly Ser Leu Glu Gly Gly Gly Gly Ser Ser Ala Leu Thr Leu Lys Gly  
 210 215 220

Thr Ser Tyr Lys Met Cys Thr Asp Lys Met Ser Phe Val Lys Asn Pro  
 225 230 235 240

Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val Lys Val Pro Lys  
 245 250 255

Gly Ala Pro Cys Lys Ile Pro Val Ile Val Ala Asp Asp Leu Thr Ala  
 260 265 270

Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro Ile Ala Ser Thr  
 275 280 285

Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro Phe Gly Asp Ser  
 290 295 300

Tyr Ile Ile Val Gly Thr Gly Asp Ser Arg Leu Thr Tyr Gln Trp His  
 305 310 315 320

Lys Glu Gly Ser Ser Ile Gly Lys Gly Gly Gly His His His His His  
 325 330 335

His

<210> 44

<211> 234

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos del polipéptido de fusión: péptido de señal BiP+SNAP+sitio de enterocinasa+EcoRV/Xmal+Etiqueta de His

<400> 44



ES 2 627 117 T3

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
 20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
 50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
 65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
 85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
 100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
 115 120 125

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
 130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
 165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
 180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
 195 200 205

Gly Ser Leu Gly Gly Gly Ser Asp Asp Asp Asp Lys Asp Ile Lys Asn  
 210 215 220

Pro Gly Gly Gly Ser His His His His His  
 225 230

- 5 <210> 45
- <211> 211
- <212> PRT
- <213> Mus musculus
- <400> 45

ES 2 627 117 T3

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

Gly Lys Met Glu Leu Ser Gly Cys Glu Arg Gly Leu His Gly Ile Arg  
 20 25 30

Leu Leu Ser Gly Lys Thr Pro Asn Thr Asp Pro Thr Glu Ala Pro Ala  
 35 40 45

Thr Pro Glu Val Leu Gly Gly Pro Glu Gly Val Pro Glu Pro Leu Val  
 50 55 60

Gln Cys Thr Ala Trp Leu Glu Ala Tyr Phe Arg Glu Pro Ala Ala Thr  
 65 70 75 80

Glu Gly Leu Pro Leu Pro Ala Leu His His Pro Val Phe Gln Gln Asp  
 85 90 95

Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe  
 100 105 110

Gly Glu Thr Val Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro  
 115 120 125

Lys Ala Ala Arg Ala Val Gly Gly Ala Met Arg Ser Asn Pro Val Pro  
 130 135 140

Ile Leu Ile Pro Cys His Arg Val Val Arg Ser Asp Gly Ala Ile Gly  
 145 150 155 160

His Tyr Ser Gly Gly Gly Gln Ala Val Lys Glu Trp Leu Leu Ala His  
 165 170 175

Glu Gly Ile Pro Thr Gly Gln Pro Ala Ser Lys Gly Leu Gly Leu Thr  
 180 185 190

Gly Thr Trp Leu Lys Ser Ser Phe Glu Ser Thr Ser Ser Glu Pro Ser  
 195 200 205

Gly Arg Asn  
 210

- 5 <210> 46
- <211> 209
- <212> PRT
- <213> Ratus norvegicus
- <400> 46

ES 2 627 117 T3

Met Ala Glu Ile Cys Lys Met Lys Tyr Thr Val Leu Asp Ser Pro Leu  
 1 5 10 15

Gly Lys Ile Glu Leu Ser Gly Cys Glu Arg Gly Leu His Gly Ile Arg  
 20 25 30

Phe Leu Ser Gly Lys Thr Pro Asn Thr Asp Pro Thr Glu Ala Pro Ala  
 35 40 45

Cys Pro Glu Val Leu Gly Gly Pro Glu Gly Val Pro Glu Pro Leu Val  
 50 55 60

Gln Cys Thr Ala Trp Leu Glu Ala Tyr Phe His Glu Pro Ala Ala Thr  
 65 70 75 80

Glu Gly Leu Pro Leu Pro Ala Leu His His Pro Val Phe Gln Gln Asp  
 85 90 95

Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe  
 100 105 110

Gly Glu Met Val Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro  
 115 120 125

Lys Ala Ala Arg Ala Val Gly Gly Ala Met Arg Ser Asn Pro Val Pro  
 130 135 140

Ile Leu Ile Pro Cys His Arg Val Ile Arg Ser Asp Gly Ala Ile Gly  
 145 150 155 160

Asn Tyr Ser Gly Gly Gly Gln Thr Val Lys Glu Trp Leu Leu Ala His  
 165 170 175

Glu Gly Ile Pro Thr Gly Gln Pro Ala Ser Lys Gly Leu Gly Leu Ile  
 180 185 190

Gly Ser Trp Leu Lys Pro Ser Phe Glu Ser Ser Ser Pro Lys Pro Ser  
 195 200 205

Gly

<210> 47

<211> 555

5 <212> ADN

<213> Secuencia artificial

<220>

<223> ADN codificante de SNAP contenido bajo de G/C

10

<400> 47

ES 2 627 117 T3

```

attgaattta cagacaaaga ctgCGAAATG AAAAGAACTA cATTGGATTC accacttggg      60
aagttggaac tgagtggatg cGAGCAAGGA ttGcatgaaa ttaagcttct gggaaaagga      120
acttctgcag ctgatgcagt tgaagttcca gcaccagcag ctgttcttgg aggtcctgag      180
cccctcatgc aagccacagc ctggcttaac gcatatttcc accagcctga ggccattgag      240
gaatttccag tccccgcctt tcaccatcct gtgtttcagc aggaaagctt caccCGCCAG      300
gtcctgtgga aattgctgaa ggtggtcaag tttggtgaag tgatttcata tcagcaactt      360
gctgcattgg cggtaaccc cgcagctaca gctgcCGTGA aaactgctct cagCGGAAAT      420
cctgtGCCA tcctgatccc ttgtcacaga gtCGTTTCAT cttCGGAGC tgtaggtggc      480
tatgaaggag gactggcagt taaggagtgg ctgctggctc atgaaggTCA tagacttggA      540
aacctggtt tggga                                                              555

<210> 48
<211> 18
5 <212> PRT
   <213> Secuencia artificial

<220>
10 <223> Secuencia de aminoácidos de BiP

<400> 48
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
1           5           10           15

Leu Gly

15 <210> 49
   <211> 738
   <212> ADN
   <213> virus TOSCANA

20 <400> 49
atggacaact atcaagagct tcgagtccag tttgctgctc aagcagtgga cgcCaatgag      60
attgaacagt ggttccgaga gtttgcttat caagggtttg atgcccgtag agttatcgaa      120
ctcttaaagc agtatggtgg ggctgactgg gagaaggatg ccaagaaaat gattgttctg      180
gctctaactc gtggcaacaa gcccaggagg atgatgatga aaatgtcgaa agaaggcaaa      240
gcaactgtgg aggtctctcat caacaagtat aagctaaagg aagggaatcc ttcccgggat      300
gagttgactc tatcacgagt tgctgcCGCC ttggctggct ggacatGCCA ggctttggtc      360
gtcttgagtg agtggcttcc tgtcactggg actaccatgg acggcctatc ccctgcatac      420
ccgaggcata tgatgcaccc cagctttgct ggcatggtgg atccttctct accaggagac      480
tatctaaggg caatattaga tgctactctc ctgtatctgc tgcagttctc cggggtcAtc      540
aaccCAAACC tccgaggtag aacaaaagag gaggttgctg caacgttcac gcagccaatg      600
aatgcagcag tgaatagcaa ctttataagc catgagaaga ggagagaatt cttgaaagcc      660
tttgacttg tggattccaa tgggaagccg tcagctgctg tcatggcagc cgctcaggct      720
tacaagacag cagcctaa                                                              738

25 <210> 50
   <211> 66
   <212> ADN

```

ES 2 627 117 T3

<213> Secuencia artificial

<220>

<223> Señal de tipo BiP: péptido de señal de insecto BiP extendida para células de mamífero

5

<400> 50  
 atgaaactat gtattctact tgcagttggt gcgttcgtag gattgtcctt acctacagct 60  
 ctggca 66

<210> 51  
 <211> 22  
 <212> PRT  
 <213> Secuencia artificial

<220>

15 <223> Señal de tipo BiP: péptido de señal de insecto BiP extendida para células de mamífero

<400> 51  
 Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15  
 Leu Pro Thr Ala Leu Ala  
 20

20 <210> 52  
 <211> 21  
 <212> ADN  
 <213> virus del grabado del tabaco

25 <400> 52  
 gaaaacctgt acttccagag c 21

<210> 53  
 <211> 7  
 <212> PRT  
 <213> Virus del grabado del tabaco

<400> 53  
 Glu Asn Leu Tyr Phe Gln Ser  
 1 5

35 <210> 54  
 <211> 699  
 <212> ADN  
 <213> Homo sapiens

40 <400> 54

ES 2 627 117 T3

gatatcatag gaggtcgaga agttattccc cactcacgcc cttacatggc atcacttcag 60  
 agaaatgggt cccacctatg cgggtggtgta ctagtccacc caaagtgggt tctaaccgca 120  
 gctcactgcc ttgccagcg gatggctcag ctaaggcttg tacttgact tcacacccta 180  
 gacagccccg gtctcacctt ccacatcaag gcagctatcc agcaccctcg atacaagcca 240  
 gtacctgcac ttgagaacga cctagctcta cttcagctag acggtaaagt aaagcctagc 300  
 cggaccatcc gaccgttggc tctacctagt aagcgcagg tagttgcagc aggtactcgg 360  
 tgcagcatgg caggctgggg acttaccac cagggtggac gcctttcccg agtacttcgg 420  
 gagctagacc ttcaagtact ggacaccgc atgtgtaaca acagccgctt ttggaacgga 480  
 agcctatccc caagcatggt ttgcctagca gctgactcca aggaccagc tcctgcaag 540  
 ggtgactcgg gtggaccct ggtttgtggc aaaggccggg tgtagccgg agttcttcc 600  
 ttcagctcca gggatgcac tgacatctc aagcctccag ttgcaaccgc tgttgcaact 660  
 tacgtttcct ggatcaggaa ggtcaccggt cgatcggcc 699

<210> 55  
 <211> 1400  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 10 <223> Secuencia de ADN de proteína de tipo BiP de insecto/SNAP/enterocinasa/GrM/Etiqueta de His para células S2 (fig. 6)

<400> 55  
 atgaagttat gcatattact ggccgtcgtg gcctttggtg gcctctcgtc cgggagatct 60  
 gacaaagact gcgaaatgaa gcgcaccacc ctggatagcc ctctgggcaa gctggaactg 120  
 tctgggtgcg aacagggcct gcacgagatc aagctgctgg gcaaaggaac atctgccgcc 180  
 gacgccgtgg aagtgcctgc cccagccgcc gtgctgggcg gaccagagcc actgatgcag 240  
 gccaccgcct ggtcaacgc ctactttcac cagcctgagg ccatcgagga gttccctgtg 300

ES 2 627 117 T3

ccagccctgc accaccaggt gttccagcag gagagcttta cccgccaggt gctgtgaaaa 360  
ctgctgaaag tggatgaagt cggagaggct atcagctacc agcagctggc cgccctggcc 420  
ggcaatcccg ccgccaccgc cgcctgaaa accgccctga gcggaaatcc cgtgccatt 480  
ctgatcccct gccaccgggt ggtgtctagc tctggcgcg tggggggcta cgagggcggg 540  
ctcgcctgga aagagtggct gctggccac gagggccaca gactgggcaa gcctgggctg 600  
ggtcctgcag gtataggcgc gccagggtcc ctagggtggc gatctgatga cgatgataaa 660  
gatatcatag gaggtcgaga agttattccc cactcacgcc cttacatggc atcacttcag 720  
agaaatgggt cccacctatg cgggtggtga ctagtccacc caaagtgggt tctaaccgca 780  
gctcactgcc ttgccagcgc gatggctcag ctaaggcttg tacttggact tcacacccta 840  
gacagccccg gtctcacctt ccacatcaag gcagctatcc agcaccctcg atacaagcca 900  
gtacctgcac ttgagaacga cctagctcta cttcagctag acggtaaagt aaagcctagc 960  
cggaccatcc gaccgttggc tctacctagt aagcgcagg tagttgcagc aggtactcgg 1020  
tgcagcatgg caggctgggg acttaccac cagggtggac gcctttccc agtacttcgg 1080  
gagctagacc ttcaagtact ggacaccgc atgtgtaaca acagccgctt ttggaacgga 1140  
agcctatccc caagcatggt ttgcctagca gctgactcca aggaccaggc tccctgcaag 1200  
ggtgactcgg gtggaccct ggtttgtggc aaagccggg tgtagccgg agttctttcc 1260  
ttcagctcca gggtatgac tgacatctc aagcctccag ttgcaaccgc tgttgacact 1320  
tacgtttcct ggatcaggaa ggtcaccggt cgatcggccc cgggcgggtg aagtcacat 1380  
catcatcacc attgaccggt 1400

<210> 56

<211> 434

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de proteína de tipo BiP/SNAP/enterocinasa/GrM/Etiqueta de His para células S2 (fig. 6)

<400> 56

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu

ES 2 627 117 T3

50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
115 120 125

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
195 200 205

Gly Ser Leu Gly Gly Gly Ser Asp Asp Asp Asp Lys Asp Ile Ile Gly  
210 215 220

Gly Arg Glu Val Ile Pro His Ser Arg Pro Tyr Met Ala Ser Leu Gln  
225 230 235 240

Arg Asn Gly Ser His Leu Cys Gly Gly Val Leu Val His Pro Lys Trp  
245 250 255

Val Leu Thr Ala Ala His Cys Leu Ala Gln Arg Met Ala Gln Leu Arg  
260 265 270

Leu Val Leu Gly Leu His Thr Leu Asp Ser Pro Gly Leu Thr Phe His  
275 280 285

Ile Lys Ala Ala Ile Gln His Pro Arg Tyr Lys Pro Val Pro Ala Leu  
290 295 300



ES 2 627 117 T3

Glu Asn Asp Leu Ala Leu Leu Gln Leu Asp Gly Lys Val Lys Pro Ser  
305 310 315 320

Arg Thr Ile Arg Pro Leu Ala Leu Pro Ser Lys Arg Gln Val Val Ala  
325 330 335

Ala Gly Thr Arg Cys Ser Met Ala Gly Trp Gly Leu Thr His Gln Gly  
340 345 350

Gly Arg Leu Ser Arg Val Leu Arg Glu Leu Asp Leu Gln Val Leu Asp  
355 360 365

Thr Arg Met Cys Asn Asn Ser Arg Phe Trp Asn Gly Ser Leu Ser Pro  
370 375 380

Ser Met Val Cys Leu Ala Ala Asp Ser Lys Asp Gln Ala Pro Cys Lys  
385 390 395 400

Gly Asp Ser Gly Gly Pro Leu Val Cys Gly Lys Gly Arg Val Leu Ala  
405 410 415

Gly Val Leu Ser Phe Ser Ser Arg Val Cys Thr Asp Ile Phe Lys Pro  
420 425 430

Pro Val

<210> 57

<211> 1243

5 <212> ADN

<213> Secuencia artificial

<220>

10 <223> Secuencia de ADN de proteína de tipo BiP/tipo SNAP/proTEV/IFN/Etiqueta de His para células de mamífero tales como células HeLa (fig. 7)

<400> 57

gctagcacca tgaagttatg catattactg gccgtcgtgg cctttgttgg cctctcgctc	60
ccaacagctc tggcagacaa agactgcgaa atgaaaagaa ctacattgga ttcaccactt	120
gggaagttgg aactgagtgg atgcgagcaa ggattgcatg aaattaagct actgggaaaa	180
ggaacttctg ctgctgatgc agttgaagtt ccagcaccag cagctgttct tggaggtcct	240
gagccccctca tgcaagccac agcctggctt aacgcatatt tccaccagcc tgaggccatt	300
gaggaatttc cagtccccgc ccttcacat cctgtgtttc agcaggagag cttcaccgc	360
caggtcctgt ggaaattgct gaaggtggtc aagtttggtg aagtgatttc atatcagcaa	420
cttgctgcat tggccggtaa ccccgagct acagctgccg tgaaaactgc tctcagcgga	480
aatcctgtgc ccactcctgat cccttgtcac agagtgcgtt catcttccgg agctgtaggt	540

ES 2 627 117 T3

```

ggctatgaag gaggactggc agttaaggag tggctgctgg ctcatgaag tcatagactt      600
ggaaagcctg ggctgggtcc tgctggtata ggcgcgccag ggtocctag tggcggatct      660
gaaaacctgt acttccagag cgatatctgt gatctccctg agaccacag cctggataac      720
aggaggacct tgatgctcct ggcacaaatg agcagaatct ctccttcctc ctgtctgatg      780
gacagacatg actttggatt tcccaggag gagtttgatg gcaaccagtt ccagaaggct      840
ccagccatct ctgtcctcca tgagctgatc cagcagattt tcaacctctt taccacaaaa      900
gattcatctg ctgcttggga tgaggacctc ctagacaaat tctgcaccga actctaccag      960
cagctgaatg acttgggaagc ctgtgtgatg caggaggaga gggtgggaga aactcccctg     1020
atgaatgagg actccatctt ggctgtgaag aaatacttcc gaagaatcac tctctatctg     1080
acagagaaga aatacagccc ttgtgcctgg gaggttgatc gagcagaaat catgagatcc     1140
ctctctttat caacaaactt gcaagaaaga ttaaggagga aggaaggcaa gtggcccggg     1200
ggtggaagtc atcatcatca tcatcattga agcttgcggc cgc                          1243

```

<210> 58  
 <211> 406  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Secuencia de aminoácidos de proteína de tipo BiP/SNAP-like/proTEV/IFN/Etiqueta de His para células de mamífero, tales como células HeLa (fig. 7)

<400> 58  
 Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15  
  
 Leu Pro Thr Ala Leu Ala Asp Lys Asp Cys Glu Met Lys Arg Thr Thr  
 20 25 30  
  
 Leu Asp Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly  
 35 40 45  
  
 Leu His Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala  
 50 55 60  
  
 Val Glu Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu  
 65 70 75 80  
  
 Met Gln Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala  
 85 90 95  
  
 Ile Glu Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln  
 100 105 110

ES 2 627 117 T3

Glu Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys  
 115 120 125  
 Phe Gly Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn  
 130 135 140  
 Pro Ala Ala Thr Ala Ala Val Lys Thr Ala Leu Ser Gly Asn Pro Val  
 145 150 155 160  
 Pro Ile Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val  
 165 170 175  
 Gly Gly Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His  
 180 185 190  
 Glu Gly His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly  
 195 200 205  
 Ala Pro Gly Ser Leu Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln Ser  
 210 215 220  
 Asp Ile Cys Asp Leu Pro Glu Thr His Ser Leu Asp Asn Arg Arg Thr  
 225 230 235 240  
 Leu Met Leu Leu Ala Gln Met Ser Arg Ile Ser Pro Ser Ser Cys Leu  
 245 250 255  
 Met Asp Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly Asn  
 260 265 270  
 Gln Phe Gln Lys Ala Pro Ala Ile Ser Val Leu His Glu Leu Ile Gln  
 275 280 285  
 Gln Ile Phe Asn Leu Phe Thr Thr Lys Asp Ser Ser Ala Ala Trp Asp  
 290 295 300  
 Glu Asp Leu Leu Asp Lys Phe Cys Thr Glu Leu Tyr Gln Gln Leu Asn  
 305 310 315 320  
 Asp Leu Glu Ala Cys Val Met Gln Glu Glu Arg Val Gly Glu Thr Pro  
 325 330 335  
 Leu Met Asn Ala Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Arg Arg  
 340 345 350  
 Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu

ES 2 627 117 T3

	355		360		365														
Val	Val	Arg	Ala	Glu	Ile	Met	Arg	Ser	Leu	Ser	Leu	Ser	Thr	Asn	Leu				
	370					375					380								
Gln	Glu	Arg	Leu	Arg	Arg	Lys	Glu	Gly	Lys	Trp	Pro	Gly	Gly	Gly	Ser				
385				390						395					400				
His	His	His	His	His	His														
				405															

- 5 <210> 59  
 <211> 798  
 <212> ADN  
 <213> Secuencia artificial
- 10 <220>  
 <223> Secuencia de ADN de deSNAP Univ:proteína tipo BiP/tipo SNAP/proTEV/MCS/proTEV/Etiqueta de His (fig. 8)
- <400> 59  
 gatcgcgagc tagcaccatg aaactatgta ttctacttgc agttgttgcg ttcgtaggat 60  
 tgtccttacc tacagctctg gcaagatctg acaaagactg cgaaatgaaa agaactacat 120  
 tggattcacc acttgggaag ttggaactga gtggatgcga gcaaggattg catgaaatta 180  
 agctactggg aaaaggaact tctgctgctg atgcagttga agttccagca ccagcagctg 240  
 ttcttgaggg tcttgagccc ctcatgcaag ccacagcctg gcttaacgca tatttccacc 300  
 agcctgaggc cattgaggaa tttccagtcc ccgcccttca ccatcctgtg tttcagcagg 360  
 agagcttcac ccgccaggtc ctgtggaat tgctgaaggt ggtcaagttt ggtgaagtga 420  
 tttcatatca gcaacttgct gcattggccg gtaaccccgc agctacagct gccgtgaaaa 480  
 ctgctctcag cggaaatcct gtgcccatcc tgatcccttg tcacagagtc gtttcatctt 540  
 ccggagctgt agtgggctat gaaggaggac tggcagttaa ggagtggctg ctggctcatg 600  
 aaggatcatag acttggaag cctgggctgg gtctgctgg tataggcgcg ccagggtccc 660  
 taggtggcgg atccgaaaac ctgtacttcc agagcgatat cggaggtgga ggcccgggag 720  
 agaatctata ttttcaaggg cccggcggag gtagtcacca tcatcaccat cactaatgac 780  
 cggtgccggc gcaagctt 798
- 15 <210> 60  
 <211> 252  
 <212> PRT  
 <213> Secuencia artificial
- 20 <220>  
 <223> Secuencia de aminoácidos de deSNAP Univ:proteína BiP/tipo SNAP/proTEV/MCS/proTEV/Etiqueta de His (fig. 8)
- <400> 60

ES 2 627 117 T3

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Pro Thr Ala Leu Ala Arg Ser Asp Lys Asp Cys Glu Met Lys Arg  
 20 25 30

Thr Thr Leu Asp Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu  
 35 40 45

Gln Gly Leu His Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala  
 50 55 60

Asp Ala Val Glu Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu  
 65 70 75 80

Pro Leu Met Gln Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro  
 85 90 95

Glu Ala Ile Glu Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe  
 100 105 110

Gln Gln Glu Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val  
 115 120 125

Val Lys Phe Gly Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala  
 130 135 140

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

Pro Val Pro Ile Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly  
 165 170 175

Ala Val Gly Gly Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu  
 180 185 190

Ala His Glu Gly His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly  
 195 200 205

Ile Gly Ala Pro Gly Ser Leu Gly Gly Gly Ser Glu Asn Leu Tyr Phe  
 210 215 220

Gln Ser Asp Ile Gly Gly Gly Gly Pro Gly Glu Asn Leu Tyr Phe Gln  
 225 230 235 240

Gly Pro Gly Gly Gly Ser His His His His His His  
 245 250

- <210> 61
- 5 <211> 5
- <212> PRT
- <213> Secuencia artificial
- <220>
- 10 <223> Sitio de corte de aminoácidos del precursor de membrana prM del virus Dengue serotipo 1
- <400> 61

ES 2 627 117 T3

```

    Pro Thr Ala Leu Ala
    1           5

<210> 62
<211> 6
5 <212> PRT
   <213> Secuencia artificial

<220>
<223> Sitio de corte de aminoácidos de enterocinasa
10

<400> 62
   Asp Asp Asp Asp Lys Asp
   1           5

<210> 63
15 <211> 4
   <212> PRT
   <213> Secuencia artificial

<220>
20 <223> Secuencia espaciadora de aminoácidos

<400> 63
   Gly Gly Gly Ser
   1

25 <210> 64
   <211> 3461
   <212> ADN
   <213> Secuencia artificial

30 <220>
   <223> DeSNAP-Secuencia universal insertada en el plásmido pUC57

<400> 64
   tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca      60
   cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg      120
   ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc      180
   accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagggcgc      240
35

```

ES 2 627 117 T3

attcgccatt caggetgccc aactgttggg aagggcgatc ggtgcgggccc tcttcgctat 300  
 tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360  
 tttcccagtc acgacgttgt aaaacgacgg ccagtggaatt cgagctcggg acctcgcgaa 420  
 tgcactataga ttcgcgagct agcaccatga aactatgtat tctacttgcg gttgttgcgt 480  
 tcgtaggatt gtccttacct acagctctgg caagatctga caaagactgc gaaatgaaaa 540  
 gaactacatt ggattcacca cttgggaagt tggaaactgag tggatgcgag caaggattgc 600  
 atgaaattaa gctactggga aaaggaactt ctgctgctga tgcagttgaa gttccagcac 660  
 cagcagctgt tcttgagggt cctgagcccc tcatgcaagc cacagcctgg cttaacgcat 720  
 atttccacca gcctgaggcc attgaggaat ttccagctcc cgccttcac catcctgtgt 780  
 ttcagcagga gagcttcacc cgccaggctc tgtggaaatt gctgaagggt gtcaagtttg 840  
 gtgaagtgat ttcatacag caacttgcct cattggccgg taaccccgca gctacagctg 900  
 ccgtgaaaac tgctctcagc ggaaatcctg tgccatcct gatccctgt cacagagtgc 960  
 tttcatcttc cggagctgta ggtggctatg aaggaggact ggcagttaag gaggggctgc 1020  
 tggctcatga aggtcataga cttggaagc ctgggctggg tctgctggt ataggcgcgc 1080  
 cagggtcctt aggtggcgga tccgaaaacc tgtactcca gagcgatata ggaggtggag 1140  
 gcccgggaga gaatctatat tttcaagggc ccggcggagg tagtcacat catcaccatc 1200  
 actaatgacc ggtgcggccg caagcttggc gtaatcatgg tcatagctgt ttctgtgtg 1260  
 aaattgttat ccgctcacia tccacacia catacgagcc ggaagcataa agtgtaaagc 1320  
 ctggggtgcc taatgagtga gctaactcac attaattgog ttgcgctcac tgcgccttt 1380  
 ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg 1440  
 cggtttgcgt attggcgct cttccgcttc ctgcctcact gactcgtgc gctcggctcgt 1500  
 tcggctgcgg cgagcgggat cagctcactc aaaggcggta atacggttat ccacagaatc 1560  
 aggggataac gcagaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 1620  
 aaaggccgcg ttgctggcgt ttttccatag gctccgccc cctgacgagc atcacaataa 1680  
 tcgacgctca agtcagaggt ggcgaaacc gacaggacta taaagatacc aggcgtttcc 1740  
 ccctggaagc tccctcgtgc gctctcctgt tccgacctg ccgcttaccg gatacctgtc 1800  
 cgcctttctc ccttcgggaa gcgtggcgt ttctcatagc tcacgctgta ggtatctcag 1860  
 ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcaac gaacccccg ttcagccga 1920  
 ccgctgcgcc ttatccggtg actatcgtct tgagtccaac ccgtaagac acgacttatc 1980  
 gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac 2040  
 agagttcttg aagtgtggc etaactcgg ctacaactaga agaacagtat ttggtatctg 2100  
 cgctctgctg aagccagtta ccttcgaaa aagagttggt agctcttgat ccggcaaaaca 2160

ES 2 627 117 T3

aaccaccgct ggtagcgggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 2220  
 aggatctcaa gaagatcctt tgatcttttc tacgggggtct gacgctcagt ggaacgaaaa 2280  
 ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt 2340  
 aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag 2400  
 ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat 2460  
 agttgcctga ctccccgtcg tgtagataac tacgatacgg gagggccttac catctggccc 2520  
 cagtgtcgca atgataccgc gagaccacag ctcaccggct ccagatttat cagcaataaa 2580  
 ccagccagcc ggaagggcgg agcgcagaag tggctctgca actttatccg cctccatcca 2640  
 gtctattaat tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa 2700  
 cgttgttgcc attgctacag gcatcgtggt gtcacgctcg tcgtttggtg tggttcatt 2760  
 cagctccggt tcccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc 2820  
 ggtagctcc ttcggctctc cgatcgttgt cagaagtaag ttggccgag tgttatcact 2880  
 catggttatg gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgcttttc 2940  
 tgtgactggt gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg 3000  
 ctcttggccg gcgtcaatac gggataatac cgcgccacat agcagaactt taaaagtgct 3060  
 catcattgga aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc 3120  
 cagttcgatg taaccactc gtgcacccaa ctgatcttca gcatctttta ctttcaccag 3180  
 cgtttctggg tgagcaaaaa caggaaggca aatgccgca aaaaaggaa taagggcgac 3240  
 acggaaatgt tgaatactca tactcttcct ttttcaatat tattgaagca tttatcaggg 3300  
 ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt 3360  
 tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac 3420  
 attaacctat aaaaataggc gtatcacgag gccctttcgt c 3461

<210> 65

<211> 7

5 <212> PRT

<213> Virus del grabado del tabaco

<400> 65

Glu Asn Leu Tyr Phe Gln Gly  
 1 5

10

<210> 66

<211> 21

<212> ADN

<213> Virus del grabado del tabaco

15

<400> 66

gaaaacctgt acttccaggg g 21

20

<210> 67

<211> 530

<212> ADN

<213> Secuencia artificial

25

<220>

<223> Secuencia de ADN de secuencia de proteína de tipo SNAP



ES 2 627 117 T3

<400> 67  
gacaaagact gcgaaatgaa aagaactaca ttggattcac cacttgggaa gttggaactg 60  
agtggatgcg agcaaggatt gcatgaaatt aagctactgg gaaaaggaac ttctgctgct 120  
gatgcagttg aagttccagc accagcagct gttcttgagg gtcctgagcc cctcatgcaa 180  
gccacagcct ggcttaacgc atatttccac cagcctgagg ccattgagga atttccagtc 240  
cccgcccttc accatcctgt gtttcagcag gagagcttca cccgccaggt cctgtggaaa 300  
ttgctgaagg tggtaagtt tgggtgaagt atttcatatc agcaacttgc tgcattggcc 360  
ggtaaccccg cagctacagc tgccgtgaaa actgctctca gcggaaatcc tgtgcccatc 420  
ctgatccctt gtcacagagt cgtttcatct tccggagctg taggtggcta tgaaggagga 480  
ctggcagtta aggagtggct gctggctcat gaaggtcata gacttggaaa 530

5 <210> 68  
<211> 711  
<212> ADN  
<213> Secuencia artificial

10 <220>  
<223> Secuencia de ADN de secuencia de MGMT modificada

<400> 68  
ctaggacaac ctgctccact agaacgattt gcttcacgac gtccacaggt ccttgcagta 60  
cgtactgttt gtgatttagt acttggaaaa atggacaaag actgcgaaat gaaaagaact 120  
acattggatt caccacttgg gaagttggaa ctgagtggat gcgagcaagg attgcatgaa 180  
attaagctac tgggaaaagg aacttctgct gctgatgcag ttgaagttcc agcaccagca 240  
gctgttcttg gaggtcctga gccctcatg caatgtacag catggcttaa cgcataattc 300  
caccagcctg aggccattga ggaatttcca gtccccgcc ttcaccatcc tgtgtttcag 360  
caggagagct tcaccgccca ggtcctgtgg aaattgctga aggtgggtcaa gtttggtgaa 420  
gtgatttcat atcagcaact tgctgcattg gccggtaacc ctaaagccgc gcgagcagtg 480  
ggaggagcaa tgagaggcaa tctgtgccc atcctgatcc cttgtcacag agtcgtttgt 540  
tcttccggag ctgtaggcaa ctattctgga ggactggcag ttaaggagtg gctgctggct 600  
catgaaggac atcgattag caaaccaggt ttaggaggtg gttcaggtct agcaggtgca 660  
tggttaagg gagcaggagc tacatctgga tcaccacctg ctggacgaaa t 711

15 <210> 69  
<211> 900  
<212> ADN  
<213> Secuencia artificial

20 <220>  
<223> Secuencia de ADN de DeMGMT Univ (tipo BIP/MGMT/proTEVx2/Etiqueta de His) (figura 9)

<400> 69

ES 2 627 117 T3

atgaaactat gtattctact tgcagttggt gcgttcgtag gattgtcctt acctacagct 60  
ctggcaagat ctctaggaca acctgctcca ctagaacgat ttgcttcacg acgtccacag 120  
gtccttgacg tacgtactgt ttgtgattta gtacttggaa aaatggacaa agactgcgaa 180  
atgaaaagaa ctacattgga ttcaccactt gggaagttgg aactgagtgg atgcgagcaa 240  
ggattgcatg aaattaagct actgggaaaa ggaacttctg ctgctgatgc agttgaagtt 300  
ccagcaccag cagctgttct tggaggtcct gagcccctca tgcaatgtac agcatggctt 360  
aacgcatatt tccaccagcc tgaggccatt gaggaatttc cagtccccgc ccttcaccat 420  
cctgtgtttc agcaggagag cttcaccocg caggtcctgt ggaaattgct gaaggtggtc 480  
aagtttgggtg aagtgatttc atatcagcaa cttgctgcat tggccggtaa ccctaaagcc 540  
gcgcgagcag tgggaggagc aatgagagc aatcctgtgc ccacctgat cccttgtcac 600  
agagtcgttt gttcttccgg agctgtaggc aactattctg gaggactggc agttaaggag 660  
tggctgctgg ctcatgaagg acatcgatta ggcaaaccag gtttaggagg tagttcaggt 720  
ctagcaggtg catggcttaa gggagcagga gctacatctg gatcaccacc tgctggacga 780  
aatggtggcg gatccgaaaa cctgtacttc cagagcgata tcggaggtgg aggcccggga 840  
gagaatctat attttcaagg gcccgcgga ggtagtcacc atcatcacca tcactaatga 900

<210> 70

<211> 297

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de DeMGMT Univ (tipo BIP/MGMT/proTEVx2/Etiqueta de His) (figura 9)

<400> 70

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

Leu Pro Thr Ala Leu Ala Arg Ser Leu Gly Gln Pro Ala Pro Leu Glu  
20 25 30

Arg Phe Ala Ser Arg Arg Pro Gln Val Leu Ala Val Arg Thr Val Cys

ES 2 627 117 T3

35 40 45

Asp Leu Val Leu Gly Lys Met Asp Lys Asp Cys Glu Met Lys Arg Thr  
50 55 60

Thr Leu Asp Ser Pro Leu Gly Lys Leu Glu Ser Gly Cys Glu Gln Gly  
65 70 75 80

Leu His Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala  
85 90 95

Val Glu Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu  
100 105 110

Met Gln Cys Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala  
115 120 125

Ile Glu Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln  
130 135 140

Glu Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys  
145 150 155 160

Phe Gly Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn  
165 170 175

Pro Lys Ala Ala Arg Ala Val Gly Gly Ala Met Arg Gly Asn Pro Val  
180 185 190

Pro Ile Leu Ile Pro Cys His Arg Val Val Cys Ser Ser Gly Ala Val  
195 200 205

Gly Asn Tyr Ser Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His  
210 215 220

Glu Gly His Arg Leu Gly Lys Pro Gly Leu Gly Gly Ser Ser Gly Leu  
225 230 235 240

Ala Gly Ala Trp Leu Lys Gly Ala Gly Ala Thr Ser Gly Ser Pro Pro  
245 250 255

Ala Gly Arg Asn Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln Ser Asp  
260 265 270

Ile Gly Gly Gly Gly Pro Gly Glu Asn Leu Tyr Phe Gln Gly Pro Gly  
275 280 285

Gly Gly Ser His His His His His His  
290 295

5 <210> 71  
 <211> 3581  
 <212> ADN  
 <213> Secuencia artificial

10 <220>

ES 2 627 117 T3

<223> Secuencia de ADN de vector lanzadera pUC57/DeMGMT (secuencia de DeMGMT insertada entre Nru I y Hind III de pUC57)

<400> 71

tcg	cg	cg	gtt	cg	gt	gat	gac	ggt	gaaa	acc	tct	gac	acat	gc	ag	ctccc	gag	ac	ggt	ca	60		
cag	ctt	gt	c	gta	ag	cg	gat	gcc	ggg	gag	ca	gaca	agccc	tc	agg	gc	gc	tc	agc	ggg	tg	120	
ttg	gc	ggg	tg	tc	ggg	ct	g	ctt	a	act	at	g	cg	gc	at	c	ag	att	gt	a	ct	180	
ac	cat	at	g	gt	g	aa	a	cc	g	ca	g	at	gc	gc	g	at	g	ag	g	gc	gc	240	
at	tc	gc	ca	tt	ag	ct	gc	aa	ct	gt	tt	gg	ag	gg	gc	at	c	g	g	gc	gc	300	
tac	gc	ca	g	ct	g	gc	g	aa	gg	g	g	g	at	g	gc	at	a	ag	tt	gg	g	360	
ttt	ccc	ag	tc	ac	g	ct	tt	gt	aaa	ac	g	ac	gc	gc	cc	ag	tc	gc	gc	gc	gc	420	
ct	ag	ca	cc	at	gt	gaa	ct	at	gt	gt	g	ct	act	tg	g	ct	ag	ta	gg	a	tt	480	
ct	ac	ag	ct	ct	g	ca	ag	at	ct	ct	ag	ca	ac	ct	g	ct	cc	act	aga	ac	g	att	540
gt	cc	ac	ag	gt	c	ct	gc	ag	ta	c	gt	act	gt	tt	gt	g	ta	act	tt	gg	aaaa	600	
act	gc	g	aa	at	g	aa	ct	g	g	at	g	at	c	ac	ct	tt	gg	ga	ag	tt	gg	660	
gc	g	ag	ca	ag	g	at	gc	at	g	aa	ta	ag	ct	ac	g	gg	g	aa	ag	g	aa	720	
tt	g	aa	gt	tc	ag	cc	ac	g	ca	g	ct	gt	tt	ct	tg	g	gc	cc	ct	ca	tg	780	
cat	g	gc	tt	aa	cg	at	at	tt	c	ac	ca	gc	ct	g	ag	gc	at	ta	g	g	aa	840	
tt	ca	cc	at	cc	tg	gt	tt	ct	ag	ca	g	g	ag	ag	ct	tc	ac	cc	gc	ca	g	900	
ag	gt	g	gt	caa	g	tt	g	gt	g	aa	gt	g	at	tt	ca	at	c	ag	ca	act	tg	960	
ct	aa	ag	cc	gc	gc	g	ac	g	ct	g	at	g	gc	gc	aa	tg	ag	ag	ca	tc	ct	1020	
ct	tt	gt	ca	ca	g	ag	ct	tt	gt	gt	gt	gc	ag	gc	gc	gc	gc	gc	gc	gc	gc	1080	
tt	a	ag	g	ag	g	at	g	at	g	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	1140	
gt	t	ca	g	gt	ct	ag	gt	ca	g	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	gc	1200	
ct	gg	ac	g	aaa	tg	gt	gg	cg	ga	acc	tg	t	act	tc	ca	g	gc	gc	at	at	c	1260	
g	ccc	gg	g	aga	ga	at	ct	at	at	ttt	ca	ag	gg	cc	gc	gc	gc	gc	gc	gc	gc	1320	
act	aat	g	acc	g	gt	gc	gg	cc	gc	ca	ag	ct	gg	gc	gt	aat	ca	tg	gc	gc	gc	1380	
aa	att	gt	at	cc	gc	tc	ac	aa	tt	cc	ac	ac	aa	ca	ta	ca	gc	gc	gc	gc	gc	1440	
5	ct	gg	g	gt	gc	ta	at	g	at	g	at	g	gc	gc	gc	gc	gc	gc	gc	gc	gc	1500	

ES 2 627 117 T3

ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg 1560  
cggtttgcgt attgggcgct cttccgcttc ctogctcact gactcgctgc gctcggtcgt 1620  
tcggctgcgg cgagcgggat cagctcactc aaaggcggta atacggttat ccacagaatc 1680  
aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 1740  
aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa 1800  
tcgacgctca agtcagaggt ggcgaaacct gacaggacta taaagatacc aggcgtttcc 1860  
ccctggaagc tcctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc 1920  
cgcctttctc ccttcgggaa gcgtggcgtc ttctcatagc tcacgctgta ggtatctcag 1980  
ttcgggtgtag gtggttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga 2040  
ccgctgcgcc ttatccggtc actatcgtct tgagtccaac ccggtaaagc acgacttacc 2100  
gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac 2160  
agagttcttg aagtgggtgc ctaactacgg ctacactaga agaacagtat ttggtatctg 2220  
cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaaa 2280  
aaccaccgct ggtagcggtg gttttttgt ttgcaagcag cagattacgc gcagaaaaaa 2340  
aggatctcaa gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa 2400  
ctcacgttaa gggattttg tcatgagatt atcaaaaagg atcttcacct agatcctttt 2460  
aaattaaaaa tgaagttta aatcaatcta aagtatatat gaggtaactt ggtctgacag 2520  
ttaccaatgc ttaatcagtg aggacacctc ctacagcagc tgtctatttc gttcatccat 2580  
agttgctga ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc 2640  
cagtgctgca atgataccgc gagaccacg ctaccggct ccagatttat cagcaataaa 2700  
ccagccagcc ggaagggcgc agcgcagaag tggctctgca actttatecc cctccatcca 2760  
gtctattaat tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgccaa 2820  
cgttgttgc attgctacag gcacgtggt gtcacgctcg tcgtttggtg tggcttcatt 2880  
cagctccggt tcccaacgat caaggcgagt tacatgatcc cccatggtgt gcaaaaaagc 2940  
ggttagctcc ttcggctctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact 3000  
catggttatg gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgcttttc 3060  
tgtgactggt gactactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg 3120  
ctcttgcggc gcgtcaatac gggataatac cgcgccacat agcagaactt taaaagtgt 3180  
catcattgga aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc 3240  
cagttcagtg taaccactc gtgcacccaa ctgatcttca gcactttta ctttcaccag 3300  
cgtttctggg tgagcaaaaa caggaaggca aatgcccga aaaaaggaa taagggcgac 3360  
acggaaatgt tgaatactca tactcttctc ttttcaatat tattgaagca tttatcaggg 3420  
ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt 3480  
tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac 3540  
attaacctat aaaaataggc gtatcacgag gccctttcgt c 3581

<210> 72  
<211> 1285  
<212> ADN

5

ES 2 627 117 T3

<213> Secuencia artificial

<220>

<223> Secuencia de ADN de DeSNAP Univ + IFN

5

<400> 72

tcgcgagcta	gcacatgaa	actatgtatt	ctacttgcag	ttgttgcgtt	cgtaggattg	60
tccttaccta	cagctctggc	aagatctgac	aaagactgcg	aaatgaaaag	aactacattg	120
gattcaccac	ttgggaagtt	ggaactgagt	ggatgcgagc	aaggattgca	tgaattaag	180
ctactgggaa	aaggaaacttc	tgctgctgat	gcagttgaag	ttccagcacc	agcagctggt	240
cttgagggtc	ctgagcccct	catgcaagcc	acagcctggc	ttaacgcata	tttccaccag	300
cctgaggcca	ttgaggaatt	tccagtcccc	gcccttcacc	atcctgtggt	tcagcaggag	360
agcttcaccc	gccaggctcct	gtggaaattg	ctgaaggtgg	tcaagtttgg	tgaagtgatt	420
tcatatcagc	aacttgctgc	attggccggt	aaccccgagc	ctacagctgc	cgtgaaaact	480
gctctcagcg	gaaatcctgt	gcccatcctg	atcccttgtc	acagagtcgt	ttcatcttcc	540
ggagctgtag	gtggctatga	aggaggactg	gcagttaagg	agtggctgct	ggctcatgaa	600
ggctcatagac	ttggaaagcc	tgggctgggt	cctgctggta	taggcgcgcc	agggctcccta	660
gggtggcgat	ccgaaaacct	gtacttccag	agcgatatct	gtgatctccc	tgagaccacc	720
agcctggata	acaggaggac	cttgatgctc	ctggcacaaa	tgagcagaat	ctctccttcc	780
tcctgtctga	tggacagaca	tgactttgga	tttccccagg	aggagtttga	tggcaaccag	840
ttccagaagg	ctccagccat	ctctgtcctc	catgagctga	tccagcagat	tttcaacctc	900
tttaccacaa	aagattcatc	tgctgcttgg	gatgaggacc	tcctagacaa	attctgcacc	960
gaactctacc	agcagctgaa	tgacttggaa	gcctgtgtga	tgcaggagga	gagggtggga	1020
gaaactcccc	tgatgaatgc	ggactccatc	ttggctgtga	agaaatactt	ccgaagaatc	1080
actctctatc	tgacagagaa	gaaatacagc	ccttgtgcct	gggaggttgt	cagagcagaa	1140
atcatgagat	ccctctcttt	atcaacaaac	ttgcaagaaa	gattaaggag	gaaggaaggc	1200
ccgggagaga	atctatattt	tcaagggccc	ggcggaggta	gtcaccatca	tcaccatcac	1260
taatgaccgg	tgcggccgca	agctt				1285

<210> 73

<211> 414

10

<212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia de aminoácidos de De SNAP Univ + IFN

15

<400> 73

ES 2 627 117 T3

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Pro Thr Ala Leu Ala Arg Ser Asp Lys Asp Cys Glu Met Lys Arg  
 20 25 30

Thr Thr Leu Asp Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu  
 35 40 45

Gln Gly Leu His Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Asp  
 50 55 60

Ala Val Glu Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro  
 65 70 75 80

Leu Met Gln Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu  
 85 90 95

Ala Ile Glu Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln  
 100 105 110

Gln Glu Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val  
 115 120 125

Lys Phe Gly Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly  
 130 135 140

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

Val Pro Ile Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala  
 165 170 175

Val Gly Gly Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala  
 180 185 190

His Glu Gly His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile  
 195 200 205

Gly Ala Pro Gly Ser Leu Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln

ES 2 627 117 T3

210		215		220
Ser Asp Ile Cys Asp Leu Pro Glu Thr His Ser Leu Asp Asn Arg Arg				
225		230		235 240
Thr Leu Met Leu Leu Ala Gln Met Ser Arg Ile Ser Pro Ser Ser Cys				
		245		250 255
Leu Met Asp Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp Gly				
		260		265 270
Asn Gln Phe Gln Lys Ala Pro Ala Ile Ser Val Leu His Glu Leu Ile				
		275		280 285
Gln Gln Ile Phe Asn Leu Phe Thr Thr Lys Asp Ser Ser Ala Ala Trp				
		290		295 300
Asp Glu Asp Leu Leu Asp Lys Phe Cys Thr Glu Leu Tyr Gln Gln Leu				
305		310		315 320
Asn Asp Leu Glu Ala Cys Val Met Gln Glu Glu Arg Val Gly Glu Thr				
		325		330 335
Pro Leu Met Asn Ala Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe Arg				
		340		345 350
Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala Trp				
		355		360 365
Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser Leu Ser Thr Asn				
		370		375 380
Leu Gln Glu Arg Leu Arg Arg Lys Glu Gly Pro Gly Glu Asn Leu Tyr				
385		390		395 400
Phe Gln Gly Pro Gly Gly Gly Ser His His His His His His				
		405		410

<210> 74

<211> 1423

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Secuencia de ADN de DeMGMT Univ + IFN

10

<400> 74

tcgcgagcta gcaccatgaa actatgtatt ctacttgcag ttgttgcgtt cgtaggattg 60

tccttaccta cagctctggc aagatctcta ggacaacctg ctccactaga acgatttgct 120



ES 2 627 117 T3

tcacgacgtc cacaggtcct tgcagtacgt actgtttgtg atttagtact tggaaaaatg 180  
gacaaaagact gcgaaatgaa aagaactaca ttggattcac cacttgggaa gttggaactg 240  
agtggatgcg agcaaggatt gcatgaaatt aagctactgg gaaaaggaac ttctgctgct 300  
gatgcagttg aagttccagc accagcagct gttcttggag gtcttgagcc cctcatgcaa 360  
tgtacagcat ggcttaacgc atatttccac cagcctgagg ccattgagga atttccagtc 420  
cccgcccttc accatcctgt gtttcagcag gagagcttca cccgccaggt cctgtggaaa 480  
ttgctgaagg tggtaagtt tgggtaagt atttcatatc agcaacttgc tgcattggcc 540  
ggtaacccta aagccgcgcg agcagtggga ggagcaatga gaggcaatcc tgtgccatc 600  
ctgatccctt gtcacagagt cgtttgttct tccggagctg taggcaacta ttctggagga 660  
ctggcagtta aggagtggct gctggctcat gaaggacatc gattaggcaa accaggttta 720  
ggaggtagtt caggtctagc aggtgcatgg ctttaaggag caggagctac atctggatca 780  
ccacctgctg gacgaaatgg tggcggatcc gaaaacctgt acttccagag cgatatctgt 840  
gatctccctg agaccacag cctggataac aggaggacct tgatgctcct ggcacaaatg 900  
agcagaatct ctcttctctc ctgtctgatg gacagacatg actttggatt tccccaggag 960  
gagtttgatg gcaaccagtt ccagaaggct ccagccatct ctgtcctcca tgagctgatc 1020  
cagcagatth tcaacctctt taccacaaaa gattcatctg ctgcttggga tgaggacctc 1080  
ctagacaaat tctgcaccga actctaccag cagctgaatg acttgggaagc ctgtgtgatg 1140  
caggaggaga ggtgaggaga aactcccctg atgaatgagg actccatctt ggctgtgaag 1200  
aaatacttcc gaagaatcac tctctatctg acagagaaga aatacagccc ttgtgcctgg 1260  
gaggttgta gagcagaaat catgagatcc ctctctttat caacaaactt gcaagaaaga 1320  
ttaaggagga aggaaggccc gggagagaat ctatattttc aagggcccgg cggaggtagt 1380  
caccatcatc accatcacta atgaccggtg cggccgcaag ctt 1423

<210> 75  
<211> 415  
5 <212> PRT  
<213> Secuencia artificial

<220>  
10 <223> Secuencia de aminoácidos de DeMGMT Univ + IFN

<400> 75  
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15  
Leu Pro Thr Ala Leu Ala Arg Ser Asp Lys Asp Cys Glu Met Lys Arg  
20 25 30

ES 2 627 117 T3

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

ES 2 627 117 T3

Ile Gln Gln Ile Phe Asn Leu Phe Thr Thr Lys Asp Ser Ser Ala Ala  
 290 295 300

Trp Asp Glu Asp Leu Leu Asp Lys Phe Cys Thr Glu Leu Tyr Gln Gln  
 305 310 315 320

Leu Asn Asp Leu Glu Ala Cys Val Met Gln Glu Glu Arg Val Gly Glu  
 325 330 335

Thr Pro Leu Met Asn Ala Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe  
 340 345 350

Arg Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala  
 355 360 365

Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser Leu Ser Thr  
 370 375 380

Asn Leu Gln Glu Arg Leu Arg Arg Lys Glu Gly Pro Gly Glu Asn Leu  
 385 390 395 400

Tyr Phe Gln Gly Pro Gly Gly Gly Ser His His His His His His  
 405 410 415

<210> 76

<211> 570

5 <212> PRT

<213> Homo sapiens

<400> 76

Gly Ala Thr Ala Thr Cys Ala Ala Cys Gly Gly Ala Gly Ala Cys Gly  
 1 5 10 15

Ala Cys Gly Cys Cys Thr Thr Thr Gly Cys Ala Ala Gly Gly Ala Gly  
 20 25 30

Ala Cys Cys Cys Ala Cys Gly Gly Thr Thr Gly Gly Thr Gly Cys Thr  
 35 40 45

Cys Ala Ala Ala Thr Ala Cys Cys Ala Gly Ala Gly Ala Ala Gly Ala  
 50 55 60

Thr Cys Cys Ala Ala Ala Ala Gly Gly Cys Cys Thr Thr Cys Gly Ala  
 65 70 75 80

Thr Gly Ala Thr Ala Thr Thr Gly Cys Cys Ala Ala Ala Thr Ala Cys  
 85 90 95

ES 2 627 117 T3

Thr Thr Cys Thr Cys Thr Ala Ala Gly Gly Ala Ala Gly Ala Gly Thr  
 100 105 110  
 Gly Gly Gly Ala Ala Ala Ala Gly Ala Thr Gly Ala Ala Ala Gly Cys  
 115 120 125  
 Cys Thr Cys Gly Gly Ala Gly Ala Ala Ala Ala Thr Cys Thr Thr Cys  
 130 135 140  
 Thr Ala Thr Gly Thr Gly Thr Ala Thr Ala Thr Gly Ala Ala Gly Ala  
 145 150 155 160  
 Gly Ala Ala Ala Gly Thr Ala Thr Gly Ala Gly Gly Cys Thr Ala Thr  
 165 170 175  
 Gly Ala Cys Thr Ala Ala Ala Cys Thr Ala Gly Gly Thr Thr Thr Cys  
 180 185 190  
 Ala Ala Gly Gly Cys Cys Ala Cys Cys Cys Thr Cys Cys Cys Ala Cys  
 195 200 205  
 Cys Thr Thr Thr Cys Ala Thr Gly Thr Gly Thr Ala Ala Thr Ala Ala  
 210 215 220  
 Ala Cys Gly Gly Gly Cys Cys Gly Ala Ala Gly Ala Cys Thr Thr Cys  
 225 230 235 240  
 Cys Ala Gly Gly Gly Gly Ala Ala Thr Gly Ala Thr Thr Thr Gly Gly  
 245 250 255  
 Ala Thr Ala Ala Thr Gly Ala Cys Cys Cys Thr Ala Ala Cys Cys Gly  
 260 265 270  
 Thr Gly Gly Gly Ala Ala Thr Cys Ala Gly Gly Thr Thr Gly Ala Ala  
 275 280 285  
 Cys Gly Thr Cys Cys Thr Cys Ala Gly Ala Thr Gly Ala Cys Thr Thr  
 290 295 300  
 Thr Cys Gly Gly Cys Ala Gly Gly Cys Thr Cys Cys Ala Gly Gly Gly  
 305 310 315 320  
 Ala Ala Thr Cys Thr Cys Cys Cys Cys Gly Ala Ala Gly Ala Thr Cys  
 325 330 335  
 Ala Thr Gly Cys Cys Cys Ala Ala Gly Ala Ala Gly Cys Cys Ala Gly

ES 2 627 117 T3

	340		345		350														
Cys	Ala	Gly	Ala	Gly	Gly	Ala	Ala	Gly	Gly	Ala	Ala	Ala	Thr	Gly	Ala				
		355						360						365					
Thr	Thr	Cys	Gly	Gly	Ala	Gly	Gly	Ala	Ala	Gly	Thr	Gly	Cys	Cys	Ala				
	370					375					380								
Gly	Ala	Ala	Gly	Cys	Ala	Thr	Cys	Thr	Gly	Gly	Cys	Cys	Cys	Ala	Cys				
385					390					395					400				
Ala	Ala	Ala	Ala	Thr	Gly	Ala	Thr	Gly	Gly	Gly	Ala	Ala	Ala	Gly	Ala				
				405					410					415					
Gly	Cys	Thr	Gly	Thr	Gly	Cys	Cys	Cys	Thr	Cys	Cys	Thr	Gly	Gly	Ala				
			420					425					430						
Ala	Ala	Ala	Cys	Cys	Ala	Ala	Cys	Thr	Ala	Cys	Cys	Thr	Cys	Thr	Gly				
		435					440					445							
Ala	Gly	Ala	Ala	Gly	Ala	Thr	Thr	Cys	Ala	Cys	Gly	Ala	Gly	Ala	Gly				
	450					455					460								
Ala	Thr	Cys	Ala	Gly	Gly	Ala	Cys	Cys	Cys	Ala	Ala	Ala	Ala	Gly	Gly				
465					470					475					480				
Gly	Gly	Gly	Gly	Ala	Ala	Cys	Ala	Thr	Gly	Cys	Cys	Thr	Gly	Gly	Ala				
				485					490					495					
Cys	Cys	Cys	Ala	Cys	Ala	Gly	Ala	Cys	Thr	Gly	Cys	Gly	Thr	Gly	Ala				
			500					505					510						
Gly	Ala	Gly	Ala	Ala	Ala	Ala	Cys	Ala	Gly	Cys	Thr	Gly	Gly	Thr	Gly				
		515					520					525							
Ala	Thr	Thr	Thr	Ala	Thr	Gly	Ala	Ala	Gly	Ala	Gly	Ala	Thr	Cys	Ala				
	530					535					540								
Gly	Cys	Gly	Ala	Cys	Cys	Cys	Thr	Gly	Ala	Gly	Gly	Ala	Ala	Gly	Ala				
545					550					555				560					
Thr	Gly	Ala	Cys	Gly	Ala	Gly	Thr	Ala	Cys										
				565					570										

<210> 77  
 <211> 1369  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Secuencia de ADN de DeSNAP Univ + SSX2

10 <400> 77

ES 2 627 117 T3

tcgcgagcta gcaccatgaa actatgtatt ctacttgcag ttgttgcggt cgtaggattg 60  
 tccttaccta cagctctggc aagatctgac aaagactgcg aatgaaaag aactacattg 120  
 gattcaccac ttgggaagtt ggaactgagt ggatgcgagc aaggattgca tgaaattaag 180  
 ctactgggaa aaggaacttc tgctgctgat gcagttgaag ttccagcacc agcagctggt 240  
 cttggaggtc ctgagcccct catgcaagcc acagcctggc ttaacgcata tttccaccag 300  
 cctgaggcca ttgaggaatt tccagtcccc gcccttcacc atcctgtggt tcagcaggag 360  
 agcttcaccc gccaggctct gtggaaattg ctgaaggtgg tcaagtttgg tgaagtgatt 420  
 tcatatcagc aacttgctgc attggccggt aaccccgag ctacagctgc cgtgaaaact 480  
 gctctcagcg gaaatcctgt gcccatcctg atcccttgtc acagagtcgt ttcattctcc 540  
 ggagctgtag gtggctatga aggaggactg gcagttaagg agtggctgct ggctcatgaa 600  
 ggtcatagac ttgaaagcc tgggctgggt cctgctggta taggcgcgcc agggctcccta 660  
 ggtgggggat cggaaaacct gtacttccag agcgatatca acggagacga cgcctttgca 720  
 aggagaccca cggttggtgc tcaaatacca gagaagatcc aaaaggcctt cgatgatatt 780  
 gccaaatact tctctaagga agagtgggaa aagatgaaag cctcggagaa aatcttctat 840  
 gtgtatatga agagaaagta tgaggctatg actaaactag gtttcaaggc caccctccca 900  
 cctttcatgt gtaataaacg ggccgaagac ttccagggga atgatttggga taatgacct 960  
 aaccgtggga atcaggttga acgtcctcag atgactttcg gcaggctcca gggaatctcc 1020  
 ccgaagatca tgcccaagaa gccagcagag gaaggaaatg attcggagga agtgccagaa 1080  
 gcatctggcc cacaaaatga tgggaaagag ctgtgccctc ctggaaaacc aactacctct 1140  
 gagaagattc acgagagatc aggacccaaa aggggggaac atgcctggac ccacagactg 1200  
 cgtgagagaa aacagctggt gatattatgaa gagatcagcg accctgagga agatgacgag 1260  
 tacgagaatc tatattttca aggcccgga gagaatctat attttcaagg gcccgcgga 1320  
 ggtagtcacc atcatcacca tcaactaatga ccggtgcggc cgcaagctt 1369

<210> 78

<211> 415

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de DeSNAP Univ + SSX2

<400> 78

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser



ES 2 627 117 T3

Cys Leu Met Asp Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Asp  
 260 265 270

Gly Asn Gln Phe Gln Lys Ala Pro Ala Ile Ser Val Leu His Glu Leu  
 275 280 285

Ile Gln Gln Ile Phe Asn Leu Phe Thr Thr Lys Asp Ser Ser Ala Ala  
 290 295 300

Trp Asp Glu Asp Leu Leu Asp Lys Phe Cys Thr Glu Leu Tyr Gln Gln  
 305 310 315 320

Leu Asn Asp Leu Glu Ala Cys Val Met Gln Glu Glu Arg Val Gly Glu  
 325 330 335

Thr Pro Leu Met Asn Ala Asp Ser Ile Leu Ala Val Lys Lys Tyr Phe  
 340 345 350

Arg Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser Pro Cys Ala  
 355 360 365

Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser Leu Ser Thr  
 370 375 380

Asn Leu Gln Glu Arg Leu Arg Arg Lys Glu Gly Pro Gly Glu Asn Leu  
 385 390 395 400

Tyr Phe Gln Gly Pro Gly Gly Gly Ser His His His His His His  
 405 410 415

<210> 79  
 <211> 1369  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Secuencia de ADN de DeMGMT Univ + SSX2

10

<400> 79  
 tcgcgagcta gcacccatgaa actatgtatt ctacttgacg ttggtgagtt cgtaggattg 60  
 tccttaccta cagctctggc aagatctgac aaagactgag aatgaaaag aactacattg 120  
 gattcaccac ttgggaagtt ggaactgagt ggatgagagc aaggattgca tgaaattaag 180  
 ctactgggaa aaggaacttc tgctgctgat gcagttgaag ttccagcacc agcagctggt 240  
 cttggaggtc ctgagcccct catgcaagcc acagcctggc ttaacgcata tttccaccag 300  
 cctgaggcca ttgaggaatt tccagtcacc gcccttcacc atcctgtggt tcagcaggag 360  
 agcttcaccc gccaggtcct gtggaaattg ctgaaggtgg tcaagtttgg tgaagtgatt 420



ES 2 627 117 T3

tcatatcagc aacttgctgc attggcgggt aaccccgag ctacagctgc cgtgaaaact 480  
gctctcagcg gaaatcctgt gcccatcctg atcccttgtc acagagtcgt ttcattctcc 540  
ggagctgtag gtggctatga aggaggactg gcagttaagg agtggctgct ggctcatgaa 600  
ggtcatagac ttgaaagcc tgggctgggt cctgctggta taggcgcgcc aggtcccta 660  
ggtggcggat ccgaaaacct gtacttccag agcgatatca acggagacga cgcctttgca 720  
aggagaccca cggttgggtgc tcaaatacca gagaagatcc aaaaggcctt cgatgatatt 780  
gccaaatact tctctaagga agagtgggaa aagatgaaag cctcggagaa aatcttctat 840  
gtgtatatga agaaaaagta tgaggctatg actaaactag gtttcaaggc caccctccca 900  
cctttcatgt gtaataaacg ggccgaagac ttccagggga atgatttggga taatgacct 960  
aacctgggga atcaggttga acgtcctcag atgactttcg gcaggctcca gggatctcc 1020  
ccgaagatca tgcccaagaa gccagcagag gaagaaatg attcggagga agtgccagaa 1080  
gcatctggcc cacaaaatga tgggaaagag ctgtgccctc ctggaaaacc aactacctct 1140  
gagaagattc acgagagatc aggacccaaa aggggggaac atgcctggac ccacagactg 1200  
cgtgagagaa aacagctggt gatttatgaa gagatcagcg accctgagga agatgacgag 1260  
tacgagaatc tatattttca aggcccgga gagaatctat attttcaagg gcccgcgga 1320  
ggtagtcacc atcatcacca tcaactaatga ccggtgcggc cgcaagctt 1369

<210> 80

<211> 489

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de DeMGMT Univ + SSX2

<400> 80

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

Leu Pro Thr Ala Leu Ala Arg Ser Leu Gly Gln Pro Ala Pro Leu Glu  
20 25 30

Arg Phe Ala Ser Arg Arg Pro Gln Val Leu Ala Val Arg Thr Val Cys  
35 40 45

Asp Leu Val Leu Gly Lys Met Asp Lys Asp Cys Glu Met Lys Arg Thr  
50 55 60

Thr Leu Asp Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln  
65 70 75 80

ES 2 627 117 T3

Gly Leu His Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp  
85 90 95

Ala Val Glu Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro  
100 105 110

Leu Met Gln Cys Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu  
115 120 125

Ala Ile Glu Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln  
130 135 140

Gln Glu Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val  
145 150 155 160

Lys Phe Gly Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly  
165 170 175

Asn Pro Lys Ala Ala Arg Ala Val Gly Gly Ala Met Arg Gly Asn Pro  
180 185 190

Val Pro Ile Leu Ile Pro Cys His Arg Val Val Cys Ser Ser Gly Ala  
195 200 205

Val Gly Asn Tyr Ser Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala  
210 215 220

His Glu Gly His Arg Leu Gly Lys Pro Gly Leu Gly Gly Ser Ser Gly  
225 230 235 240

Leu Ala Gly Ala Trp Leu Lys Gly Ala Gly Ala Thr Ser Gly Ser Pro  
245 250 255

Pro Ala Gly Arg Asn Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln Ser  
260 265 270

Asp Ile Asn Gly Asp Asp Ala Phe Ala Arg Arg Pro Thr Val Gly Ala  
275 280 285

Gln Ile Pro Glu Lys Ile Gln Lys Ala Phe Asp Asp Ile Ala Lys Tyr  
290 295 300

Phe Ser Lys Glu Glu Trp Glu Lys Met Lys Ala Ser Glu Lys Ile Phe  
305 310 315 320

Tyr Val Tyr Met Lys Arg Lys Tyr Glu Ala Met Thr Lys Leu Gly Phe  
325 330 335

ES 2 627 117 T3

Lys Ala Thr Leu Pro Pro Phe Met Cys Asn Lys Arg Ala Glu Asp Phe  
 340 345 350

Gln Gly Asn Asp Leu Asp Asn Asp Pro Asn Arg Gly Asn Gln Val Glu  
 355 360

Arg Pro Gln Met Thr Phe Gly Arg Leu Gln Gly Ile Ser Pro Lys Ile  
 370 375 380

Met Pro Lys Lys Pro Ala Glu Glu Gly Asn Asp Ser Glu Glu Val Pro  
 385 390 395 400

Glu Ala Ser Gly Pro Gln Asn Asp Gly Lys Glu Leu Cys Pro Pro Gly  
 405 410 415

Lys Pro Thr Thr Ser Glu Lys Ile His Glu Arg Ser Gly Pro Lys Arg  
 420 425 430

Gly Glu His Ala Trp Thr His Arg Leu Arg Glu Arg Lys Gln Leu Val  
 435 440 445

Ile Tyr Glu Glu Ile Ser Asp Pro Glu Glu Asp Asp Glu Tyr Glu Asn  
 450 455 460

Leu Tyr Phe Gln Gly Pro Gly Glu Asn Leu Tyr Phe Gln Gly Pro Gly  
 465 470 475 480

Gly Gly Ser His His His His His His  
 485

<210> 81  
 <211> 1299  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 10 <223> Secuencia de ADN de ssBiP+SNAP+proTEV site + SSX2 + proTEV + Etiqueta de His

<400> 81  
 atgaagttat gcatattact ggccgtcgtg gcctttggtg gcctctcgct cgggagatct 60  
 gacaaaagact gcgaaatgaa gcgcaccacc ctggatagcc ctctgggcaa gctggaactg 120  
 tctgggtgcg aacagggcct gcacgagatc aagctgctgg gcaaaggaac atctgccgcc 180  
 gacgccgtgg aagtgcctgc cccagccgcc gtgctgggcg gaccagagcc actgatgcag 240  
 gccaccgcct ggtcaacgc ctactttcac cagcctgagg ccatcgagga gttccctgtg 300  
 ccagccctgc accaccagt gttccagcag gagagcttta cccgccaggt gctgtggaag 360  
 ctgctgaaag tggatgaagt cggagaggtc atcagctacc agcagctggc cgccctggcc 420

ES 2 627 117 T3

ggcaatcccg cggccaccgc cgccgtgaaa accgacctga gcggaaatcc cgtgccatt 480  
ctgatcccct gccaccgggt ggtgtctagc tctggcgccg tggggggcta cgagggcggg 540  
ctcgccgtga aagagtggct gctggccac gagggccaca gactgggcaa gcctgggctg 600  
ggtcctgcag gtataggcgc gccagggtcc ctaggtggcg gatctgaaaa cctctacttc 660  
cagagtgata tcaacggaga cgacgccttt gcaaggagac ccacggttgg tgctcaaata 720  
ccagagaaga tccaaaaggc cttcgatgat attgccaaat acttctctaa ggaagagtgg 780  
gaaaagatga aagcctcggg gaaaatcttc tatgtgtata tgaagagaaa gtatgaggct 840  
atgactaaac taggtttcaa ggccaccctc ccacctttca tgtgtaataa acgggccgaa 900  
gacttccagg ggaatgattt ggataatgac cctaaccgtg ggaatcaggt tgaacgtcct 960  
cagatgactt tcggcaggct ccagggaatc tccccgaaga tcatgcccaa gaagccagca 1020  
gaggaaggaa atgattcggg ggaagtgccg gaagcatctg gccacaaaa tgatgggaaa 1080  
gagctgtgcc ctctctggaaa accaactacc tctgagaaga ttcacgagag atcaggacct 1140  
aaaagggggg aacatgcctg gaccacaga ctgcgtgaga gaaaacagct ggtgatttat 1200  
gaagagatca gcgaccctga ggaagatgac gagtacgaga atctatattt tcaaggcccg 1260  
ggcggtgga gtcacatca tcacatcac tgaccggta 1299

<210> 82

<211> 430

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de BiPmc+SNAP+sitio proTEV + SSX2+ proTEV + Etiqueta de His

<400> 82

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
20 25 30

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
65 70 75 80

ES 2 627 117 T3

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
115 120 125

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
195 200 205

Gly Ser Leu Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln Ser Asp Ile  
210 215 220

Asn Gly Asp Asp Ala Phe Ala Arg Arg Pro Thr Val Gly Ala Gln Ile  
225 230 235 240

Pro Glu Lys Ile Gln Lys Ala Phe Asp Asp Ile Ala Lys Tyr Phe Ser  
245 250 255

Lys Glu Glu Trp Glu Lys Met Lys Ala Ser Glu Lys Ile Phe Tyr Val  
260 265 270

Tyr Met Lys Arg Lys Tyr Glu Ala Met Thr Lys Leu Gly Phe Lys Ala  
275 280 285

Thr Leu Pro Pro Phe Met Cys Asn Lys Arg Ala Glu Asp Phe Gln Gly  
290 295 300

Asn Asp Leu Asp Asn Asp Pro Asn Arg Gly Asn Gln Val Glu Arg Pro  
305 310 315 320

Gln Met Thr Phe Gly Arg Leu Gln Gly Ile Ser Pro Lys Ile Met Pro  
325 330 335

ES 2 627 117 T3

Lys Lys Pro Ala Glu Glu Gly Asn Asp Ser Glu Glu Val Pro Glu Ala  
 340 345 350

Ser Gly Pro Gln Asn Asp Gly Lys Glu Leu Cys Pro Pro Gly Lys Pro  
 355 360 365

Thr Thr Ser Glu Lys Ile His Glu Arg Ser Gly Pro Lys Arg Gly Glu  
 370 375 380

His Ala Trp Thr His Arg Leu Arg Glu Arg Lys Gln Leu Val Ile Tyr  
 385 390 395 400

Glu Glu Ile Ser Asp Pro Glu Glu Asp Asp Glu Tyr Glu Asn Leu Tyr  
 405 410 415

Phe Gln Gly Pro Gly Gly Gly Ser His His His His His His  
 420 425 430

<210> 83

<211> 798

5 <212> ADN

<213> Homo sapiens

<400> 83

tctctagctg	gagaaacagg	tcaagaagct	gcacctcttg	atggagtact	agcaaatcca	60
cctaataattt	caagtctatc	acctcgacaa	cttcttggat	ttccatgtgc	agaagtatct	120
ggactaagta	cagaacgtgt	tcgagaacta	gctgtagcat	tagcacagaa	aatgtaaaa	180
ctatcaacag	aacaacttcg	atgtctagct	catcgacttt	ctgaaccacc	tgaggatcta	240
gatgcacttc	cattogatct	acttctatct	ctaaatccag	atgcattttc	aggacctcaa	300
gcatgtactc	gatttttttc	tcgaattaca	aaagcaaag	tcgatctact	tccaagagga	360
gcaccagaac	gacaacgact	actacctgca	gctctagcat	gttggggagt	acgaggatct	420
ctacttagtg	aagcagatgt	acgagctcta	ggaggtctag	cttgtgatct	acctggacga	480
tttgtagcag	aatctgcaga	agtactacta	ccacgacttg	ttagttgtcc	tgacacctta	540
gatcaagatc	aacaagaagc	tgctagagca	gctcttcaag	gtggtggacc	tccttatgga	600
cctccatcaa	catggtctgt	atcaacaatg	gatgcactac	gaggacttct	tcctgtacta	660
ggtcaacctt	ttattcgaag	tattccacaa	ggtattgtag	cagcatggcg	acaacgatct	720
tctcgagatc	catcttggcg	acaacctgaa	cgaactattc	ttcgaccacg	cccgggagag	780
aatctatatt	ttcaaggg					798

10 <210> 84

<211> 1558

<212> ADN

<213> Secuencia artificial

15 <220>

<223> Secuencia de ADN de DeSNAP Univ -NERCMSL

<400> 84

ES 2 627 117 T3

tcgcgagcta gcacccatgaa actatgtatt ctacttgacag ttgttgcggt cgtaggattg 60  
 tccttaccta cagctctggc aagatctgac aaagactgag aatgaaaag aactacattg 120  
 gattcaccac ttgggaagtt ggaactgagt ggatgagagc aaggattgca tgaaattaag 180  
 ctactgggaa aaggaacttc tgctgctgat gcagttgaag ttccagcacc agcagctggt 240  
 cttggaggtc ctgagcccct catgcaagcc acagcctggc ttaacgcata tttccaccag 300  
 cctgaggcca ttgaggaatt tccagtcccc gcccttcacc atcctgtggt tcagcaggag 360  
 agcttcaccc gccaggtcct gtggaaattg ctgaaggtgg tcaagtttgg tgaagtgatt 420  
 tcatatcagc aacttgctgc attggccggt aaccccgag ctacagctgc cgtgaaaact 480  
 gctctcagcg gaaatcctgt gcccatcctg atcccttgtc acagagtcgt ttcattctcc 540  
 ggagctgtag gtggctatga aggaggactg gcagttaagg agtggctgct ggctcatgaa 600  
 ggtcatagac ttgaaagcc tgggctgggt cctgctggta taggcgagcc agggctcccta 660  
 ggtggcggat ccgaaaacct gtacttccag agcgatatca ggtctctagc tggagaaaca 720  
 ggtcaagaag ctgcacctct tgatggagta ctagcaaatc cacctaatat ttcaagtcta 780  
 tcacctcgac aacttcttgg atttccatgt gcagaagtat ctggactaag tacagaacgt 840  
 gttcgagaac tagctgtagc attagcacag aaaaatgtaa aactatcaac agaacaactt 900  
 cgatgtctag ctcatcgact ttctgaacca cctgaggatc tagatgcact tccattcgat 960  
 ctacttctat ttctaaatcc agatgcattt tcaggacctc aagcatgtac tcgatttttt 1020  
 tctcgaatta caaaagcaaa tgtcgatcta cttccaagag gagcaccaga acgacaacga 1080  
 ctactacctg cagctctagc atgttgggga gtacgaggat ctctacttag tgaagcagat 1140  
 gtacgagctc taggaggtct agcttgtgat ctacctggac gatttgtagc agaatctgca 1200  
 gaagtactac taccacgact tgttagttgt cctggacctc tagatcaaga tcaacaagaa 1260  
 gctgctagag cagctcttca aggtggtgga cctccttatg gacctccatc aacatggtct 1320  
 gtatcaacaa tggatgcact acgaggactt cttcctgtac taggtcaacc tattattcga 1380  
 agtattccac aaggtattgt agcagcatgg cgacaacgat cttctcgaga tccatcttgg 1440  
 cgacaacctg aacgaactat tcttcgacca cgcccgggag agaatctata ttttcaaggg 1500  
 cccggcggag gtagtcacca tcatcaccat caactaatgac cggtgaggcc gcaagctt 1558

<210> 85

<211> 504

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de DeSNAP Univ -NERCMSL

<400> 85

ES 2 627 117 T3

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Pro Thr Ala Leu Ala Arg Ser Asp Lys Asp Cys Glu Met Lys Arg  
 20 25 30

Thr Thr Leu Asp Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu  
 35 40 45

Gln Gly Leu His Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala  
 50 55 60

Asp Ala Val Glu Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu  
 65 70 75 80

Pro Leu Met Gln Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro  
 85 90 95

Glu Ala Ile Glu Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe  
 100 105 110

Gln Gln Glu Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val  
 115 120 125

Val Lys Phe Gly Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala  
 130 135 140

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

Pro Val Ile Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala  
 165 170 175

Val Gly Gly Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala  
 180 185 190

His Glu Gly His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile  
 195 200 205

Gly Ala Pro Gly Ser Leu Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln  
 210 215 220

Ser Asp Ile Arg Ser Leu Ala Gly Glu Thr Gly Gln Glu Ala Ala Pro





ES 2 627 117 T3

<213> Secuencia artificial

<220>

<223> Secuencia de ADN de DeMGMT Univ -NERCMSL

5

<400> 86

tcgcgagcta	gcacccatgaa	actatgtatt	ctacttgcag	ttgttgcgtt	cgtaggattg	60
tccttaccta	cagctctggc	aagatctgac	aaagactgcg	aatgaaaag	aactacattg	120
gattcaccac	ttgggaagtt	ggaactgagt	ggatgcgagc	aaggattgca	tgaaattaag	180
ctactgggaa	aaggaacttc	tgctgctgat	gcagttgaag	ttccagcacc	agcagctggt	240
cttgagggtc	ctgagccctc	catgcaagcc	acagcctggc	ttaacgcata	tttccaccag	300
cctgaggcca	ttgaggaatt	tccagtcccc	gcccttcacc	atcctgtggt	tcagcaggag	360
agcttcaccc	gccaggctct	gtggaaattg	ctgaaggtgg	tcaagtttgg	tgaagtgatt	420
tcatatcagc	aacttgctgc	attggccggt	aaccccgagc	ctacagctgc	cgtgaaaact	480
gctctcagcg	gaaatcctgt	gcccatcctg	atcccttgtc	acagagtcgt	ttcatcttcc	540
ggagctgtag	gtggctatga	aggaggactg	gcagttaagg	agtggctgct	ggctcatgaa	600
ggctcatagac	ttgaaagcc	tgggctgggt	cctgctggta	taggcgcgcc	agggtcctta	660
ggtgccggat	ccgaaaacct	gtacttccag	agcgatatca	ggtctctagc	tggagaaaca	720
ggtaagaag	ctgcacctct	tgatggagta	ctagcaaatc	cacctaatat	ttcaagtcta	780
tcacctcgac	aacttcttgg	atctccatgt	gcagaagtat	ctggactaag	tacagaacgt	840
gttcgagaac	tagctgtagc	attagcacag	aaaaatgtaa	aactatcaac	agaacaactt	900
cgatgtctag	ctcatcgact	ttctgaacca	cctgaggatc	tagatgcact	tccattogat	960
ctacttctat	ttctaaatcc	agatgcattt	tcaggacctc	aagcatgtac	togatttttt	1020
tctcgaatta	caaaagcaaa	tgtcgatcta	cttccaagag	gagcaccaga	acgacaacga	1080
ctactacctg	cagctctagc	atgttgggga	gtacgaggat	ctctacttag	tgaagcagat	1140
gtacgagctc	taggaggtct	agcttgtgat	ctacctggac	gatttgtagc	agaatctgca	1200
gaagtactac	taccacgact	tgttagttgt	cctggacctc	tagatcaaga	tcaacaagaa	1260
gctgctagag	cagctcttca	aggtggtgga	cctccttatg	gacctccatc	aacatggtct	1320
gtatcaacaa	tggatgcact	acgaggactt	cttctgttac	taggtcaacc	tattatttga	1380
agtattccac	aaggtattgt	agcagcatgg	cgacaacgat	cttctcgaga	tccatcttgg	1440
cgacaacctg	aacgaactat	tcttcgacca	cgcccgggag	agaatctata	ttttcaaggg	1500
cccggcggag	gtagtcacca	tcatacccat	cactaatgac	cggtgcggcc	gcaagctt	1558

10

<210> 87

<211> 551

<212> PRT

<213> Secuencia artificial

15

<220>

<223> Secuencia de aminoácidos de DeMGMT Univ -NERCMSL

<400> 87

ES 2 627 117 T3

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15

Leu Pro Thr Ala Leu Ala Arg Ser Leu Gly Gln Pro Ala Pro Leu Glu  
 20 25 30

Arg Phe Ala Ser Arg Arg Pro Gln Val Leu Ala Val Arg Thr Val Cys  
 35 40 45

Asp Leu Val Leu Gly Lys Met Asp Lys Asp Cys Glu Met Lys Arg Thr  
 50 55 60

Thr Leu Asp Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln  
 65 70 75 80

Gly Leu His Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp  
 85 90 95

Ala Val Glu Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro  
 100 105 110

Leu Met Gln Cys Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu  
 115 120 125

Ala Ile Glu Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln  
 130 135 140

Gln Glu Ser Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val  
 145 150 155 160

Lys Phe Gly Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly  
 165 170 175

Asn Pro Lys Ala Ala Arg Ala Val Gly Gly Ala Met Arg Gly Asn Pro  
 180 185 190

ES 2 627 117 T3

Val Pro Ile Leu Ile Pro Cys His Arg Val Val Cys Ser Ser Gly Ala  
 195 200 205

Val Gly Asn Tyr Ser Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala  
 210 215 220

His Glu Gly His Arg Gly Lys Pro Gly Leu Gly Gly Ser Ser Gly Leu  
 225 230 235 240

Ala Gly Ala Trp Leu Lys Gly Ala Gly Ala Thr Ser Gly Ser Pro Pro  
 245 250 255

Ala Gly Arg Asn Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln Ser Asp  
 260 265 270

Ile Arg Ser Leu Ala Gly Glu Thr Gly Gln Glu Ala Ala Pro Leu Asp  
 275 280 285

Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser Leu Ser Pro Arg Gln  
 290 295 300

Leu Leu Gly Phe Pro Cys Ala Glu Val Ser Gly Leu Ser Thr Glu Arg  
 305 310 315 320

Val Arg Glu Leu Ala Val Ala Leu Ala Gln Lys Asn Val Lys Leu Ser  
 325 330 335

Thr Glu Gln Leu Arg Cys Leu Ala His Arg Leu Ser Glu Pro Pro Glu  
 340 345 350

Asp Leu Asp Ala Leu Pro Phe Asp Leu Leu Leu Phe Leu Asn Pro Asp  
 355 360 365

Ala Phe Ser Gly Pro Gln Ala Cys Thr Arg Phe Phe Ser Arg Ile Thr  
 370 375 380

Lys Ala Asn Val Asp Leu Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg  
 385 390 395 400

Leu Leu Pro Ala Ala Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu  
 405 410 415

Ser Glu Ala Asp Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro  
 420 425 430

Gly Arg Phe Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val

ES 2 627 117 T3

435	440	445
Ser Cys Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala		
450	455	460
Ala Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp Ser		
465	470	475
Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu Gly Gln		
485	490	495
Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala Trp Arg Gln		
500	505	510
Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu Arg Thr Ile Leu		
515	520	525
Arg Pro Arg Pro Gly Glu Asn Leu Tyr Phe Gln Gly Pro Gly Gly Gly		
530	535	540
Ser His His His His His His		
545	550	

<210> 88  
 <211> 543  
 <212> ADN  
 <213> Homo sapiens  
 <400> 88  
 gatatccagc tcttcacact acagaaggag ctggcagaac tccgagagtc taccagccag 60  
 atgcacacag catcatcttt ggagaagcaa atagccacc ccagtccacc ccctgaaaa 120  
 aaggagctga ggaaagtggc ccatttaaca ggcaagtcca actcaaggtc catgcctctg 180  
 gaatgggaag acacctatgg aattgtcctg ctttctggag tgaagtataa gaagggtggc 240  
 cttgtgatca atgaaactgg gctgtacttt gtatattcca aagtatactt ccgggggtcaa 300  
 tcttgcaaca acctgccct gagccacaag gtctacatga ggaactctaa gtatccccag 360  
 gatctggtga tgatggaggg gaagatgatg agctactgca ctactgggca gatgtgggcc 420  
 cgcagcagct acctgggggc agtgttcaat cttaccagtg ctgatcattt atatgtcaac 480  
 gtatctgagc tctctctggt caatcttgag gaatctcaga cgtttttcgg cttatataag 540  
 etc 543

10 <210> 89  
 <211> 1250  
 <212> ADN  
 <213> Secuencia artificial

15 <220>  
 <223> Secuencia de ADN de BiPmc-SNAP-sitio proTEV-sFasL  
 <400> 89

ES 2 627 117 T3

atgaagttat gcatattact ggccgtcgtg gcctttggtg gcctctcgct cgggagatct 60  
gacaaaagact gcgaaatgaa gcgcaccacc ctggatagcc ctctgggcaa gctggaactg 120  
tctgggtgcg aacagggcct gcacgagatc aagctgctgg gcaaaggaac atctgccgcc 180  
gacgccgtgg aagtgcctgc cccagccgcc gtgctgggcg gaccagagcc actgatgcag 240  
gccaccgcct ggctcaacgc ctactttcac cagcctgagg ccatcgagga gttccctgtg 300  
ccagccctgc accaccagcgt gttccagcag gagagcttta cccgccaggt gctgtggaaa 360  
ctgctgaaaag tggatgaagt cggagaggtc atcagctacc agcagctggc cgccctggcc 420  
ggcaatcccg ccgccaccgc cgcctgaaa accgccctga gcggaaatcc cgtgccatt 480  
ctgatcccct gccaccgggt ggtgtctagc tctggcgcg tggggggcta cgagggcggg 540  
ctgcctgtga aagagtggct gctggccac gagggccaca gactgggcaa gcctgggctg 600  
ggctcctgag gtataggggc gccagggctc ctaggtggcg gatctgaaaa cctctaattc 660  
cagagtgata tccagctctt ccacctacag aaggagctgg cagaactccg agagtctacc 720  
agccagatgc acacagcatc atctttggag aagcaaatag gccaccccag tccaccccct 780  
gaaaaaaagg agctgaggaa agtggcccat ttaacaggca agtccaactc aaggtccatg 840  
cctctggaat ggaagacac ctatggaatt gtctgcttt ctggagtga gtataagaag 900  
ggtggccttg tgatcaatga aactgggctg tactttgtat attccaaagt atacttccgg 960  
ggtcaatctt gcaacaacct gccctgagc cacaaggtct acatgaggaa ctctaagtat 1020  
ccccaggatc tggatgatg ggaggggaag atgatgagct actgcactac tgggcagatg 1080  
tgggcccgca gcagctacct gggggcagtg ttcaatctta ccagtgctga tcatttatat 1140  
gtcaacgtat ctgagctctc tctggtcaat tttgaggaat ctgagacgtt tttcgctta 1200  
tataagctcc cgggcgggtg aagtcacat catcatcatc attgaccggt 1250

<210> 90  
<211> 414  
5 <212> PRT  
<213> Secuencia artificial

<220>  
10 <223> Secuencia de aminoácidos de BiPmc-SNAP-sitio proTEV-sFasL

<400> 90  
Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

Leu Gly Arg Ser Asp Lys Asp Cys Glu Met Lys Arg Thr Thr Leu Asp  
20 25 30

ES 2 627 117 T3

Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu Gln Gly Leu His  
 35 40 45

Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala Asp Ala Val Glu  
 50 55 60

Val Pro Ala Pro Ala Ala Val Leu Gly Gly Pro Glu Pro Leu Met Gln  
 65 70 75 80

Ala Thr Ala Trp Leu Asn Ala Tyr Phe His Gln Pro Glu Ala Ile Glu  
 85 90 95

Glu Phe Pro Val Pro Ala Leu His His Pro Val Phe Gln Gln Glu Ser  
 100 105 110

Phe Thr Arg Gln Val Leu Trp Lys Leu Leu Lys Val Val Lys Phe Gly  
 115 120 125

Glu Val Ile Ser Tyr Gln Gln Leu Ala Ala Leu Ala Gly Asn Pro Ala  
 130 135 140

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

Leu Ile Pro Cys His Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly  
 165 170 175

Tyr Glu Gly Gly Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly  
 180 185 190

His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro  
 195 200 205

Gly Ser Leu Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln Ser Asp Ile  
 210 215 220

Gln Leu Phe His Leu Gln Lys Glu Leu Ala Glu Leu Arg Glu Ser Thr  
 225 230 235 240

Ser Gln Met His Thr Ala Ser Ser Leu Glu Lys Gln Ile Gly His Pro  
 245 250 255

Ser Pro Pro Pro Glu Lys Lys Glu Leu Arg Lys Val Ala His Leu Thr  
 260 265 270

Gly Lys Ser Asn Ser Arg Ser Met Pro Leu Glu Trp Glu Asp Thr Tyr

# ES 2 627 117 T3

275	280	285
Gly Ile Val Leu Leu Ser 290	Gly Val Lys Tyr Lys 295	Lys Gly Gly Leu Val 300
Ile Asn Glu Thr Gly 305	Leu Tyr Phe Val Tyr 310	Ser Lys Val Tyr Phe Arg 315 320
Gly Gln Ser Cys Asn 325	Asn Leu Pro Leu Ser 330	His Lys Val Tyr Met Arg 335
Asn Ser Lys Tyr Pro 340	Gln Asp Leu Val Met 345	Met Glu Gly Lys Met Met 350
Ser Tyr Cys Thr Thr 355	Gly Gln Met Trp Ala 360	Arg Ser Ser Tyr Leu Gly 365
Ala Val Phe Asn Leu Thr 370	Ser Ala Asp His Leu 375	Tyr Val Asn Val Ser 380
Glu Leu Ser Leu Val 385	Asn Phe Glu Glu Ser 390	Gln Thr Phe Phe Gly Leu 395 400
Tyr Lys Leu Pro 405	Gly Gly Gly Ser His 410	His His His His His

<210> 91  
 <211> 2976  
 5 <212> ADN  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 10 <223> secuencia de CNTN4 entre los aminoácidos 19 y 990

<b>&lt;400&gt; 91</b>	agatctatga ggttgccatg ggaactgctg gtactgcaat cattcatttt gtgccttgca	60
	gatgattcca cactgcatgg cccgattttt attcaagaac caagtctctgt aatgttcctt	120
	ttggattctg aggagaaaaa agtgaagctc aattgtgaag ttaaaggaaa tccaaaacct	180
	catatcaggt ggaagttaaa tggaacagat gttgacactg gtatggattt ccgctacagt	240
	gttgttgaag ggagcttggt gatcaataac cccaataaaa cccaagatgc tggaacgtac	300
	cagtgcacag cgacaaactc gtttgaaca attgttagca gagaagcaaa gctgcagttt	360
	gcttatcttg acaactttaa aacaagaaca agaagcactg tgtctgtccg tcgaggtcaa	420
	ggaatggtgc tactgtgtgg cccgccaccc cattctggag agctgagtta tgcctggatc	480
	ttcaatgaat acccttcccta tcaggataat cgccgctttg tttctcaaga gactgggaat	540



ES 2 627 117 T3

ctgtatattg ccaaagtaga aaaatcagat gttgggaatt atacctgtgt ggttaccaat	600
accgtgacaa accacaaggt cctggggcca cctacaccac taatattgag aaatgatgga	660
gtgatgggtg aatatgagcc caaaatagaa gtgcagttcc cagaacacagt tccgactgca	720
aaaggagcaa cggatgaagct ggaatgcttt gctttaggaa atccagtacc aactattatc	780
tggcgaagag ctgatggaaa gccaatagca aggaaagcca gaagacacaa gtcaaatgga	840
attcttgaga tccctaattt tcagcaggag gatgctggtt tatatgaatg tgtagctgaa	900
aattccagag ggaaaaatgt agcaagggga cagctaactt tctatgctca acctaattgg	960
attcaaaaaa taaatgatat tcacgtggcc atggaagaaa atgtcttttg ggaatgtaa	1020
gcaaatggaa ggccctagcc tacatacaag tggctaaaaa atggcgaacc tctgctaact	1080
cgggatagaa ttcaaattga gcaaggaaca ctcaacataa caatagtgaa cctctcagat	1140
gctggcatgt atcagtgttt ggcagagaat aaacatggag ttatcttttc caacgcagag	1200
cttagtgta tagctgtagg tccagatfff tcaagaacac tcttgaaaag agtaactctt	1260
gtcaaatggg gaggtgaagt tgtcattgag tgtaagccaa aagcgtctcc aaaacctggt	1320
tacacctgga agaaaggaag ggatatatta aaagaaaatg aaagaattac ctttctgaa	1380
gatgaaacc tcagaatcat caacgttact aaatcagacg ctgggagtta tacctgtata	1440
gccactaacc attttggaac tgctagcagt actggaaact tggtagtgaa agatccaaca	1500
agggtaatgg taccctctc cagtatggat gtcactgttg gagagagtat tgttttaccg	1560
tgccaggtaa cgcatgatca ctgcctagac atcgtgttta cttggtcatt taatggacac	1620
ctgatagact ttgacagaga tggggaccac tttgaaagag ttggagggca ggattcagct	1680
ggtgatttga tgatccgaaa catccaactg aagcatgctg ggaaatatgt ctgcatggtc	1740
caacaagtg tggaacaggt atctgctgct gcagacctga ttgtaagagg tccctcaggt	1800
ccccagagg ctgtgacaat agacgaaatc acagatacca ctgctcagct ctcctggaga	1860
cccgtcctg acaaccacag ccccatcacc atgtatgtca ttcaagccag gactccattc	1920
tccgtgggct ggcaagcagt cagtacagtc ccagaactca ttgatggga gacattcaca	1980
gcgacctgg tgggtttgaa cccttgggtt gaatatgaat tccgcacagt tgcagccaac	2040
gtgattggga ttggggagcc cagccgcccc tcagagaaac ggagaacaga agaagctctc	2100
cccgaagtca caccagcgaa tgctcagtggt ggcggaggca gcaaatctga actggttata	2160
acctgggaga cggctcctga ggaattacag aatggctcag gctttggtta tgtggtggcc	2220
ttccggccct acggtaaaat gatctggatg ctgacagtgc tggcctcagc tgatgcctct	2280
agatacgtgt tcaggaatga gagcgtgcac cccttctctc cctttgaggt taaagtaggt	2340
gtcttcaaca acaaaggaga aggccctttc agtcccacca cgggtggtgta ttctgcagaa	2400

ES 2 627 117 T3

gaagaaccca ccaaaccacc agccagtatc tttgccagaa gtctttctgc cacagatatt 2460  
 gaagttttct gggcctcccc actggagaag aatagaggac gaatacaagg ttatgaggtt 2520  
 aaatattgga gacatgaaga caaagaagaa aatgctagaa aaatacgaac agttggaat 2580  
 cagacatcaa caaaaatcac gaacttaaaa ggcagtgctg tgtatcactt agctgtcaag 2640  
 gcatataatt ctgctgggac aggccctct agtgcaacag tcaatgtgac aaccgaaag 2700  
 ccaccaccaa gtcaaccccc cggaaacatc atatggaatt catcagactc caaaattatc 2760  
 ctgaattggg atcaagtgaa ggccctggat aatgagtcgg aagtaaaagg atacaaagtc 2820  
 ttgtacagat ggaacagaca aagcagcaca tctgtcattg aaacaaataa aacatcggtg 2880  
 gagctttctt tgcctttcga tgaagattat ataatagaaa ttaagccatt cagcgacgga 2940  
 ggagatggca gcagcagtga acaaattcga attccc 2976

<210> 92  
 <211> 3688  
 <212> ADN  
 <213> Secuencia artificial

5

<220>  
 <223> Secuencia de ADN de BiP-CNTN419-990-SNAP-ProTEV-Etiqueta de His

10

<400> 92  
 atgaagtat gcatattact ggccgtcgtg gcctttgttg gcctctcgtc cgggagatct 60  
 atgaggttgc catgggaact gctggtagct caatcattca ttttgtgcct tgcagatgat 120  
 tccacactgc atggcccgat ttttattcaa gaaccaagtc ctgtaatgtt ccctttggat 180  
 tctgaggaga aaaaagtgaa gctcaattgt gaagttaaag gaaatccaaa acctcatatc 240  
 aggtggaagt taaatggaac agatgttgac actggtatgg atttccgcta cagtgttgtt 300  
 gaagggagct tgttgatcaa taaccccaat aaaacccaag atgctggaac gtaccagtgc 360  
 acagcgacaa actcgttttg aacaattgtt agcagagaag caaagctgca gtttgcttat 420  
 cttgacaact ttaaaacaag aacaagaagc actgtgtctg tccgtcgagg tcaaggaatg 480  
 gtgctactgt gtggcccgcc accccattct ggagagctga gttatgctg gatcttcaat 540  
 gaataccctt cctatcagga taatcgccgc tttgtttctc aagagactgg gaatctgtat 600  
 attgccaaag tagaaaaatc agatgttggg aattatacct gtgtggttac caataccgtg 660  
 acaaaccaca aggtcctggg gccacctaca ccactaatat tgagaaatga tggagtgatg 720  
 ggtgaatatg agcccaaat agaagtgcag ttcccagaaa cagttccgac tgcaaaagga 780  
 gcaacggtga agctggaatg ctttgcttta ggaaatccag taccaactat tatctggcga 840  
 agagctgatg gaaagccaat agcaaggaaa gccagaagac acaagtcaaa tggattctt 900  
 gagatcccta attttcagca ggaggatgct ggtttatatg aatgtgtagc tgaaaattcc 960  
 agagggaaaa atgtagcaag gggacagcta actttctatg ctcaacctaa ttggattcaa 1020

ES 2 627 117 T3

aaaataaatg atattcacgt ggccatggaa gaaaatgtct tttgggaatg taaagcaaat 1080  
 ggaaggccta agcctacata caagtggcta aaaaatggcg aacctctgct aactcgggat 1140  
 agaattcaaa ttgagcaag aacactcaac ataacaatag tgaacctctc agatgctggc 1200  
 atgtatcagt gtttggcaga gaataaacat ggagttatct tttccaacgc agagcttagt 1260  
 gttatagctg taggtccaga tttttcaaga aacctcttga aaagagtaac tcttgtcaaa 1320  
 gtgggaggtg aagttgtcat tgagtgtgtaag ccaaaagcgt ctccaaaacc tgtttacacc 1380  
 tggaaгааag gaagggatat attaaaagaa aatgaaagaa ttaccatttc tgaagatgga 1440  
 aacctcagaa tcatacaactg tactaaatca gacgctggga gttatacctg tatagccact 1500  
 aaccatthttg gaactgctag cagtactgga aacttggtag tgaagatcc aacaagggtg 1560  
 atggtacccc cttccagtat ggatgtcact gttggagaga gtattgtttt accgtgccag 1620  
 gtaacgcatg atcactcgtc agacatcgtg tttacttggc catttaatgg acacctgata 1680  
 gactttgaca gagatgggga ccactttgaa agagttggag ggcaggattc agctggtgat 1740  
 ttgatgatcc gaaacatcca actgaagcat gctgggaaat atgtctgcat ggtccaaaca 1800  
 agtgtggaca ggctatctgc tgctgcagac ctgattgtaa gaggtcctcc aggtcccca 1860  
 gaggtctgta caatagacga aatcacagat accactgctc agctctcctg gagaccgggt 1920  
 cctgacaacc acagcccat caccatgtat gtcattcaag ccaggactcc attctccgtg 1980  
 ggctggcaag cagtcagtac agtcccagaa ctcatatgat ggaagacatt cacagcgacc 2040  
 gtggtgggtt tgaacccttg ggttgaatat gaattccgca cagttgcagc caacgtgatt 2100  
 gggattgggg agcccagccg ccctcagag aaacggagaa cagaagaagc tctcccggaa 2160  
 gtcacaccag cgaatgtcag tgggtggcgga ggcagcaaat ctgaactggt tataacctgg 2220  
 gagacggtcc ctgaggaatt acagaatggt cgaggctttg gttatgtggt ggccttcggg 2280  
 ccctacggta aatgatctg gatgctgaca gtgctggcct cagctgatgc ctctagatac 2340  
 gtgttcagga atgagagcgt gcaccccttc tctccctttg aggttaaagt aggtgtcttc 2400  
 aacaacaaag gagaaggccc tttcagtccc accacggtgg tgtattctgc agaagaagaa 2460  
 cccaccaaac caccagccag tatctttgcc agaagtcttt ctgccacaga tattgaagtt 2520  
 ttctgggctt cccactgga gaagaataga ggacgaatac aaggttatga ggttaaatat 2580  
 tggagacatg aagacaaaga agaaaatgct agaaaaatac gaacagttgg aatcagaca 2640  
 tcaacaaaaa tcacgaactt aaaaggcagt gtgctgtatc acttagctgt caaggcatat 2700  
 aattctgctg ggacaggccc ctctagtgca acagtcaatg tgacaaccgg aaagccacca 2760  
 ccaagtcaac cccccgaaa catcatatgg aattcatcag actccaaaat tatcctgaat 2820  
 tgggatcaag tgaaggccct ggataatgag tcggaagtaa aaggatacaa agtcttgtac 2880

ES 2 627 117 T3

```

agatggaaca gacaaagcag cacatctgtc attgaaacaa ataaaacatc ggtggagctt      2940
tctttgcctt tcgatgaaga ttatataata gaaattaagc cattcagcga cggaggagat      3000
ggcagcagca gtgaacaaat tcgaattccc gggggaggta gcaaagactg cgaaatgaag      3060
cgcaccaccc tggatagccc tctgggcaag ctggaactgt ctgggtgcga acagggcctg      3120
cacgagatca agctgctggg caaaggaaca tctgccgccg acgccgtgga agtgcctgcc      3180
ccagccgccg tgctgggagg accagagcca ctgatgcagg ccaccgcctg gctcaacgcc      3240
tactttcacc agcctgaggc catcgaggag ttccctgtgc cagccctgca ccaccagtg      3300
ttccagcagg agagctttac ccgccaggtg ctgtggaaac tgctgaaagt ggtgaagttc      3360
ggagaggtea tcagctacca gcagctggcc gccctggccg gcaatcccgc cgccaccgcc      3420
gccgtgaaaa ccgccctgag cggaaatccc gtgccattc tgatcccctg ccaccgggtg      3480
gtgtctagct ctggcgccgt ggggggctac gagggcgggc tcgccgtgaa agagtggctg      3540
ctggcccacg agggccacag actgggcaag cctgggctgg gtcctgcagg tataggcgcg      3600
ccaggttccc tggagaatct atattttcaa agtggcggag gtagccatca tcatcatcat      3660
cattgatgac cggtaagctt gcggccgc                                          3688

```

<210> 93  
 <211> 1221  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Secuencia de aminoácidos de BiP-CNTN419-990-SNAP-ProTEV-Etiqueta de His

<400> 93  
 Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
 1 5 10 15  
  
 Leu Gly Arg Ser Met Arg Leu Pro Trp Glu Leu Leu Val Leu Gln Ser  
 20 25 30  
  
 Phe Ile Leu Cys Leu Ala Asp Asp Ser Thr Leu His Gly Pro Ile Phe  
 35 40 45  
  
 Ile Gln Glu Pro Ser Pro Val Met Phe Pro Leu Asp Ser Glu Glu Lys  
 50 55 60  
  
 Lys Val Lys Leu Asn Cys Glu Val Lys Gly Asn Pro Lys Pro His Ile  
 65 70 75 80  
  
 Arg Trp Lys Leu Asn Gly Thr Asp Val Asp Thr Gly Met Asp Phe Arg  
 85 90 95

ES 2 627 117 T3

Tyr Ser Val Val Glu Gly Ser Leu Leu Ile Asn Asn Pro Asn Lys Thr  
 100 105 110

Gln Asp Ala Gly Thr Tyr Gln Cys Thr Ala Thr Asn Ser Phe Gly Thr  
 115 120 125

Ile Val Ser Arg Glu Ala Lys Leu Gln Phe Ala Tyr Leu Asp Asn Phe  
 130 135 140

Lys Thr Arg Thr Arg Ser Thr Val Ser Val Arg Arg Gly Gln Gly Met  
 145 150 155 160

Val Leu Leu Cys Gly Pro Pro Pro His Ser Gly Glu Leu Ser Tyr Ala  
 165 170 175

Trp Ile Phe Asn Glu Tyr Pro Ser Tyr Gln Asp Asn Arg Arg Phe Val  
 180 185 190

Ser Gln Glu Thr Gly Asn Leu Tyr Ile Ala Lys Val Glu Lys Ser Asp  
 195 200 205

Val Gly Asn Tyr Thr Cys Val Val Thr Asn Thr Val Thr Asn His Lys  
 210 215 220

Val Leu Gly Pro Pro Thr Pro Leu Ile Leu Arg Asn Asp Gly Val Met  
 225 230 235 240

Gly Glu Tyr Glu Pro Lys Ile Glu Val Gln Phe Pro Glu Thr Val Pro  
 245 250 255

Thr Ala Lys Gly Ala Thr Val Lys Leu Glu Cys Phe Ala Leu Gly Asn  
 260 265 270

Pro Val Pro Thr Ile Ile Trp Arg Arg Ala Asp Gly Lys Pro Ile Ala  
 275 280 285

Arg Lys Ala Arg Arg His Lys Ser Asn Gly Ile Leu Glu Ile Pro Asn  
 290 295 300

Phe Gln Gln Glu Asp Ala Gly Leu Tyr Glu Cys Val Ala Glu Asn Ser  
 305 310 315 320

Arg Gly Lys Asn Val Ala Arg Gly Gln Leu Thr Phe Tyr Ala Gln Pro  
 325 330 335

Asn Trp Ile Gln Lys Ile Asn Asp Ile His Val Ala Met Glu Glu Asn  
 340 345 350

ES 2 627 117 T3

Val Phe Trp Glu Cys Lys Ala Asn Gly Arg Pro Lys Pro Thr Tyr Lys  
 355 360 365

Trp Leu Lys Asn Gly Glu Pro Leu Leu Thr Arg Asp Arg Ile Gln Ile  
 370 375 380

Glu Gln Gly Thr Leu Asn Ile Thr Ile Val Asn Leu Ser Asp Ala Gly  
 385 390 395 400

Met Tyr Gln Cys Leu Ala Glu Asn Lys His Gly Val Ile Phe Ser Asn  
 405 410 415

Ala Glu Leu Ser Val Ile Ala Val Gly Pro Asp Phe Ser Arg Thr Leu  
 420 425 430

Leu Lys Arg Val Thr Leu Val Lys Val Gly Gly Glu Val Val Ile Glu  
 435 440 445

Cys Lys Pro Lys Ala Ser Pro Lys Pro Val Tyr Thr Trp Lys Lys Gly  
 450 455 460

Arg Asp Ile Leu Lys Glu Asn Glu Arg Ile Thr Ile Ser Glu Asp Gly  
 465 470 475 480

Asn Leu Arg Ile Ile Asn Val Thr Lys Ser Asp Ala Gly Ser Tyr Thr  
 485 490 495

Cys Ile Ala Thr Asn His Phe Gly Thr Ala Ser Ser Thr Gly Asn Leu  
 500 505 510

Val Val Lys Asp Pro Thr Arg Val Met Val Pro Pro Ser Ser Met Asp  
 515 520 525

Val Thr Val Gly Glu Ser Ile Val Leu Pro Cys Gln Val Thr His Asp  
 530 535 540

His Ser Leu Asp Ile Val Phe Thr Trp Ser Phe Asn Gly His Leu Ile  
 545 550 555 560

Asp Phe Asp Arg Asp Gly Asp His Phe Glu Arg Val Gly Gly Gln Asp  
 565 570 575

Ser Ala Gly Asp Leu Met Ile Arg Asn Ile Gln Leu Lys His Ala Gly  
 580 585 590

Lys Tyr Val Cys Met Val Gln Thr Ser Val Asp Arg Leu Ser Ala Ala  
 595 600 605

ES 2 627 117 T3

Ala Asp Leu Ile Val Arg Gly Pro Pro Gly Pro Pro Glu Ala Val Thr  
610 615 620

Ile Asp Glu Ile Thr Asp Thr Thr Ala Gln Leu Ser Trp Arg Pro Gly  
625 630 635 640

Pro Asp Asn His Ser Pro Ile Thr Met Tyr Val Ile Gln Ala Arg Thr  
645 650 655

Pro Phe Ser Val Gly Trp Gln Ala Val Ser Thr Val Pro Glu Leu Ile  
660 665 670

Asp Gly Lys Thr Phe Thr Ala Thr Val Val Gly Leu Asn Pro Trp Val  
675 680 685

Glu Tyr Glu Phe Arg Thr Val Ala Ala Asn Val Ile Gly Ile Gly Glu  
690 695 700

Pro Ser Arg Pro Ser Glu Lys Arg Arg Thr Glu Glu Ala Leu Pro Glu  
705 710 715 720

Val Thr Pro Ala Asn Val Ser Gly Gly Gly Gly Ser Lys Ser Glu Leu  
725 730 735

Val Ile Thr Trp Glu Thr Val Pro Glu Glu Leu Gln Asn Gly Arg Gly  
740 745 750

Phe Gly Tyr Val Val Ala Phe Arg Pro Tyr Gly Lys Met Ile Trp Met  
755 760 765

Leu Thr Val Leu Ala Ser Ala Asp Ala Ser Arg Tyr Val Phe Arg Asn  
770 775 780

Glu Ser Val His Pro Phe Ser Pro Phe Glu Val Lys Val Gly Val Phe  
785 790 795 800

Asn Asn Lys Gly Glu Gly Pro Phe Ser Pro Thr Thr Val Val Tyr Ser  
805 810 815

Ala Glu Glu Glu Pro Thr Lys Pro Pro Ala Ser Ile Phe Ala Arg Ser  
820 825 830

Leu Ser Ala Thr Asp Ile Glu Val Phe Trp Ala Ser Pro Leu Glu Lys  
835 840 845

Asn Arg Gly Arg Ile Gln Gly Tyr Glu Val Lys Tyr Trp Arg His Glu





ES 2 627 117 T3

Pro Val Phe Gln Gln Glu Ser Phe Thr Arg Gln Val Leu Trp Lys  
 1100 1105 1110

Leu Leu Lys Val Val Lys Phe Gly Glu Val Ile Ser Tyr Gln Gln  
 1115 1120 1125

Leu Ala Ala Leu Ala Gly Asn Pro Ala Ala Thr Ala Ala Val Lys  
 1130 1135 1140

Thr Ala Leu Ser Gly Asn Pro Val Pro Ile Leu Ile Pro Cys His  
 1145 1150 1155

Arg Val Val Ser Ser Ser Gly Ala Val Gly Gly Tyr Glu Gly Gly  
 1160 1165 1170

Leu Ala Val Lys Glu Trp Leu Leu Ala His Glu Gly His Arg Leu  
 1175 1180 1185

Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala Pro Gly Ser  
 1190 1195 1200

Leu Glu Asn Leu Tyr Phe Gln Ser Gly Gly Gly Ser His His His  
 1205 1210 1215

His His His  
 1220

<210> 94

<211> 2538

<212> ADN

<213> Homo sapiens

<400> 94

ttcctgtcgc accaccgcct gaaaggcagg ttccagaggg accgcaggaa catccgcccc 60

aacatcatcc tgggtgctgac ggacgaccag gatgtggagc tgggttccat gcaggtgatg 120

aacaagaccc ggcgcatcat ggagcagggc ggggcgcact tcatcaacgc cttcgtgacc 180

acaccatgt gctgcccctc acgctcctcc atcctcactg gcaagtacgt ccacaaccac 240

aacacctaca ccaacaatga gaactgctcc tcgccctcct ggcaggcaca gcacgagagc 300

cgcacctttg ccgtgtacct caatagcact ggctaccgga cagctttctt cgggaagtat 360

cttaatgaat acaacggctc ctacgtgcca cccggctgga aggagtgggt cggactcctt 420

aaaaactccc gcttttataa ctacacgctg tgtcggaacg gggtgaaaga gaagcacggc 480

tccgactact ccaaggatta cctcacagac ctcatcacca atgacagcgt gagcttcttc 540

cgcacgtcca agaagatgta cccgcacagg ccagtcctca tggatcatcag ccatgcagcc 600

ccccacggcc ctgaggattc agccccacaa tattcacgcc tcttcccaaa cgcattctcag 660

5

ES 2 627 117 T3

cacatcacgc cgagctacaa ctacgcgccc aaccggaca aacctggat catgcgctac 720  
acggggccca tgaagcccat ccacatggaa ttcaccaaca tgctccagcg gaagcgcttg 780  
cagaccctca tgtcgggtga cgactccatg gagacgattt acaacatgct gggtgagacg 840  
ggcgagctgg acaacacgta catcgtatac accgcccacc acggttacca catcggccag 900  
tttggcctgg tgaaggaa atccatgcca tatgagtttg acatcagggt cccgttctac 960  
gtgaggggccc ccaacgtgga agccggctgt ctgaatcccc acatcgtcct caacattgac 1020  
ctggccccc ccatcctgga cattgcaggc ctggacatac ctgcggatat ggacgggaaa 1080  
tccatcctca agctgctgga cacggagcgg ccggtgaatc ggtttcaact gaaaaagaag 1140  
atgaggtct ggcgggactc cttcttggtg gagagaggca agctgctaca caagagagac 1200  
aatgacaagg tggacgccc ggaggagaac tttctgccc agtaccagcg tgtgaaggac 1260  
ctgtgtcagc gtgctgagta ccagacggcg tgtgagcagc tgggacagaa gtggcagtgt 1320  
gtggaggacg ccacggggaa gctgaagctg cataagtgca agggcccat gcggctgggc 1380  
ggcagcagag ccctctcaa cctcgtgcc aagtactacg ggcagggcag cgaggcctgc 1440  
acctgtgaca gcggggacta caagctcagc ctggccggac gccgaaaaa actcttcaag 1500  
aagaagtaca aggccagcta tgtcccgagt cgctccatcc gctcagtggc catcgaggtg 1560  
gacggcaggg tgtaccacgt aggcctgggt gatgccgcc agccccgaaa cctaccaag 1620  
cggcactggc caggggcccc tgaggaccaa gatgacaagg atggtgggga cttcagtggc 1680  
actggaggcc tccccgacta ctcagccgcc aaccattta aagtacaca tcggtgctac 1740  
atcctagaga acgacacagt ccagtgtgac ctggacctgt acaagtccct gcaggcctgg 1800  
aaagaccaca agctgcacat cgaccacgag attgaaacct tgcagaaca aattaagaac 1860  
ctgaggggaa tccgaggtca cctgaagaaa aagcggccag aagaatgtga ctgtcacaaa 1920  
atcagctacc acaccagca caaaggccgc ctcaagcaca gaggctccag tctgcatcct 1980  
ttcaggaagg gcctgcaaga gaaggacaag gtgtggctgt tgcgggagca gaagcgcaag 2040  
aagaaactcc gcaagctgct caagcgcctg cagaacaacg acacgtgag catgccaggc 2100  
ctcacgtgct tcaccacga caaccagcac tggcagacgg cgccttctg gacactgggg 2160  
cctttctgtg cctgcaccag cgccaacaat aacacgtact ggtgcatgag gaccatcaat 2220  
gagactcaca atttctctt ctgtgaattt gcaactggct tcctagagta ctttgatctc 2280  
aacacagacc cctaccagct gatgaatgca gtgaacacac tggacagga tgtcctcaac 2340  
cagctacacg tacagctcat ggagctgagg agctgcaagg gttacaagca gtgtaacccc 2400  
cggactcgaa acatggacct gggacttaa gatggaggaa gctatgagca atacaggcag 2460  
tttcagcgtc gaaagtggcc agaaatgaag agaccttctt ccaaatcact gggacaactg 2520  
tggaaggct gggaaggc 2538

- 5 <210> 95
- <211> 879
- <212> PRT
- <213> Homo sapiens
- <400> 95

ES 2 627 117 T3

Gly His Arg Leu Gly Lys Pro Gly Leu Gly Pro Ala Gly Ile Gly Ala  
 1 5 10 15

Pro Gly Ser Leu Gly Gly Gly Ser Glu Asn Leu Tyr Phe Gln Ser Asp  
 20 25 30

Ile Phe Leu Ser His His Arg Leu Lys Gly Arg Phe Gln Arg Asp Arg  
 35 40 45

Arg Asn Ile Arg Pro Asn Ile Ile Leu Val Leu Thr Asp Asp Gln Asp  
 50 55 60

Val Glu Leu Gly Ser Met Gln Val Met Asn Lys Thr Arg Arg Ile Met  
 65 70 75 80

Glu Gln Gly Gly Ala His Phe Ile Asn Ala Phe Val Thr Thr Pro Met  
 85 90 95

Cys Cys Pro Ser Arg Ser Ser Ile Leu Thr Gly Lys Tyr Val His Asn  
 100 105 110

His Asn Thr Tyr Thr Asn Asn Glu Asn Cys Ser Ser Pro Ser Trp Gln  
 115 120 125

Ala Gln His Glu Ser Arg Thr Phe Ala Val Tyr Leu Asn Ser Thr Gly  
 130 135 140

Tyr Arg Thr Ala Phe Phe Gly Lys Tyr Leu Asn Glu Tyr Asn Gly Ser  
 145 150 155 160

Tyr Val Pro Pro Gly Trp Lys Glu Trp Val Gly Leu Leu Lys Asn Ser  
 165 170 175

Arg Phe Tyr Asn Tyr Thr Leu Cys Arg Asn Gly Val Lys Glu Lys His  
 180 185 190

Gly Ser Asp Tyr Ser Lys Asp Tyr Leu Thr Asp Leu Ile Thr Asn Asp  
 195 200 205

Ser Val Ser Phe Phe Arg Thr Ser Lys Lys Met Tyr Pro His Arg Pro



ES 2 627 117 T3

Glu Gln Leu Gly Gln Lys Trp Gln Cys Val Glu Asp Ala Thr Gly Lys  
 465 470 475 480  
 Leu Lys Leu His Lys Cys Lys Gly Pro Met Arg Leu Gly Gly Ser Arg  
 485 490 495  
 Ala Leu Ser Asn Leu Val Pro Lys Tyr Tyr Gly Gln Gly Ser Glu Ala  
 500 505 510  
 Cys Thr Cys Asp Ser Gly Asp Tyr Lys Leu Ser Leu Ala Gly Arg Arg  
 515 520 525  
 Lys Lys Leu Phe Lys Lys Lys Tyr Lys Ala Ser Tyr Val Arg Ser Arg  
 530 535 540  
 Ser Ile Arg Ser Val Ala Ile Glu Val Asp Gly Arg Val Tyr His Val  
 545 550 555 560  
 Gly Leu Gly Asp Ala Ala Gln Pro Arg Asn Leu Thr Lys Arg His Trp  
 565 570 575  
 Pro Gly Ala Pro Glu Asp Gln Asp Asp Lys Asp Gly Gly Asp Phe Ser  
 580 585 590  
 Gly Thr Gly Gly Leu Pro Asp Tyr Ser Ala Ala Asn Pro Ile Lys Val  
 595 600 605  
 Thr His Arg Cys Tyr Ile Leu Glu Asn Asp Thr Val Gln Cys Asp Leu  
 610 615 620  
 Asp Leu Tyr Lys Ser Leu Gln Ala Trp Lys Asp His Lys Leu His Ile  
 625 630 635 640  
 Asp His Glu Ile Glu Thr Leu Gln Asn Lys Ile Lys Asn Leu Arg Glu  
 645 650 655  
 Val Arg Gly His Leu Lys Lys Lys Arg Pro Glu Glu Cys Asp Cys His  
 660 665 670  
 Lys Ile Ser Tyr His Thr Gln His Lys Gly Arg Leu Lys His Arg Gly  
 675 680 685  
 Ser Ser Leu His Pro Phe Arg Lys Gly Leu Gln Glu Lys Asp Lys Val  
 690 695 700  
 Trp Leu Leu Arg Glu Gln Lys Arg Lys Lys Lys Leu Arg Lys Leu Leu  
 705 710 715 720

ES 2 627 117 T3

Lys Arg Leu Gln Asn Asn Asp Thr Cys Ser Met Pro Gly Leu Thr Cys  
 725 730 735

Phe Thr His Asp Asn Gln His Trp Gln Thr Ala Pro Phe Trp Thr Leu  
 740 745 750

Gly Pro Phe Cys Ala Cys Thr Ser Ala Asn Asn Asn Thr Tyr Trp Cys  
 755 760 765

Met Arg Thr Ile Asn Glu Thr His Asn Phe Leu Phe Cys Glu Phe Ala  
 770 775 780

Thr Gly Phe Leu Glu Tyr Phe Asp Leu Asn Thr Asp Pro Tyr Gln Leu  
 785 790 795 800

Met Asn Ala Val Asn Thr Leu Asp Arg Asp Val Leu Asn Gln Leu His  
 805 810 815

Val Gln Leu Met Glu Leu Arg Ser Cys Lys Gly Tyr Lys Gln Cys Asn  
 820 825 830

Pro Arg Thr Arg Asn Met Asp Leu Gly Leu Lys Asp Gly Gly Ser Tyr  
 835 840 845

Glu Gln Tyr Arg Gln Phe Gln Arg Arg Lys Trp Pro Glu Met Lys Arg  
 850 855 860

Pro Ser Ser Lys Ser Leu Gly Gln Leu Trp Glu Gly Trp Glu Gly  
 865 870 875

<210> 96  
 <211> 3322  
 5 <212> ADN  
 <213> Secuencia artificial

<220>  
 10 <223> Secuencia de ADN de proteína de tipo BiP-SNAP-hSULF2

<400> 96  
 tcgcgagcta gcacccatgaa actatgtatt ctacttgacg ttggtgctgt cgtaggattg 60  
 tccttaccta cagctctggc aagatctgac aaagactgcg aaatgaaaag aactacattg 120  
 gattcaccac ttgggaagtt ggaactgagt ggatgacgagc aaggattgca tgaaattaag 180  
 ctactgggaa aaggaacttc tgctgctgat gcagttgaag ttccagcacc agcagctggt 240  
 cttggaggtc ctgagcccct catgcaagcc acagcctggc ttaacgcata tttccaccag 300  
 cctgaggcca ttgaggaatt tccagtcacc gcccttcacc atcctgtggt tcagcaggag 360  
 agcttcaccc gccaggtcct gtggaaattg ctgaaggtgg tcaagtttgg tgaagtgatt 420

ES 2 627 117 T3

tcatatcagc aacttgctgc attggccggg aaccccgag ctacagctgc cgtgaaaact	480
gctctcagcg gaaatcctgt gcccatcctg atcccttgtc acagagtcgt ttcattctcc	540
ggagctgtag gtggctatga aggaggactg gcagttaagg agtggctgct ggctcatgaa	600
ggtcatagac ttggaagcc tgggctgggt cctgctggta taggcgcgcc agggcccta	660
ggtggcggat ccgaaaacct gtacttccag agcgatatct tcctgtcgca ccaccgctg	720
aaaggcaggt ttcagagga ccgcaggaac atccgcccc aatcatcct ggtgctgacg	780
gacgaccagg atgtggagct gggttccatg caggtgatga acaagaccg gcgcatcatg	840
gagcagggcg gggcgactt catcaacgcc ttcgtgacca caccatgtg ctgcccctca	900
cgctcctcca tcctcactgg caagtacgtc cacaaccaca acacctacac caacaatgag	960
aactgctcct cgccctcctg gcaggcacag cacgagagcc gcaccttgc cgtgtacctc	1020
aatagcactg gctaccggac agctttcttc gggaagtatc ttaatgaata caacggctcc	1080
tacgtgccac ccggctggaa ggagtgggtc ggactcotta aaaactcccg cttttataac	1140
tacacgctgt gtcggaacgg ggtgaaagag aagcacggct ccgactactc caaggattac	1200
ctcacagacc tcatcaccaa tgacagcgtg agcttcttcc gcacgtcca gaagatgtac	1260
ccgcacaggg cagtccatc ggtcatcagc catgcagccc cccacggccc tgaggattca	1320
gccccacaat attcacgcct cttcccaaac gcctctcagc acatcacgcc gagctacaac	1380
tacgcgcccc acccggacaa aactggatc atgcgctaca cggggcccat gaagccatc	1440
cacatggaat tcaccaacat gctccagcgg aagcgcttgc agaccctcat gtcggtggac	1500
gactccatgg agacgattta caacatgctg gttgagacgg gcgagctgga caacacgtac	1560
atcgtataca ccgccgacca cggttaccac atcggccagt ttggcctggt gaaagggaaa	1620
tccatgceat atgagtttga catcagggtc ccgttctacg tgaggggccc caacgtggaa	1680
gccggctgtc tgaatccca catcgtctc aacattgacc tggccccac catcctggac	1740
attgcaggcc tggacatacc tgcggatatg gacgggaaat ccatcctcaa gctgctggac	1800
acggagcggc cggtgaatcg gttcacttg aaaaagaaga tgagggctctg gcgggactcc	1860
ttcttgggtg agagagcga gctgctacac aagagagaca atgacaaggt ggacgcccag	1920
gaggagaact ttctgcccga gtaccagcgt gtgaaggacc tgtgtcagcg tgctgagtac	1980
cagacggcgt gtgagcagct gggacagaag tggcagtgtg tggaggacgc cacggggaag	2040
ctgaagctgc ataagtgcaa gggccccatg cggctgggcg gcagcagagc cctctccaac	2100
ctcgtgcccga agtactacgg gcagggcagc gaggcctgca cctgtgacag cggggactac	2160
aagctcagcc tggccggacg ccggaaaaaa ctcttcaaga agaagtacaa ggccagctat	2220
gtccgcagtc gctccatccg ctcagtgccc atcgaggtgg acggcagggg gtaccacgta	2280

ES 2 627 117 T3

```

ggcctgggtg atgccgcccc gccccgaaac ctcaccaagc ggcaactggcc aggggcccct      2340
gaggaccaag atgacaagga tgggtggggac ttcagtggca ctggaggcct tcccgactac      2400
tcagccgcca accccattaa agtgacacat cggtgctaca tcctagagaa cgacacagtc      2460
cagtgtgacc tggacctgta caagtccttg caggcctgga aagaccacaa gctgcacatc      2520
gaccacgaga ttgaaaccct gcagaacaaa attaagaacc tgaggggaagt ccgaggtcac      2580
ctgaagaaaa agcggccaga agaatgtgac tgtcacaaaa tcagctacca caccagcac      2640
aaaggccgcc tcaagcacag aggctccagt ctgcatcctt tcaggaaggg cctgcaagag      2700
aaggacaagg tgtggctggt gcgggagcag aagcgcaaga agaaactccg caagctgctc      2760
aagcgctgc agaacaacga cacgtgcagc atgccaggcc tcacgtgctt caccacagac      2820
aaccagcact ggcagacggc gcctttctgg aactggggc ctttctgtgc ctgcaccagc      2880
gccaacaata acacgtactg gtgcatgagg accatcaatg agactcacia tttcctcttc      2940
tgtgaatttg caactggctt cctagagtac tttgatctca acacagacc caccagctg      3000
atgaatgcag tgaacacact ggacagggat gtctcaacc agctacacgt acagctcatg      3060
gagctgagga gctgcaaggg ttacaagcag tgtaaccccc ggactcgaaa catggacctg      3120
ggacttaaag atggaggaag ctatgagcaa tacaggcagt ttcagcgtcg aaagtggcca      3180
gaaatgaaga gaccttcttc caaatcactg ggacaactgt ggaaggctg ggaaggcccg      3240
ggagagaatc tataatttca agggcccggc ggaggtagtc accatcatca ccatcactaa      3300
tgaccggtgc ggcgcaagc tt                                     3322

```

<210> 97

<211> 1094

5 <212> PRT

<213> Secuencia artificial

<220>

10 <223> Secuencia de aminoácidos de proteína de tipo BiP-SNAP-SULF2\_h

<400> 97

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser  
1 5 10 15

Leu Pro Thr Ala Leu Ala Arg Ser Asp Lys Asp Cys Glu Met Lys Arg  
20 25 30

Thr Thr Leu Asp Ser Pro Leu Gly Lys Leu Glu Leu Ser Gly Cys Glu  
35 40 45

Gln Gly Leu His Glu Ile Lys Leu Leu Gly Lys Gly Thr Ser Ala Ala  
50 55 60



ES 2 627 117 T3

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

ES 2 627 117 T3

Ser Trp Gln Ala Gln His Glu Ser Arg Thr Phe Ala Val Tyr Leu Asn  
 325 330 335

Ser Thr Gly Tyr Arg Thr Ala Phe Phe Gly Lys Tyr Leu Asn Glu Tyr  
 340 345 350

Asn Gly Ser Tyr Val Pro Pro Gly Trp Lys Glu Trp Val Gly Leu Leu  
 355 360 365

Lys Asn Ser Arg Phe Tyr Asn Tyr Thr Leu Cys Arg Asn Gly Val Lys  
 370 375 380

Glu Lys His Gly Ser Asp Tyr Ser Lys Asp Tyr Leu Thr Asp Leu Ile  
 385 390 400

Thr Asn Asp Ser Val Ser Phe Phe Arg Thr Ser Lys Lys Met Tyr Pro  
 405 410 415

His Arg Pro Val Leu Met Val Ile Ser His Ala Ala Pro His Gly Pro  
 420 425 430

Glu Asp Ser Ala Pro Gln Tyr Ser Arg Leu Phe Pro Asn Ala Ser Gln  
 435 440 445

His Ile Thr Pro Ser Tyr Asn Tyr Ala Pro Asn Pro Asp Lys His Trp  
 450 455 460

Ile Met Arg Tyr Thr Gly Pro Met Lys Pro Ile His Met Glu Phe Thr  
 465 470 475 480

Asn Met Leu Gln Arg Lys Arg Leu Gln Thr Leu Met Ser Val Asp Asp  
 485 490 495

Ser Met Glu Thr Ile Tyr Asn Met Leu Val Glu Thr Gly Glu Leu Asp  
 500 505 510

Asn Thr Tyr Ile Val Tyr Thr Ala Asp His Gly Tyr His Ile Gly Gln  
 515 520 525

Phe Gly Leu Val Lys Gly Lys Ser Met Pro Tyr Glu Phe Asp Ile Arg  
 530 535 540

Val Pro Phe Tyr Val Arg Gly Pro Asn Val Glu Ala Gly Cys Leu Asn  
 545 550 555 560

Pro His Ile Val Leu Asn Ile Asp Leu Ala Pro Thr Ile Leu Asp Ile  
 565 570 575

ES 2 627 117 T3

Ala Gly Leu Asp Ile Pro Ala Asp Met Asp Gly Lys Ser Ile Leu Lys  
580 585 590

Leu Leu Asp Thr Glu Arg Pro Val Asn Arg Phe His Leu Lys Lys Lys  
595 600 605

Met Arg Val Trp Arg Asp Ser Phe Leu Val Glu Arg Gly Lys Leu Leu  
610 615 620

His Lys Arg Asp Asn Asp Lys Val Asp Ala Gln Glu Glu Asn Phe Leu  
625 630 635 640

Pro Lys Tyr Gln Arg Val Lys Asp Leu Cys Gln Arg Ala Glu Tyr Gln  
645 650 655

Thr Ala Cys Glu Gln Leu Gly Gln Lys Trp Gln Cys Val Glu Asp Ala  
660 665 670

Thr Gly Lys Leu Lys Leu His Lys Cys Lys Gly Pro Met Arg Leu Gly  
675 680 685

Gly Ser Arg Ala Leu Ser Asn Leu Val Pro Lys Tyr Tyr Gly Gln Gly  
690 695 700

Ser Glu Ala Cys Thr Cys Asp Ser Gly Asp Tyr Lys Leu Ser Leu Ala  
705 710 715 720

Gly Arg Arg Lys Lys Leu Phe Lys Lys Lys Tyr Lys Ala Ser Tyr Val  
725 730 735

Arg Ser Arg Ser Ile Arg Ser Val Ala Ile Glu Val Asp Gly Arg Val  
740 745 750

Tyr His Val Gly Leu Gly Asp Ala Ala Gln Pro Arg Asn Leu Thr Lys  
755 760 765

Arg His Trp Pro Gly Ala Pro Glu Asp Gln Asp Asp Lys Asp Gly Gly  
770 775 780

Asp Phe Ser Gly Thr Gly Gly Leu Pro Asp Tyr Ser Ala Ala Asn Pro  
785 790 795 800

Ile Lys Val Thr His Arg Cys Tyr Ile Leu Glu Asn Asp Thr Val Gln  
805 810 815

Cys Asp Leu Asp Leu Tyr Lys Ser Leu Gln Ala Trp Lys Asp His Lys

ES 2 627 117 T3

820	825	830																			
Leu	His	Ile	Asp	His	Glu	Ile	Glu	Thr	Leu	Gln	Asn	Lys	Ile	Lys	Asn						
		835					840					845									
Leu	Arg	Glu	Val	Arg	Gly	His	Leu	Lys	Lys	Lys	Arg	Pro	Glu	Glu	Cys						
	850					855					860										
Asp	Cys	His	Lys	Ile	Ser	Tyr	His	Thr	Gln	His	Lys	Gly	Arg	Leu	Lys						
865					870					875					880						
His	Arg	Gly	Ser	Ser	Leu	His	Pro	Phe	Arg	Lys	Gly	Leu	Gln	Glu	Lys						
				885					890						895						
Asp	Lys	Val	Trp	Leu	Leu	Arg	Glu	Gln	Lys	Arg	Lys	Lys	Lys	Leu	Arg						
			900					905						910							
Lys	Leu	Leu	Lys	Arg	Leu	Gln	Asn	Asn	Asp	Thr	Cys	Ser	Met	Pro	Gly						
		915					920						925								
Leu	Thr	Cys	Phe	Thr	His	Asp	Asn	Gln	His	Trp	Gln	Thr	Ala	Pro	Phe						
	930					935					940										
Trp	Thr	Leu	Gly	Pro	Phe	Cys	Ala	Cys	Thr	Ser	Ala	Asn	Asn	Asn	Thr						
945					950					955					960						
Tyr	Trp	Cys	Met	Arg	Thr	Ile	Asn	Glu	Thr	His	Asn	Phe	Leu	Phe	Cys						
				965					970						975						
Glu	Phe	Ala	Thr	Gly	Phe	Leu	Glu	Tyr	Phe	Asp	Leu	Asn	Thr	Asp	Pro						
			980					985						990							
Tyr	Gln	Leu	Met	Asn	Ala	Val	Asn	Thr	Leu	Asp	Arg	Asp	Val	Leu	Asn						
		995					1000						1005								
Gln	Leu	His	Val	Gln	Leu	Met	Glu	Leu	Arg	Ser	Cys	Lys	Gly	Tyr							
	1010					1015					1020										
Lys	Gln	Cys	Asn	Pro	Arg	Thr	Arg	Asn	Met	Asp	Leu	Gly	Leu	Lys							
	1025					1030					1035										
Asp	Gly	Gly	Ser	Tyr	Glu	Gln	Tyr	Arg	Gln	Phe	Gln	Arg	Arg	Lys							
	1040					1045					1050										
Trp	Pro	Glu	Met	Lys	Arg	Pro	Ser	Ser	Lys	Ser	Leu	Gly	Gln	Leu							
	1055					1060					1065										
Trp	Glu	Gly	Trp	Glu	Gly	Pro	Gly	Glu	Asn	Leu	Tyr	Phe	Gln	Gly							
	1070					1075					1080										
Pro	Gly	Gly	Gly	Ser	His	His	His	His	His	His											
	1085					1090															

## REIVINDICACIONES

1. Utilización *in vitro* de:
- 5 i) el enzima 6-metilguanina-ADN-metiltransferasa (MGMT, EC 2.1.1.63) o un homólogo del mismo, cuya secuencia presenta una identidad de por lo menos 70% con SEC ID nº: 4, o
- 10 ii) el mutante SNAP de SEC ID nº: 2 o un homólogo del mismo, cuya secuencia presenta una identidad de por lo menos 80% con SEC ID nº: 2
- para aumentar la producción de una proteína heteróloga en células de insecto infectadas con vectores replicativos o defectuosos.
2. Utilización según la reivindicación 1, en la que dicha proteína heteróloga es secretada por dichas células de insecto.
3. Utilización según la reivindicación 1 o 2, en la que dicha proteína heteróloga es codificada por un vector que comprende una secuencia de nucleótidos que codifica en un marco de lectura abierto único, de 5' a 3':
- 20 a) una señal de secreción peptídica que es funcional en dichas células de insecto,
- b) dichos enzima 6-metilguanina-ADN-metiltransferasa (MGMT, EC 2.1.1.63), mutante SNAP u homólogo de los mismos, y
- 25 c) dicha proteína heteróloga.
4. Utilización según cualquiera de las reivindicaciones 1 a 3, en la que dicho enzima MGMT es la proteína de SEC ID nº: 4 o la proteína SNAP de SEC ID nº: 2.
- 30 5. Utilización según cualquiera de las reivindicaciones 1 a 4, en la que dicha célula de insecto es una célula S2 de *Drosophila* o un derivado de la misma.
6. Utilización según cualquiera de las reivindicaciones 1 a 5, en la que dicha proteína heteróloga se selecciona de entre el grupo que consiste en proteínas y/o polipéptidos diagnósticos y/o terapéuticos, preferentemente seleccionados de entre el grupo que consiste en IFN $\alpha$ , granzima M, FasL, SSX2, NERCMSL, hSULF2<sup>ΔTMD</sup> y CNTN4.
- 35 7. Vector para expresar proteínas recombinantes en células de insecto, que comprende una secuencia de nucleótidos que codifica en un marco de lectura abierto único, de 5' a 3':
- 40 a) una señal de secreción peptídica que es funcional en dichas células de insecto,
- b) un enzima 6-metilguanina-ADN-metiltransferasa (MGMT, EC 2.1.1.63) o un homólogo del mismo, cuya secuencia presenta una identidad de por lo menos 70% con SEC ID nº: 4, o el mutante SNAP de SEC ID nº: 2 o un homólogo del mismo, cuya secuencia presenta una identidad de por lo menos 80% con SEC ID nº: 2, y
- 45 c) una proteína recombinante.
- 50 8. Vector de expresión según la reivindicación 7, en el que dicho marco de lectura abierto está asociado funcionalmente con un promotor inducible que es funcional en la misma célula de insecto que la señal peptídica, siendo preferentemente dicha célula de insecto una célula S2 de *Drosophila*.
9. Vector de expresión según la reivindicación 7 u 8, en el que dicha proteína recombinante se selecciona de entre el grupo que consiste en proteínas y/o polipéptidos diagnósticos y/o terapéuticos, preferentemente seleccionados de entre el grupo que consiste en IFN $\alpha$ , granzima M, FasL, SSX2, NERCMSL, hSULF2<sup>ΔTMD</sup> y CNTN4.
- 55 10. Célula recombinante de insecto que es transfectada de manera estable mediante el vector de expresión según cualquiera de las reivindicaciones 7 a 9.
- 60 11. Célula recombinante según la reivindicación 10, en la que es una célula S2 de *Drosophila*.
12. Célula transfectada de manera estable según la reivindicación 11, caracterizada por que
- 65 i) dicho vector de expresión se selecciona de entre el grupo que consiste en:

- 5 - un vector que comprende la SEC ID nº 19 o la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, París, Francia, el 19 de agosto de 2010 con el número CNCM I-4357,
- el vector según la reivindicación 7 que comprende la SEC ID nº 22 o la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 27 de octubre de 2010 con el CNCM I-4381,
- 10 - el vector según la reivindicación 7 que comprende la SEC ID nº 21 o la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 27 de octubre de 2010 con el número CNCM I-4382,
- 15 - un vector que comprende la SEC ID nº 9 o la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 29 de septiembre de 2010 con el número CNCM I-4368, y
- el vector según la reivindicación 7 que comprende la SEC ID nº 20 la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 29 de septiembre de 2010 con el número CNCM I-4369,
- 20 - el vector según la reivindicación 17 que comprende la SEC ID nº 10 o 59 o 69,
- el vector de SEC ID nº 64 o la secuencia de nucleótidos clonada en la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 9 de diciembre de 2011 con el número CNCM I-4581,
- 25 - el vector de la SEC ID nº 71,
- 30 - el vector según la reivindicación 9 que comprende la SEC ID nº 55, SEC ID nº 57 o 72 o 74, SEC ID nº 77, 79 u 81, SEC ID nº 89, SEC ID nº 84 u 86, SEC ID nº 92 o SEC ID nº 96, o
- ii) se selecciona de entre el grupo que consiste en:
  - 35 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 19 de agosto de 2010 con el número CNCM I-4357,
  - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 27 de octubre de 2010 con el CNCM I-4381,
  - 40 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 27 de octubre de 2010 con el número CNCM I-4382,
  - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 29 de septiembre de 2010 con el número CNCM I-4368,
  - 45 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur, el 29 de septiembre de 2010 con el número CNCM I-4369,
  - 50 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 París Cedex 15, Francia) el 5 de diciembre de 2011, con el número CNCM I-4565,
  - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 París Cedex 15, Francia) el 5 de diciembre de 2011, con el número CNCM I-4566,
  - 55 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 París Cedex 15, Francia) el 5 de diciembre de 2011, con el número CNCM I-4567,
  - 60 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 París Cedex 15, Francia) el 5 de diciembre de 2011, con el número CNCM I-4568,
  - 65 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM),

Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 5 de diciembre de 2011, con el número CNCM I-4569,

- 5 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 5 de diciembre de 2011, con el número CNCM I-4570,
- 10 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 5 de diciembre de 2011, con el número CNCM I-4571,
- 15 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 5 de diciembre de 2011, con el número CNCM I-4572,
- la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 8 de diciembre de 2011, con el número CNCM I-4576,
- 20 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 8 de diciembre de 2011, con el número CNCM I-4577,
- 25 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 8 de diciembre de 2011, con el número CNCM I-4578,
- 30 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 8 de diciembre de 2011, con el número CNCM I-4579,
- 35 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 8 de diciembre de 2011, con el número CNCM I-4580,
- 40 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 9 de diciembre de 2011, con el número CNCM I-4584,
- 45 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 9 de diciembre de 2011, con el número CNCM I-4585, y
- 50 - la célula que ha sido depositada en el Centre National de Culture et de Microorganismes (CNCM), Institut Pasteur (25 rue du Docteur Roux, 75724 Paris Cedex 15, Francia) el 9 de diciembre de 2011, con el número CNCM I-4586.

13. Polipéptido de fusión que comprende una señal de secreción peptídica que es funcional en células de insecto y la proteína SNAP de SEC ID nº: 2, preferentemente comprendiendo asimismo una proteína recombinante.

55 14. Utilización *in vitro* del enzima 6-metilguanina-ADN-metiltransferasa (MGMT, EC 2.1.1.63) o un homólogo del mismo, cuya secuencia presenta una identidad de por lo menos 70% con SEC ID nº: 4 o el mutante SNAP de SEC ID nº: 2 o un homólogo del mismo, cuya secuencia presenta una identidad de por lo menos 80% con SEC ID nº: 2, como polipéptido protector fusionado o unido a proteínas recombinantes para mejorar la semivida de la proteína recombinante en un medio de almacenamiento, en plasma o en amortiguador, o para mejorar la semivida de la proteína recombinante utilizada en kits diagnósticos.

65 15. Polipéptido de fusión que comprende el enzima 6-metilguanina-ADN-metiltransferasa (MGMT, EC 2.1.1.63) o un homólogo del mismo cuya secuencia presenta una identidad de por lo menos 70% con SEC ID nº: 4, o el mutante SNAP de SEC ID nº: 2 o un homólogo del mismo cuya secuencia presenta una identidad de por lo menos 80% con SEC ID nº: 2, y una proteína recombinante, para la utilización como un medicamento o una vacuna.

Figura 1A

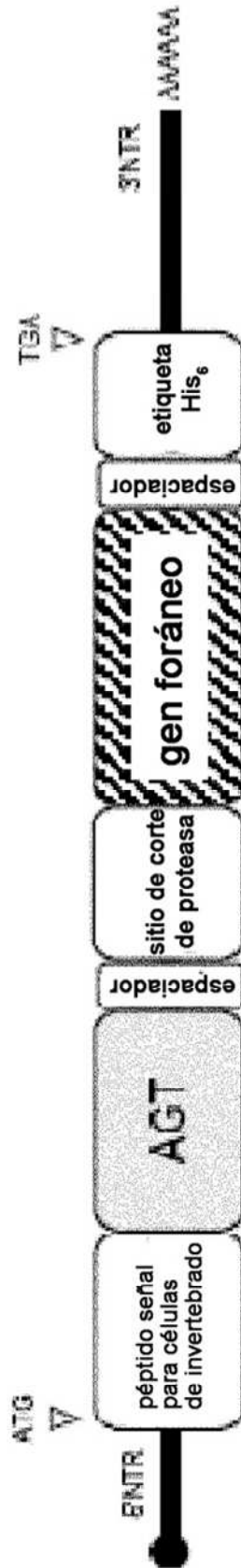






Figure 2A

```

10 100 110 120 130 140 150 160 170 180 190
20 200 210 220 230 240 250 260 270 280 290
30 300 310 320 330 340 350 360 370 380 390
40 400 410 420 430 440 450 460 470 480
50 500 510 520 530 540 550 560 570 580 590
60 600 610 620 630 640 650 660 670 680 690
70 700 710 720 730 740 750 760 770 780
80 800 810 820 830 840 850 860 870 880 890
90 900 910 920 930 940 950 960 970 980 990
100 1000 1010 1020 1030 1040 1050 1060 1070 1080 1090
1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
1200 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
1300 1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
1400 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
1500 1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
1600 1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
1800 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900
1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
2000 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100
2100 2110 2120 2130 2140 2150 2160 2170 2180 2190 2200
2200 2210 2220 2230 2240 2250 2260 2270 2280 2290 2300
2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400
2400 2410 2420 2430 2440 2450 2460 2470 2480 2490 2500
2500 2510 2520 2530 2540 2550 2560 2570 2580 2590 2600
2600 2610 2620 2630 2640 2650 2660 2670 2680 2690 2700
2700 2710 2720 2730 2740 2750 2760 2770 2780 2790 2800
2800 2810 2820 2830 2840 2850 2860 2870 2880 2890 2900
2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000
3000 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100
3100 3110 3120 3130 3140 3150 3160 3170 3180 3190 3200
3200 3210 3220 3230 3240 3250 3260 3270 3280 3290 3300
3300 3310 3320 3330 3340 3350 3360 3370 3380 3390 3400
3400 3410 3420 3430 3440 3450 3460 3470 3480 3490 3500
3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
3600 3610 3620 3630 3640 3650 3660 3670 3680 3690 3700
3700 3710 3720 3730 3740 3750 3760 3770 3780 3790 3800
3800 3810 3820 3830 3840 3850 3860 3870 3880 3890 3900
3900 3910 3920 3930 3940 3950 3960 3970 3980 3990 4000
4000 4010 4020 4030 4040 4050 4060 4070 4080 4090 4100
4100 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
4200 4210 4220 4230 4240 4250 4260 4270 4280 4290 4300
4300 4310 4320 4330 4340 4350 4360 4370 4380 4390 4400
4400 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500
4500 4510 4520 4530 4540 4550 4560 4570 4580 4590 4600
4600 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700
4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
4800 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900
4900 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000
5000 5010 5020 5030 5040 5050 5060 5070 5080 5090 5100
5100 5110 5120 5130 5140 5150 5160 5170 5180 5190 5200
5200 5210 5220 5230 5240 5250 5260 5270 5280 5290 5300
5300 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400
5400 5410 5420 5430 5440 5450 5460 5470 5480 5490 5500
5500 5510 5520 5530 5540 5550 5560 5570 5580 5590 5600
5600 5610 5620 5630 5640 5650 5660 5670 5680 5690 5700
5700 5710 5720 5730 5740 5750 5760 5770 5780 5790 5800
5800 5810 5820 5830 5840 5850 5860 5870 5880 5890 5900
5900 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000
6000 6010 6020 6030 6040 6050 6060 6070 6080 6090 6100
6100 6110 6120 6130 6140 6150 6160 6170 6180 6190 6200
6200 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300
6300 6310 6320 6330 6340 6350 6360 6370 6380 6390 6400
6400 6410 6420 6430 6440 6450 6460 6470 6480 6490 6500
6500 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600
6600 6610 6620 6630 6640 6650 6660 6670 6680 6690 6700
6700 6710 6720 6730 6740 6750 6760 6770 6780 6790 6800
6800 6810 6820 6830 6840 6850 6860 6870 6880 6890 6900
6900 6910 6920 6930 6940 6950 6960 6970 6980 6990 7000
7000 7010 7020 7030 7040 7050 7060 7070 7080 7090 7100
7100 7110 7120 7130 7140 7150 7160 7170 7180 7190 7200
7200 7210 7220 7230 7240 7250 7260 7270 7280 7290 7300
7300 7310 7320 7330 7340 7350 7360 7370 7380 7390 7400
7400 7410 7420 7430 7440 7450 7460 7470 7480 7490 7500
7500 7510 7520 7530 7540 7550 7560 7570 7580 7590 7600
7600 7610 7620 7630 7640 7650 7660 7670 7680 7690 7700
7700 7710 7720 7730 7740 7750 7760 7770 7780 7790 7800
7800 7810 7820 7830 7840 7850 7860 7870 7880 7890 7900
7900 7910 7920 7930 7940 7950 7960 7970 7980 7990 8000
8000 8010 8020 8030 8040 8050 8060 8070 8080 8090 8100
8100 8110 8120 8130 8140 8150 8160 8170 8180 8190 8200
8200 8210 8220 8230 8240 8250 8260 8270 8280 8290 8300
8300 8310 8320 8330 8340 8350 8360 8370 8380 8390 8400
8400 8410 8420 8430 8440 8450 8460 8470 8480 8490 8500
8500 8510 8520 8530 8540 8550 8560 8570 8580 8590 8600
8600 8610 8620 8630 8640 8650 8660 8670 8680 8690 8700
8700 8710 8720 8730 8740 8750 8760 8770 8780 8790 8800
8800 8810 8820 8830 8840 8850 8860 8870 8880 8890 8900
8900 8910 8920 8930 8940 8950 8960 8970 8980 8990 9000
9000 9010 9020 9030 9040 9050 9060 9070 9080 9090 9100
9100 9110 9120 9130 9140 9150 9160 9170 9180 9190 9200
9200 9210 9220 9230 9240 9250 9260 9270 9280 9290 9300
9300 9310 9320 9330 9340 9350 9360 9370 9380 9390 9400
9400 9410 9420 9430 9440 9450 9460 9470 9480 9490 9500
9500 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600
9600 9610 9620 9630 9640 9650 9660 9670 9680 9690 9700
9700 9710 9720 9730 9740 9750 9760 9770 9780 9790 9800
9800 9810 9820 9830 9840 9850 9860 9870 9880 9890 9900
9900 9910 9920 9930 9940 9950 9960 9970 9980 9990 10000

```



Figura 2C

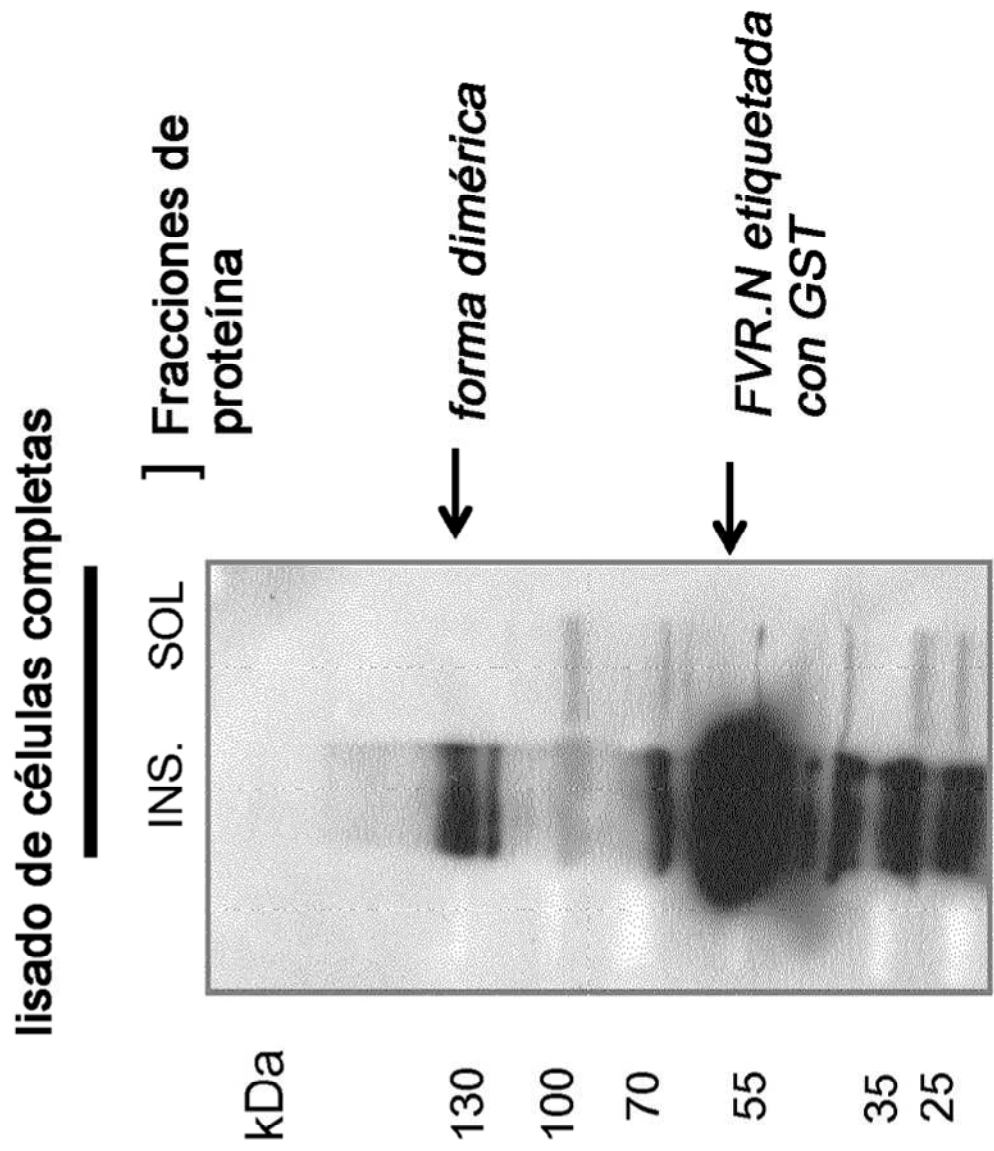


Figura 2D

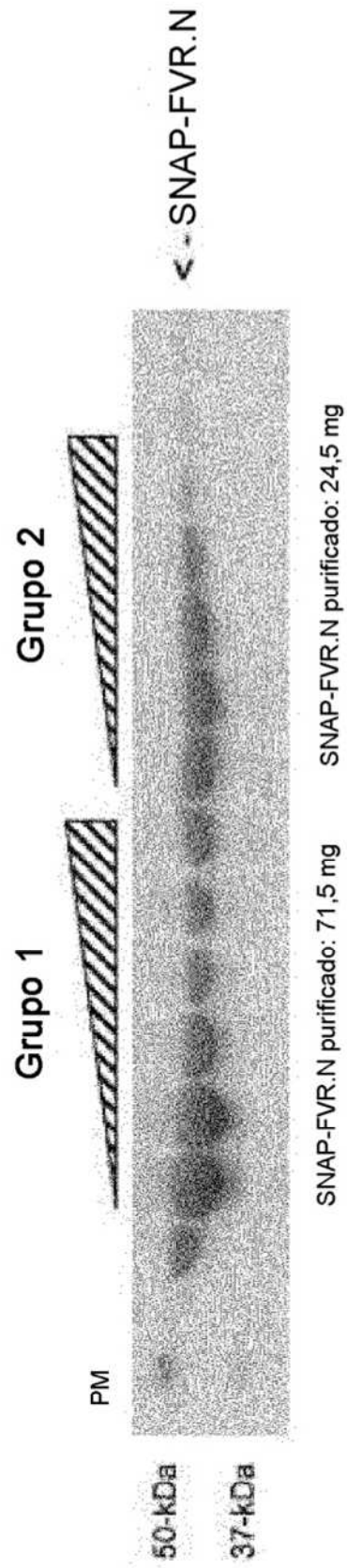






Figura 3B

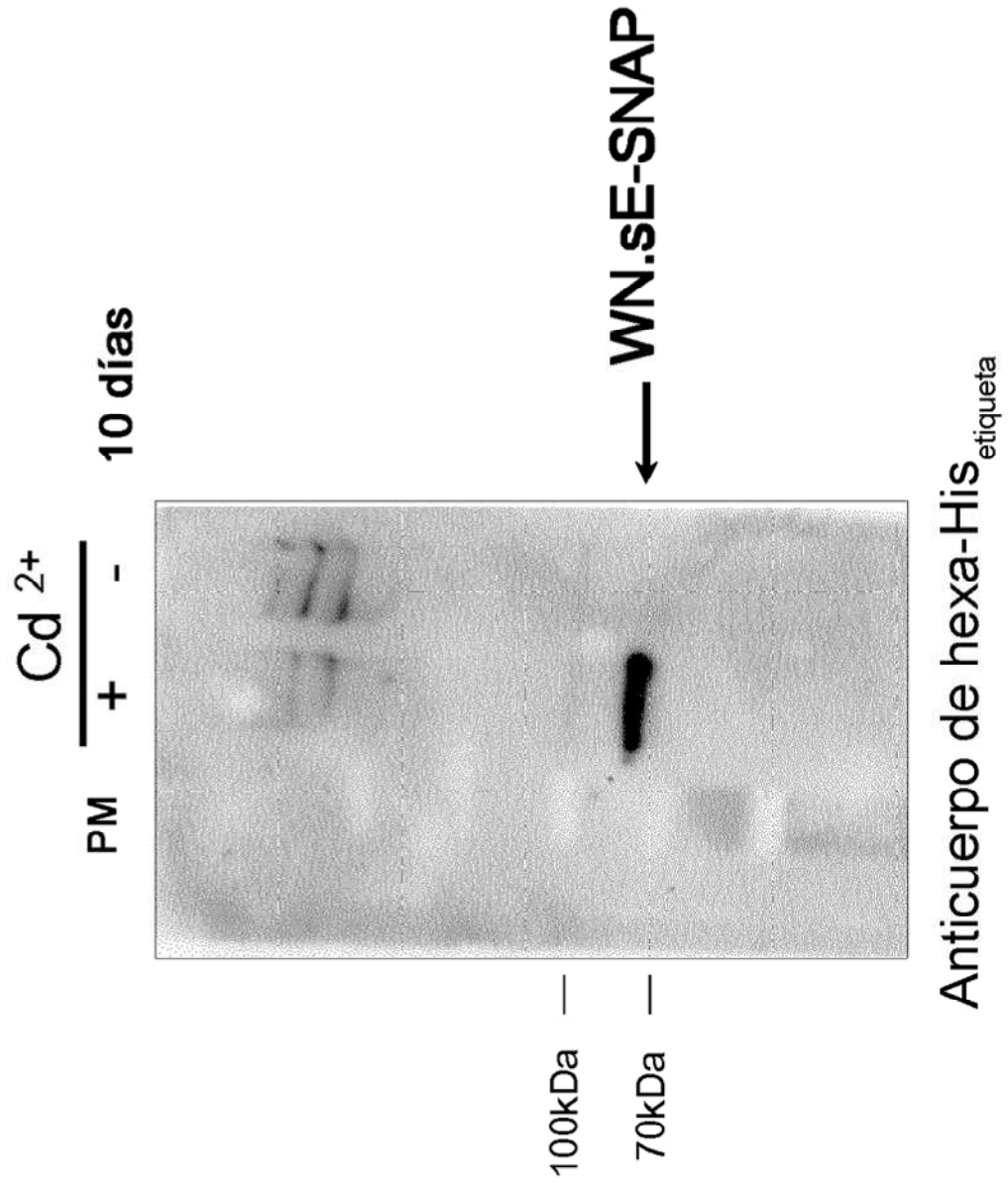




Figura 4A

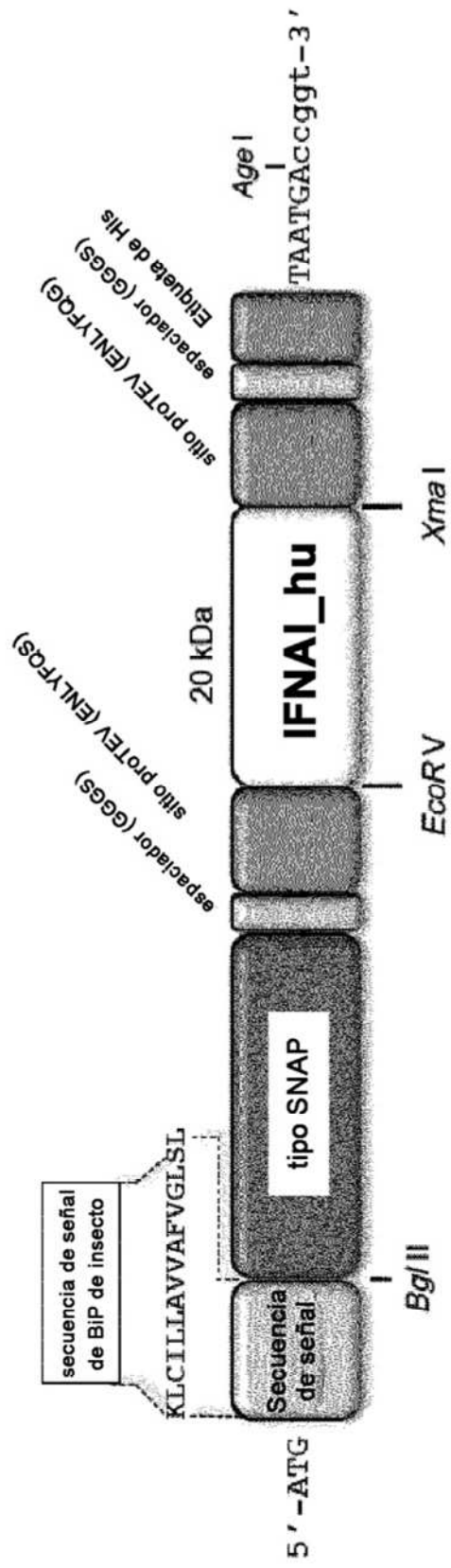




Figura 4B (continuación)

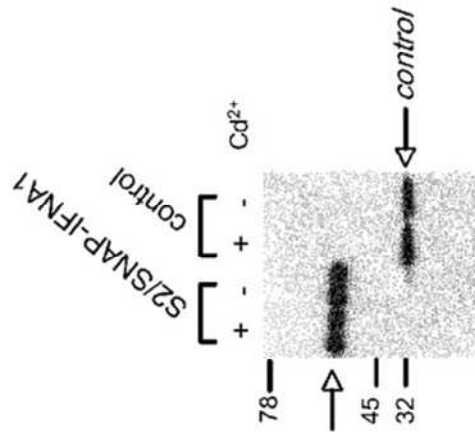
```

gag gac ctc cta gac aaa tcc tgc acc gaa ctc tac cag cag ctg aat gac ttg gaa gcc tgt gtc atg sag gag agg gtc ggc gaa acc ccc ctg a < 1000
E D L L D K F C T E L Y Q Q L N D L E A C V M Q E E R V G E T P L M
910 920 930 940 950 960 970 980 990
tg aat gcg gac tcc atc ttg gct gtc mag asa tac ttc cga aga atc acc ctc tat ctg aca gag sag aaa tac agc cct tgt gcc tgg gag gtt gtc ag < 1100
N A D S I L A V K K Y F R R I T L Y L T E K K Y S P C A W E V V R
1010 1020 1030 1040 1050 1060 1070 1080 1090
c gaa gaa ctc aag acc ccc ctg tcc ttc ttc cga aga tac ttc cga aga atc acc ctc tat ctg aca gag sag aaa tac agc cct tgt gcc tgg gag gtt gtc ag < 1200
A E I M R S L S L S T N L Q E R L R R K E G K W C C S S H H H H H
1110 1120 1130 1140 1150 1160 1170 1180 1190

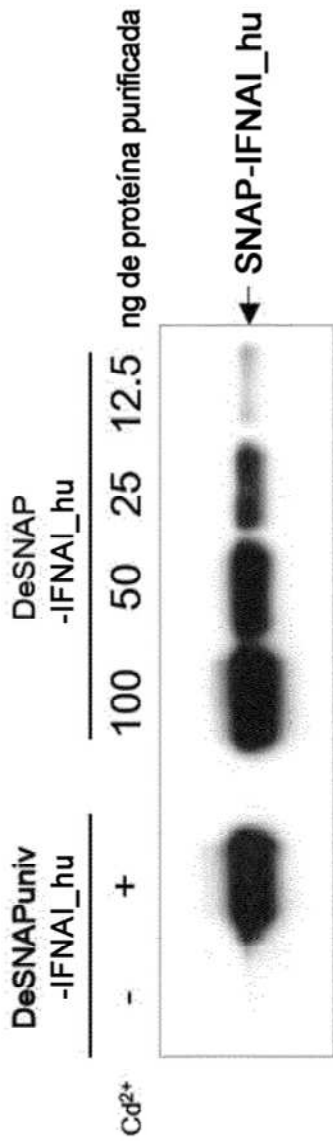
```

AgeI  
|  
cct taa ccc g. < 1211  
H ,

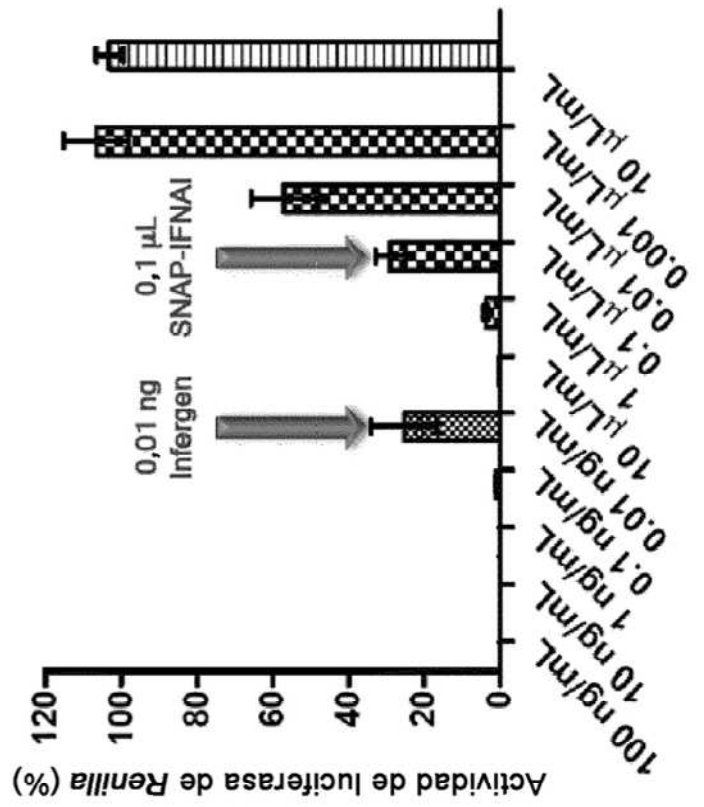
Figura 4C



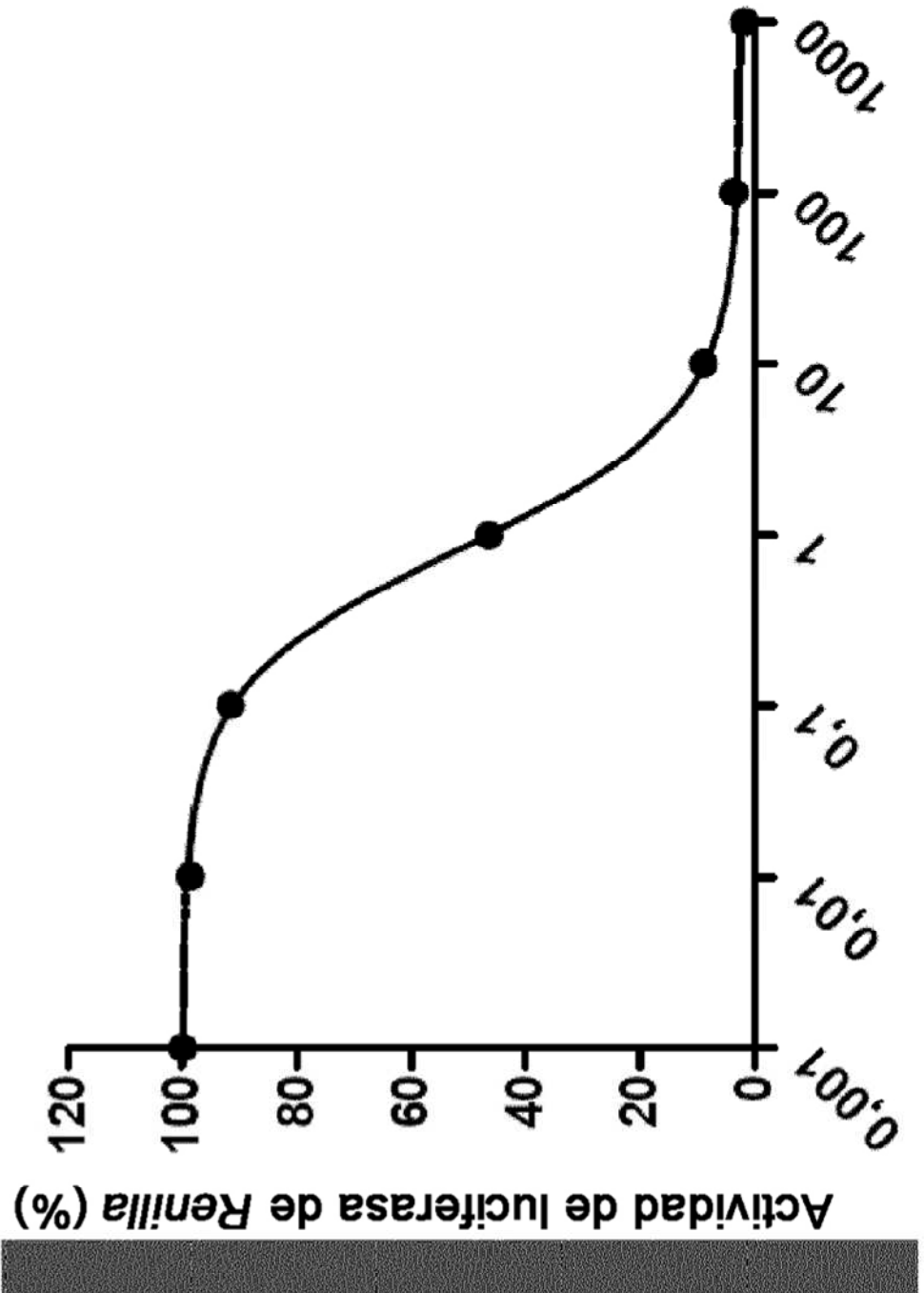
**Figura 4D**



**Figura 4E**



**Figura 4F**



**Figura 5**

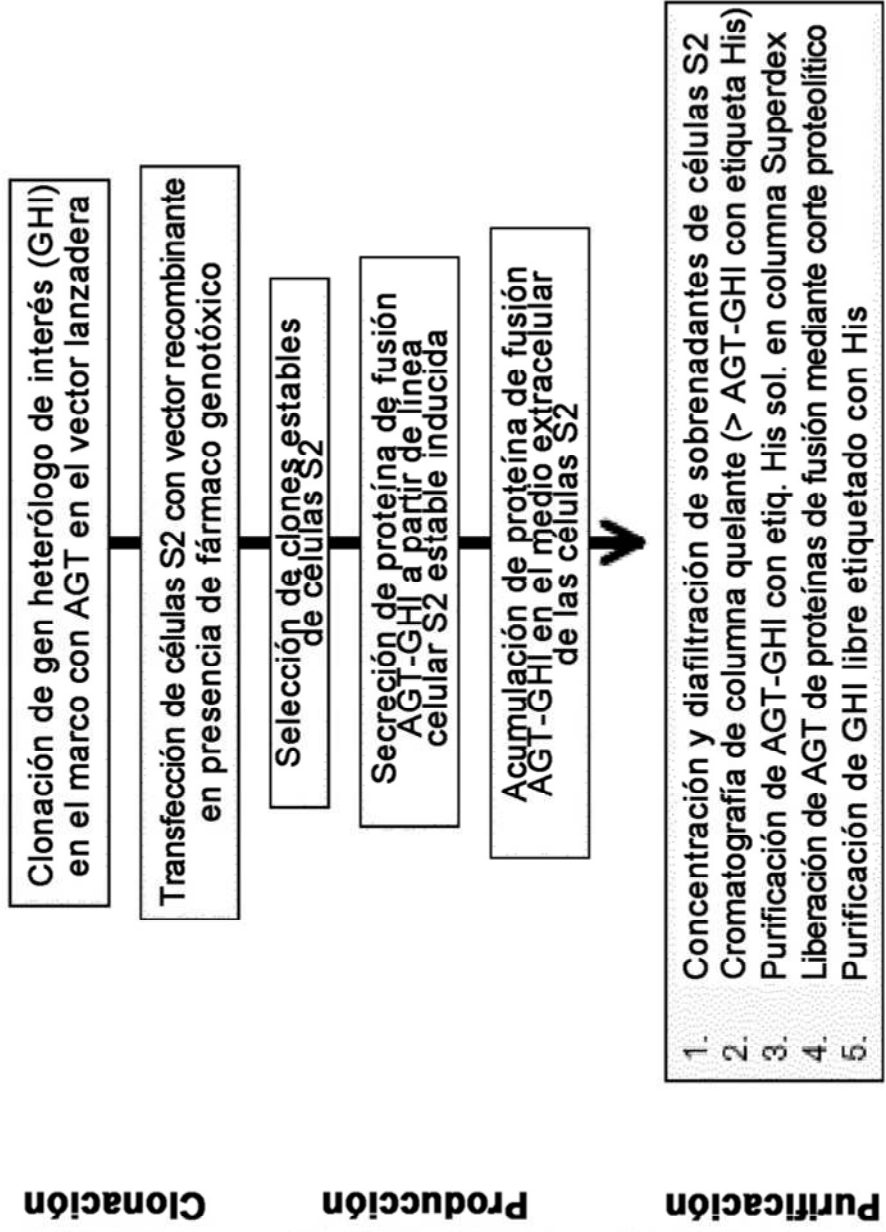


Figura 6 A

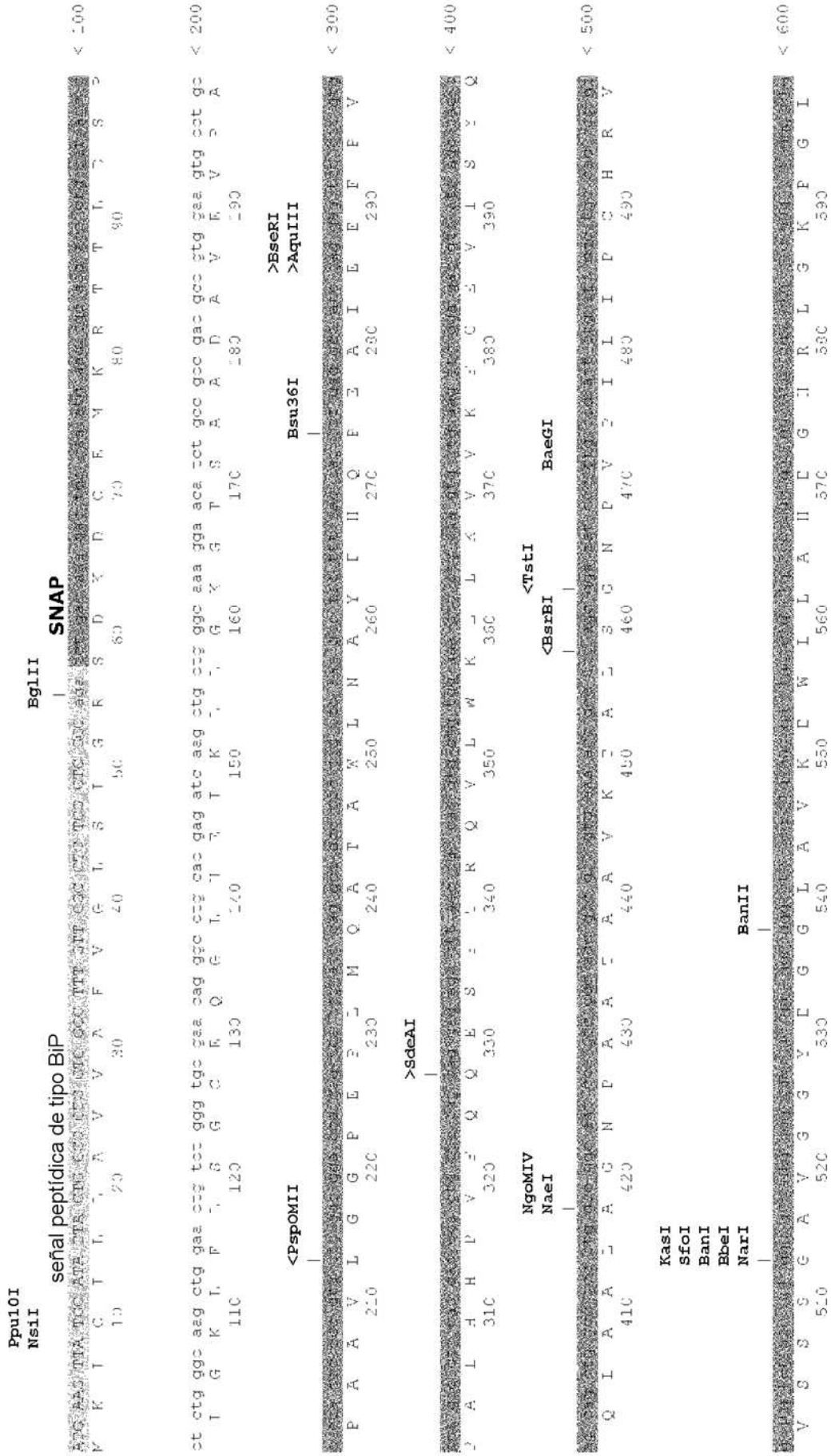


Figura 6A (continuación)

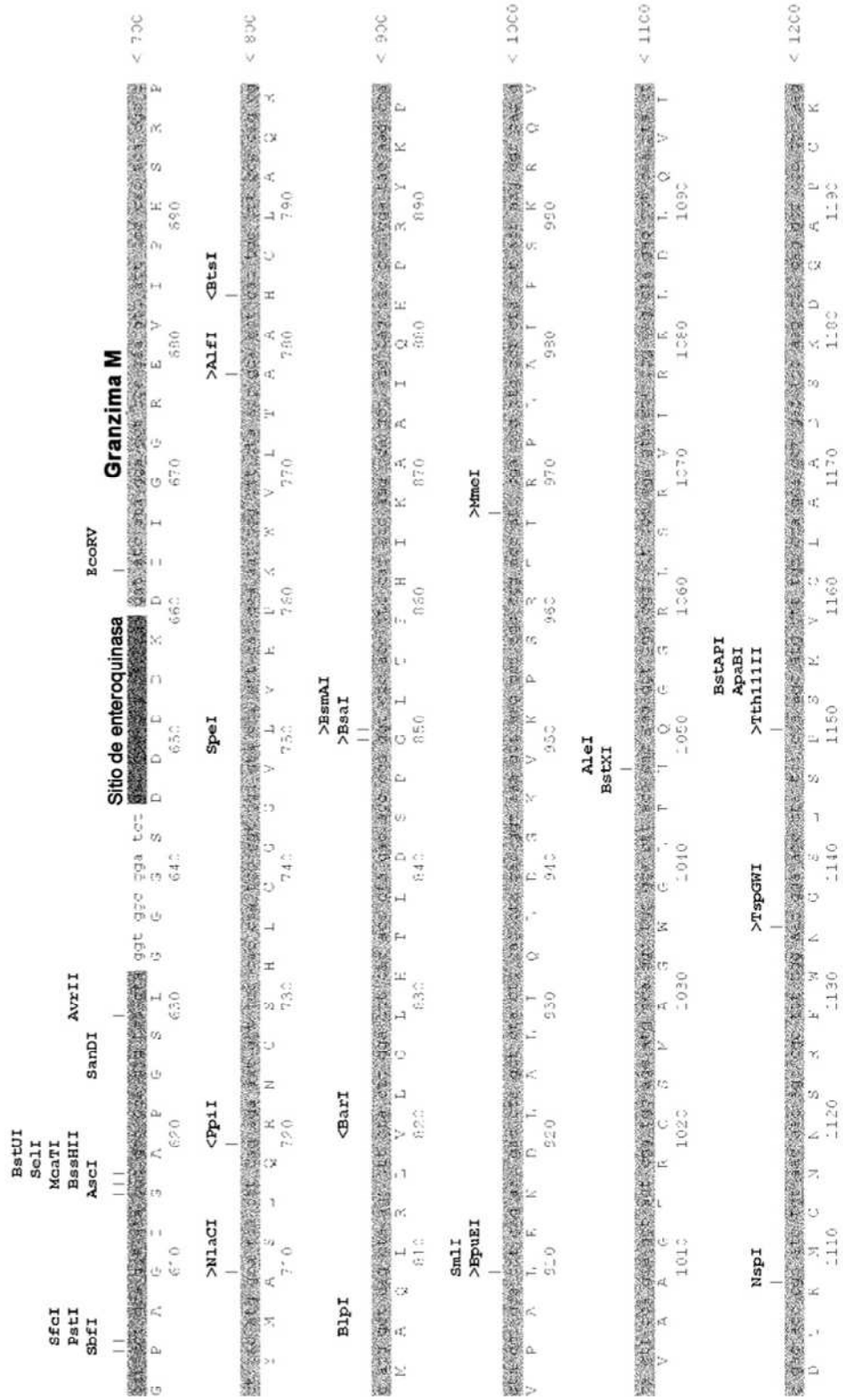
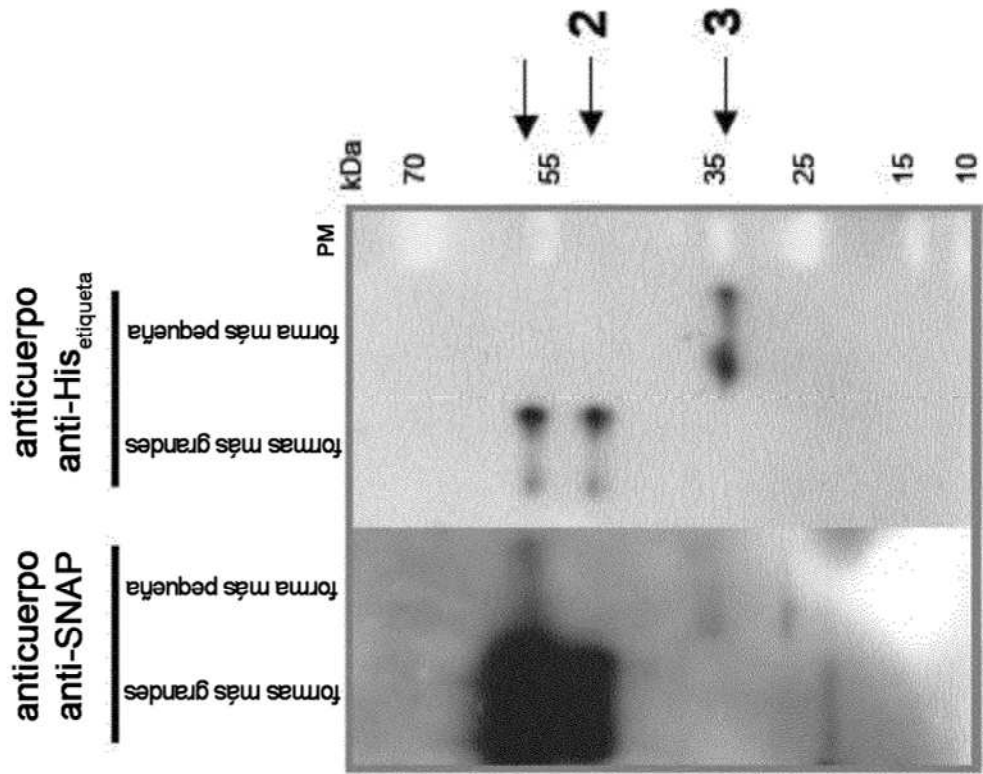






Figura 6C



Ensayos de inmunotransferencia de SNAP-GrM purificado (200 ng por ensayo)

Figura 7A

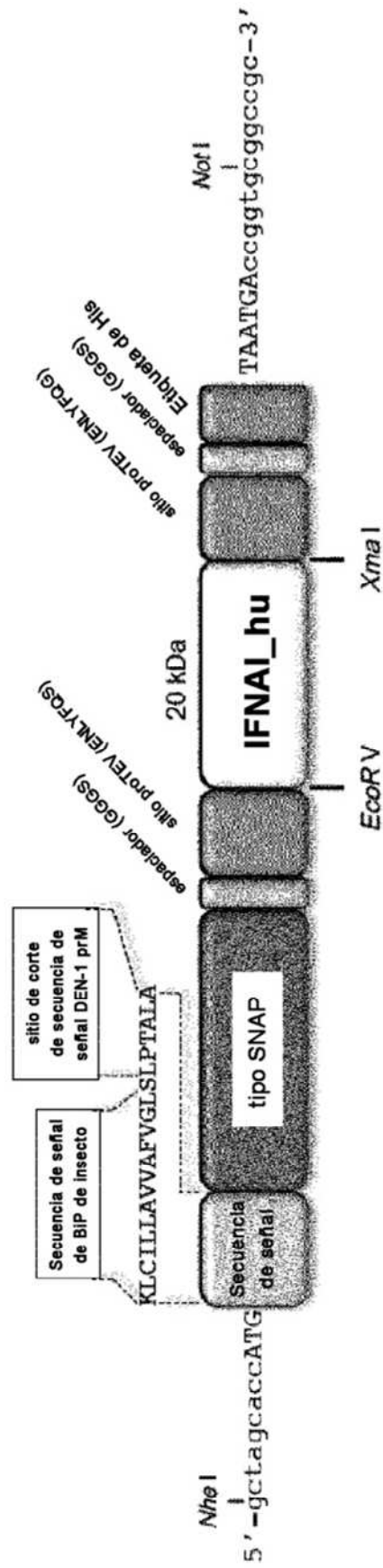




Figura 7B (continuación)

```

>RleAI
<BsmAI
<BsaI
|
gt gat ctc gag acc cac agc ctg gat acc agg agc acc ttg atg ctc ctg gca caa atg agc aga atc tct cct tcc tcc tgt ctg atg gac aga ca < 80C
D L P E T H S L D N R R T L M L L A Q M S R I S P S S C L M D R H
710 720 730 740 750 760 770 780 790

t gac ttc gga ttt ccc cag gag gac ttt gat ggc aac cag ttc cag agg gct cca ggc acc tct ctg ctc cca gag ctg atc cag cag att ttc aac ctg < 90C
D F G F P Q E E F D G N Q F Q K A P A I S V L H E L I Q Q I F N L
810 820 830 840 850 860 870 880 890

<Tth1111I
|
ttt acc aca aaa gat tca tct gct tgg gar gag gac ctc cta gac aaa ttc tgc acc gaa ctc tac cag cag ctg aat gac ttg gaa gcc tgc gty a < 100C
F T T K D S S A A W D E D L L D K F C T E L Y Q Q L N D L E A C V M
910 920 930 940 950 960 970 980 990

<BsgI
|
tg caq caq caq aq ctg aqa qua act ccc ctg atg caa gca qac tcc atc ttg gct qct atq acc tac ttc acc agc atc cct ctc tat ctg acc gaq ca < 110C
Q E R V G E T P L M N A D S I L A V K K Y F R R I T L Y L T E K
1010 1020 1030 1040 1050 1060 1070 1080 1090

>BsmI
|
tg caq caq caq aq ctg aqa qua act ccc ctg atg caa gca qac tcc atc ttg gct qct atq acc tac ttc acc agc atc cct ctc tat ctg acc gaq ca < 110C
Q E R V G E T P L M N A D S I L A V K K Y F R R I T L Y L T E K
1010 1020 1030 1040 1050 1060 1070 1080 1090

XmaI
AvaI
SmaI
N113877I
|
g asa tac agc cct tct gcc tgg gag gtt gtc aga gca gaa atc atg aga tcc ctc tct tta tca aca aac ttg caa gaa aga tta agy agy agg gaa gaa < 120C
K Y S P C A W E V R A E I M R S L S L S T N L Q E R L R R K E S
1110 1120 1130 1140 1150 1160 1170 1180 1190

PspOMI
ApaI
|
ccc gga gag aat cta taa ttt caa ggg ccc gcc gca ggt agt cgc cal ca' cgc gat cac taa tga ccc gtc ggc ccc cca gct : < 1265
P G E A L Y F Q G P G C S E H A H H E *
1210 1220 1230 1240 1250 1260

```

Figura 7C

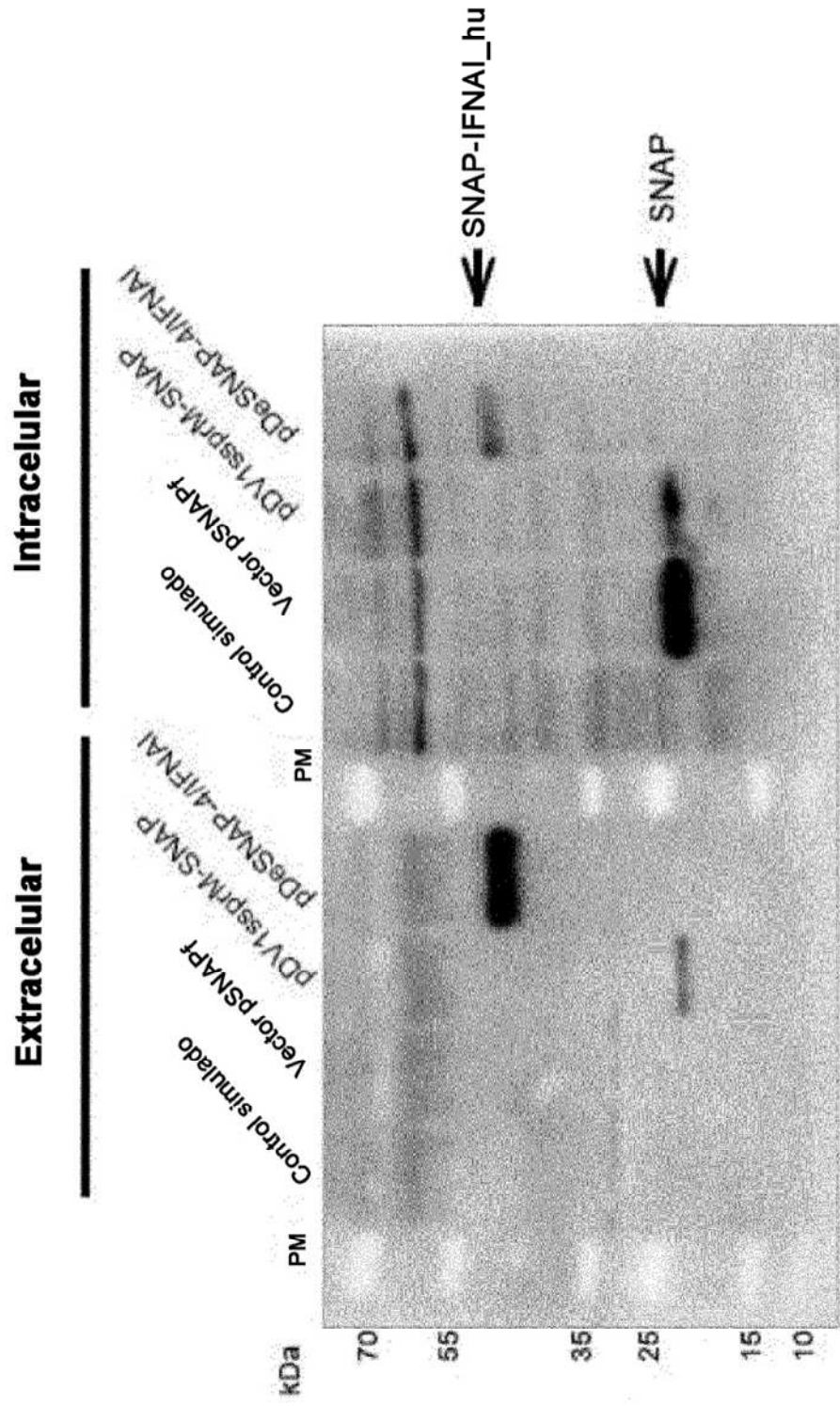




Figura 8B

### Vector lanzadera: DeMGMT univ

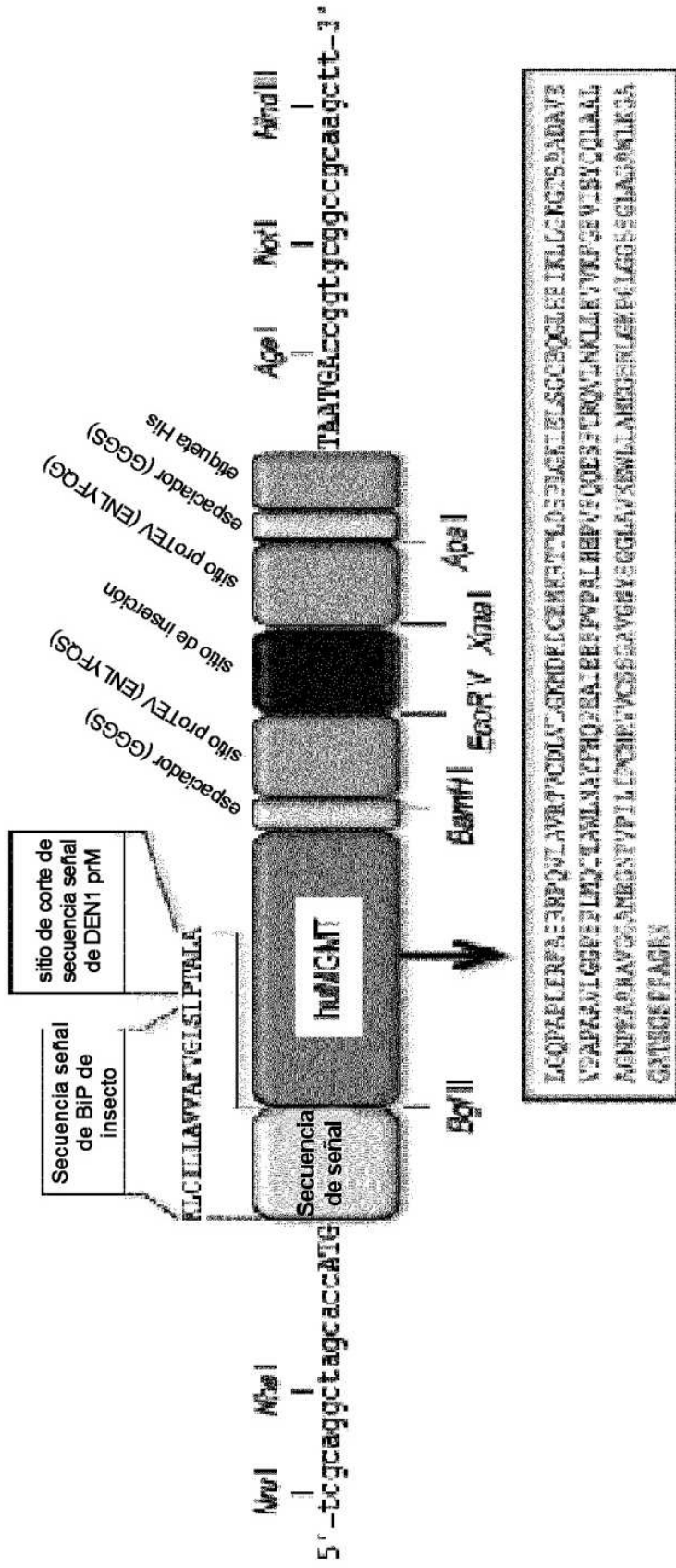




Figura 9

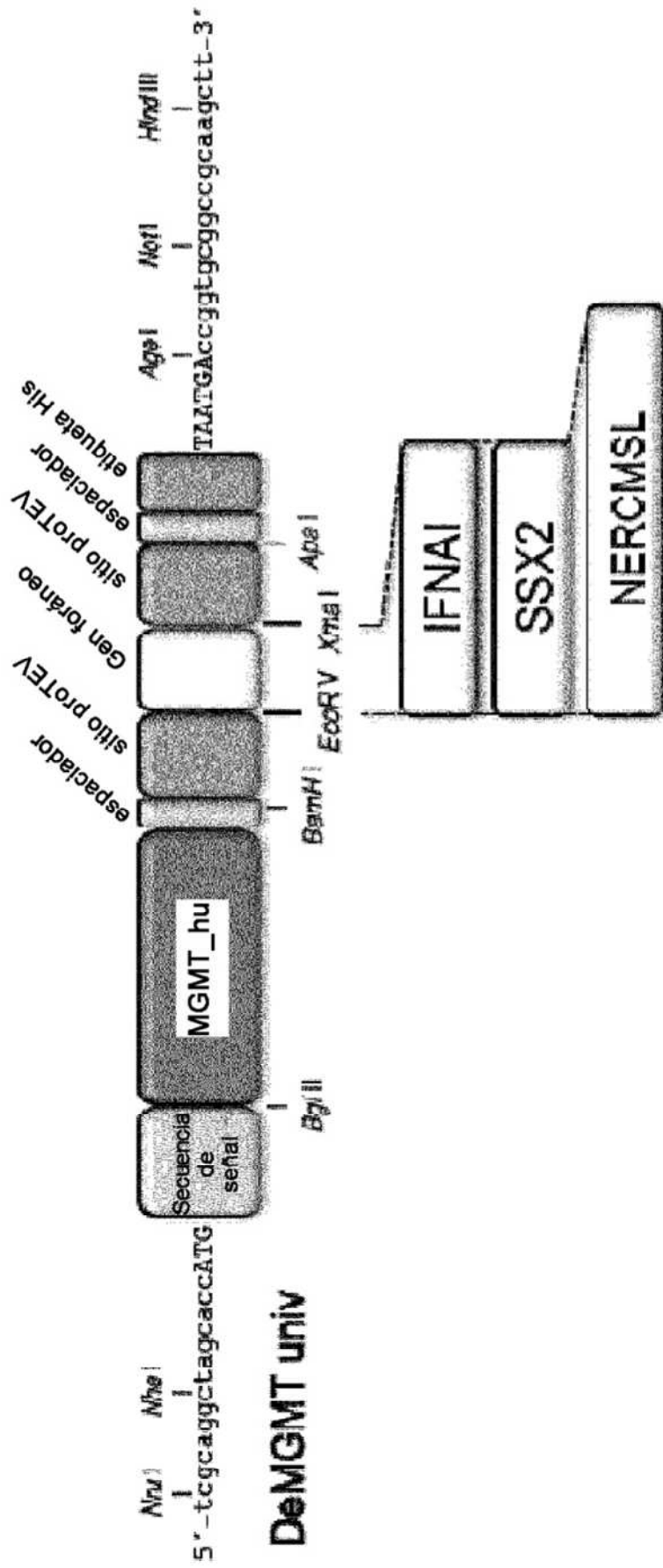


Figura 10A

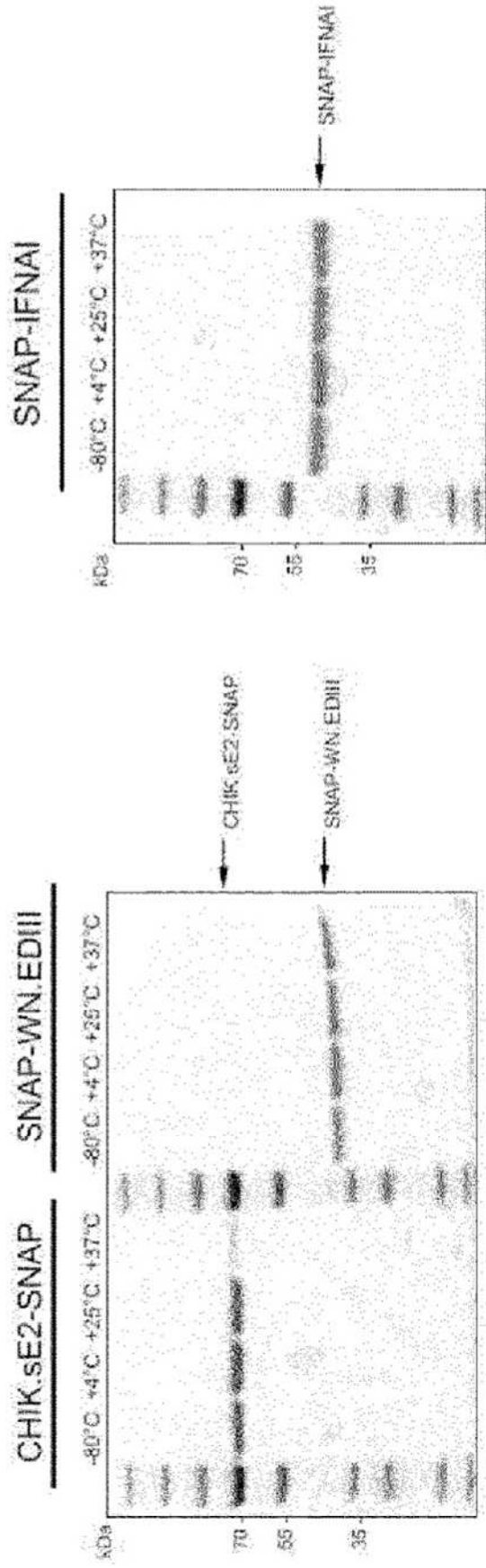


Figura 10B

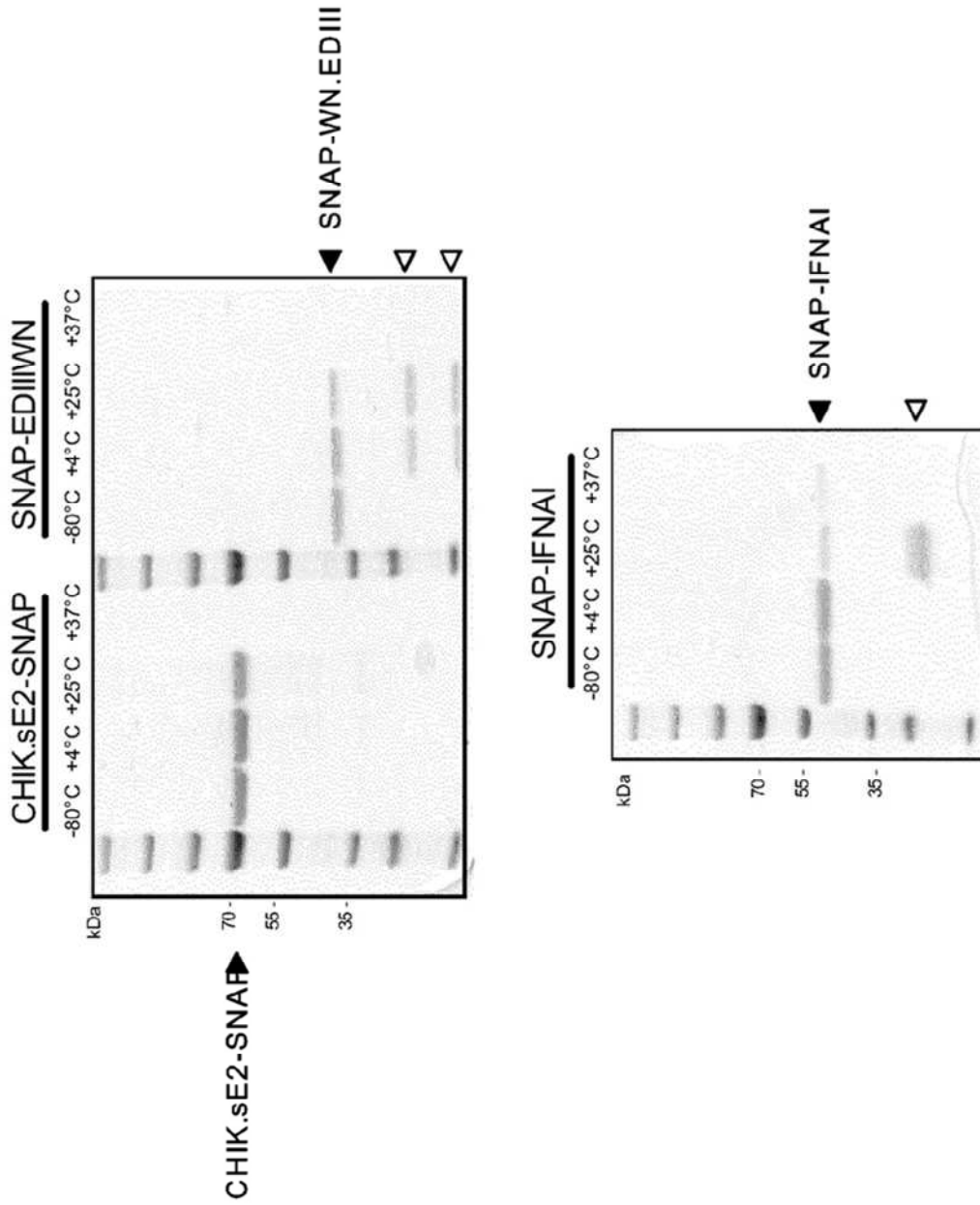
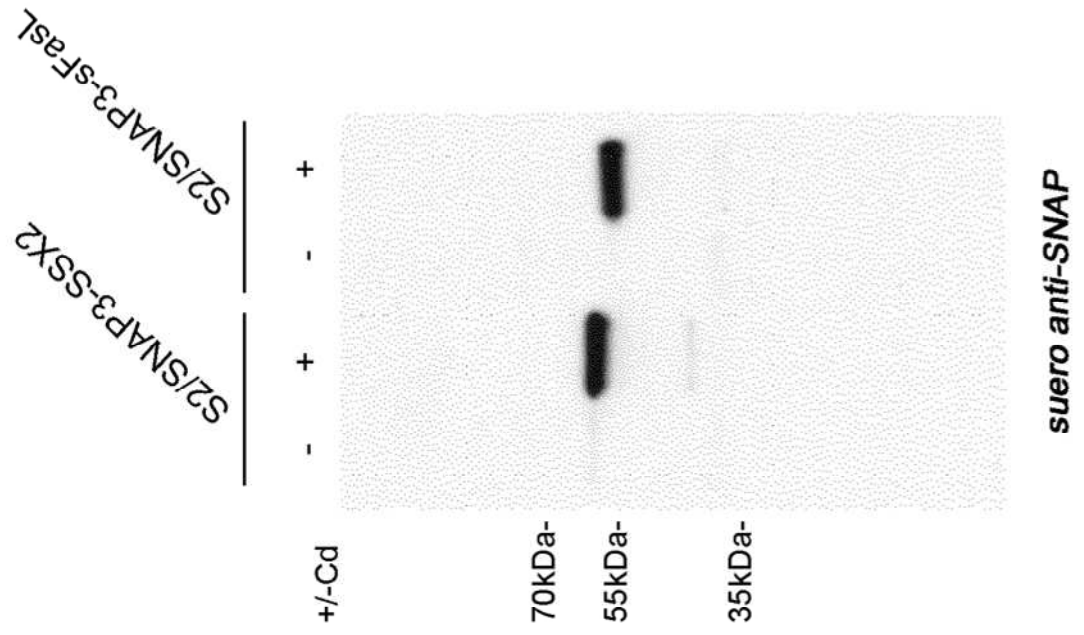


Figura 11A



**Figura 11B**

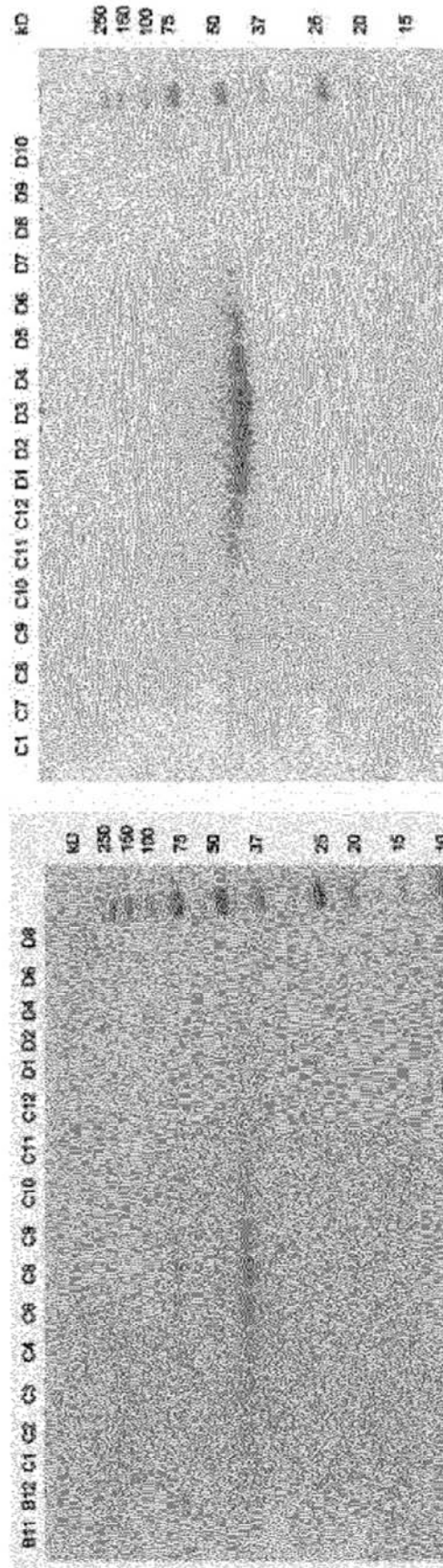


Figura 12A

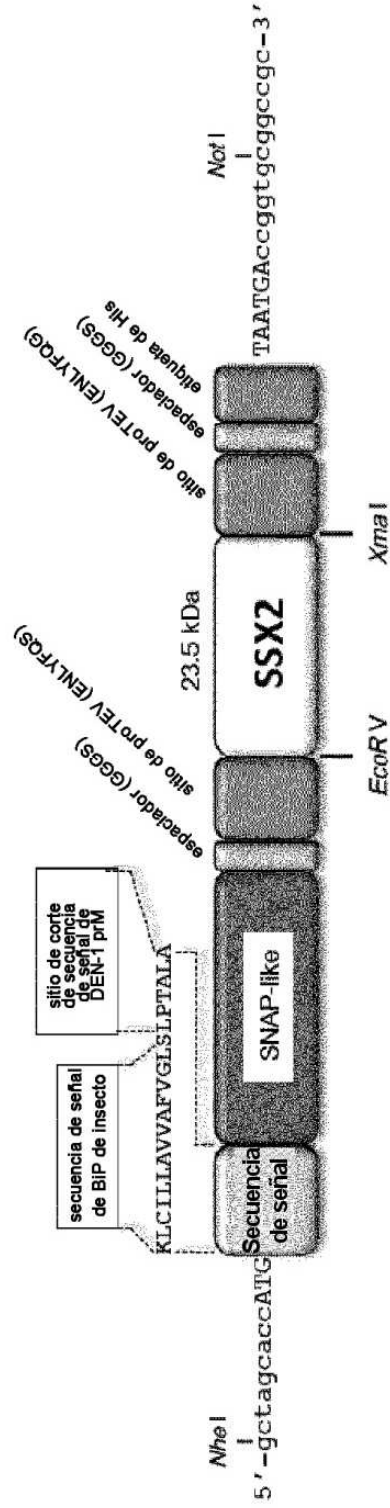


Figura 12B

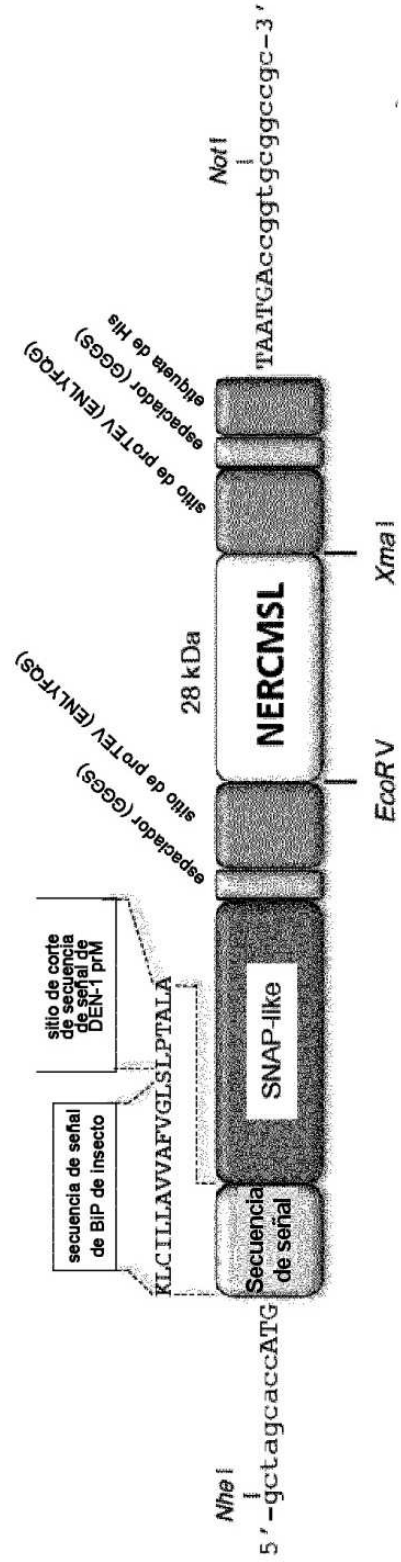


Figura 12C

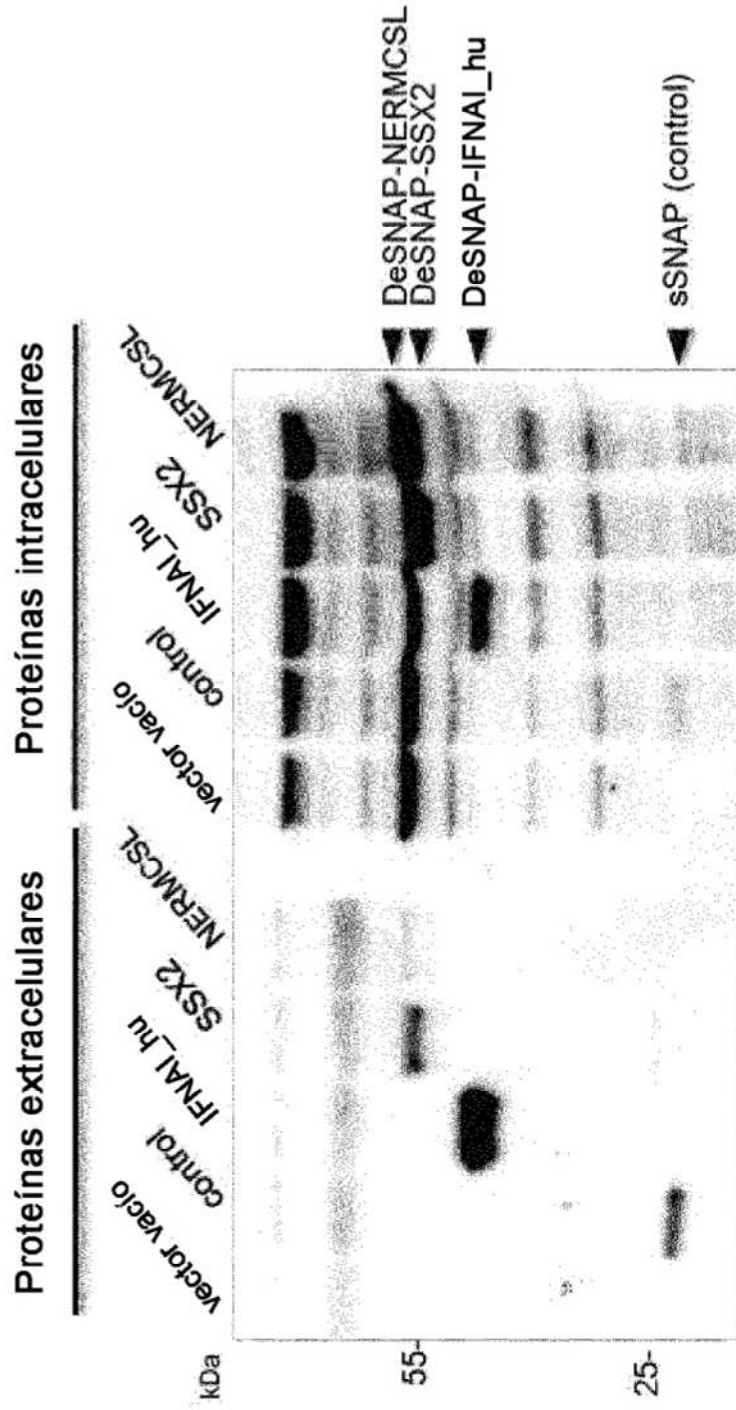


Figura 13A

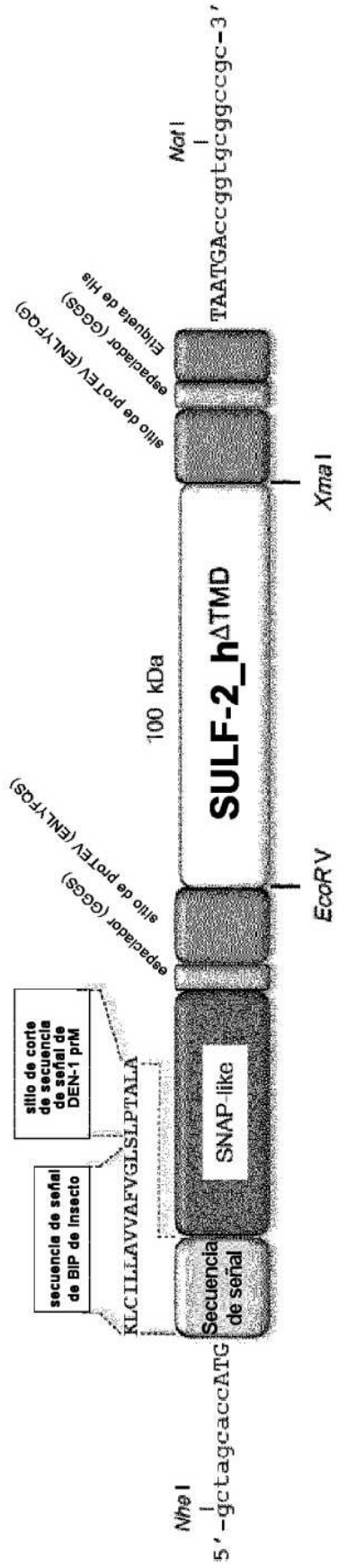






Figura 13B (continuación)

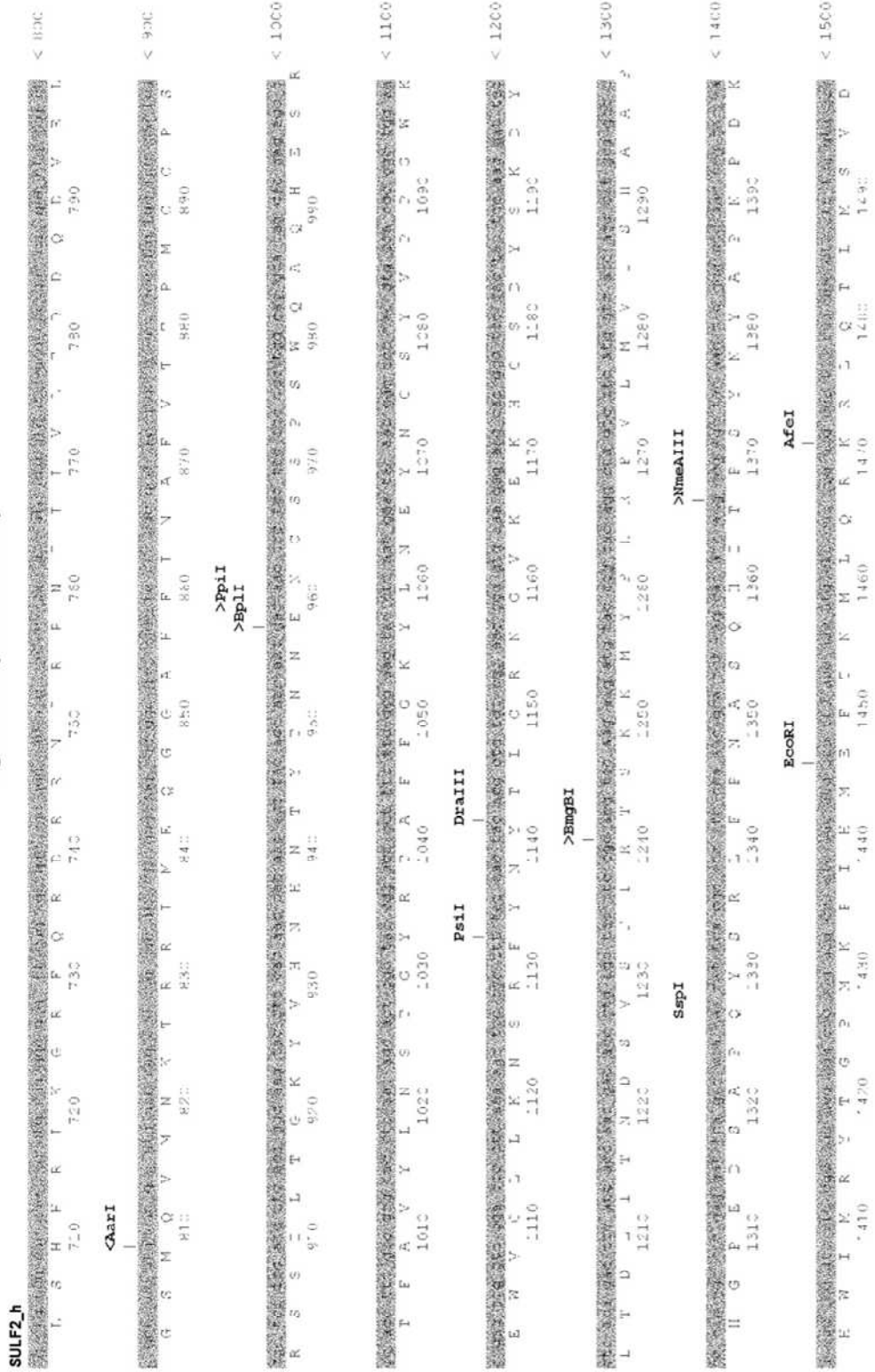


Figura 13B (continuación)

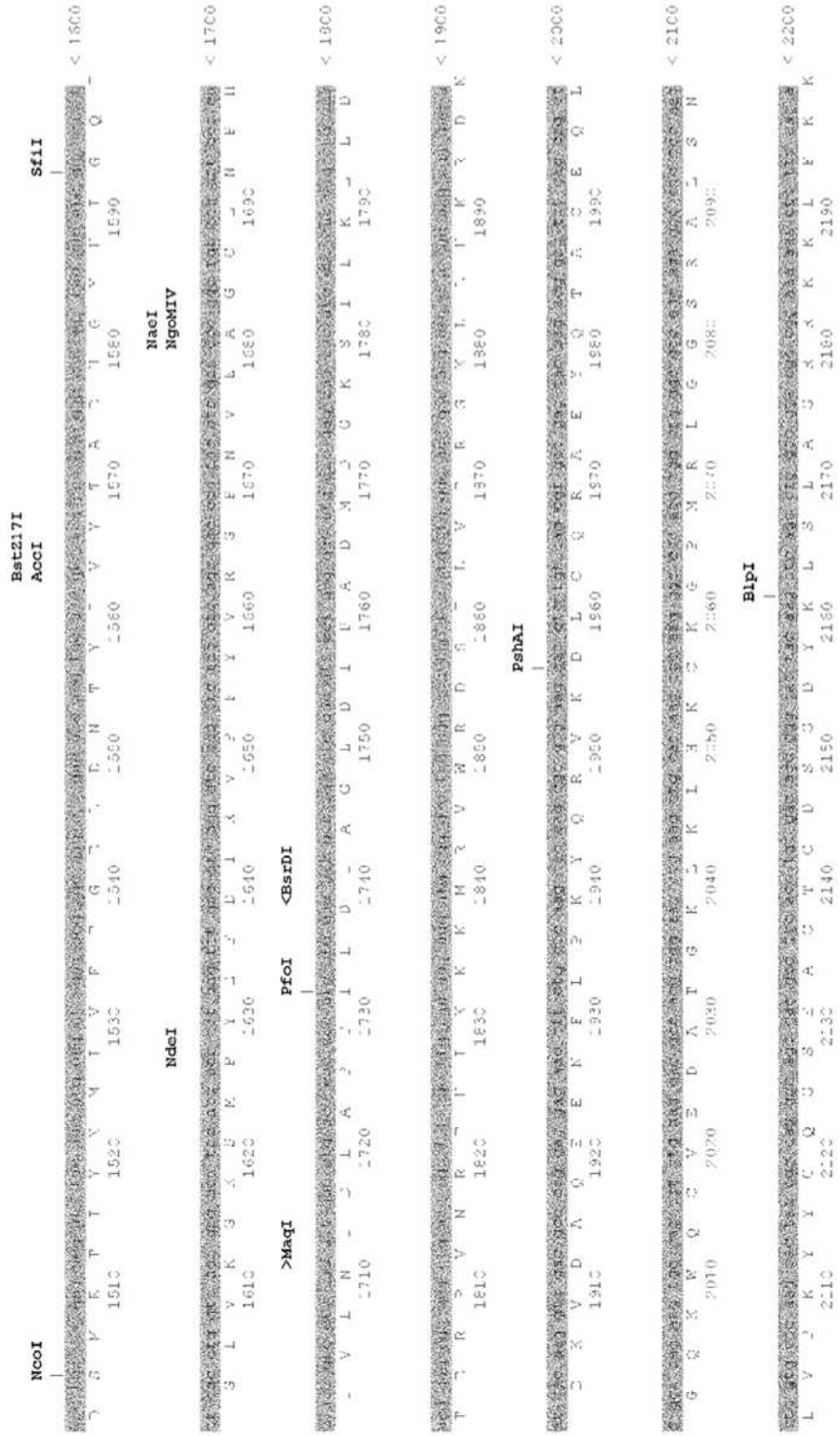


Figura 13B (continuación)

```

K Y K A S V V R S K S T R S V A - E V E G R V V F V G , G E A A G < 2300
2210 2220 2230 2240 2250 2260 2270 2280 2290
>CapCI
F A N L T A S A W P C A P A D Q D D A D C C D E S C H C G L E D Y < 2400
2310 2320 2330 2340 2350 2360 2370 2380 2390
Tch111 BsrGI SbfI
S A A N F I A V I J A C Y I L J N D I V Q C D L D L Y A E L Q A W S < 2500
2410 2420 2430 2440 2450 2460 2470 2480 2490
>ApyPI
D H K L H I E H E I E T L Q N K E K N E R E V R G H E K K K R P E < 2600
2510 2520 2530 2540 2550 2560 2570 2580 2590
E C D C H X I S Y H T Q H K E R L X H R G S S L H P F R X G L Q E < 2700
2610 2620 2630 2640 2650 2660 2670 2680 2690
Sphi
A D A V W L L A E Q A A A A K L L A K A L Q N N D T C S M F C L < 2800
2710 2720 2730 2740 2750 2760 2770 2780 2790
Kasi SfoI Nari Bbei Bani
T C P T H C N G T W G T A P T W S L G S - G A C C S A V V A V W < 2900
2810 2820 2830 2840 2850 2860 2870 2880 2890

```



Figura 13C

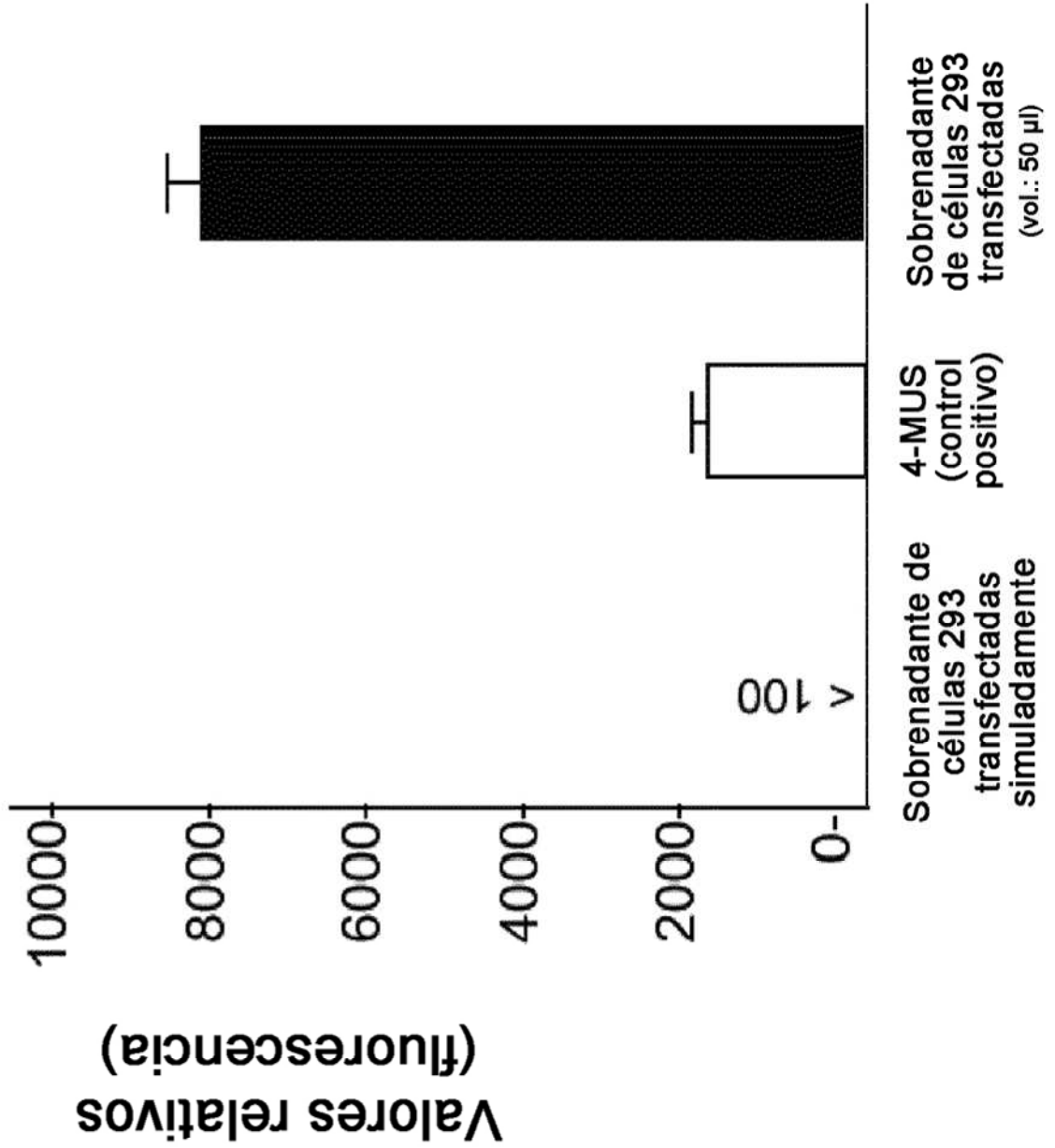


Figura 14A

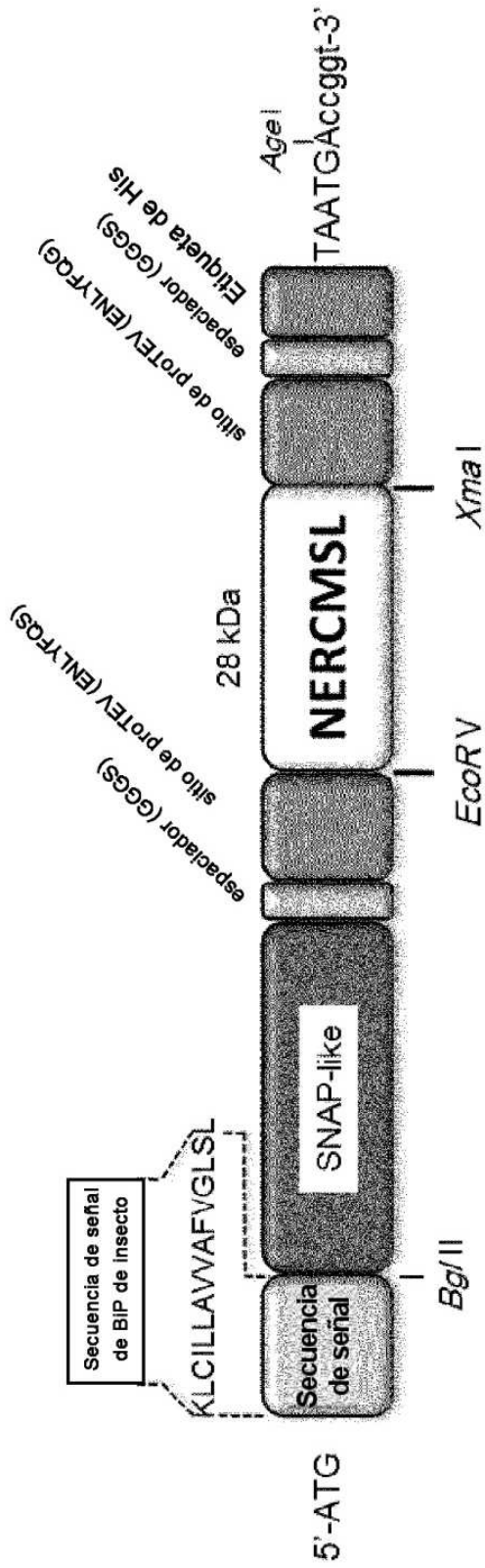


Figura 14B

