



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

① Número de publicación: **2 296 423**

② Número de solicitud: 200301830

⑤ Int. Cl.:  
**C12N 15/67** (2006.01)

**C07K 16/18** (2006.01)

⑫

PATENTE DE INVENCION

B1

⑫ Fecha de presentación: **31.07.2003**

⑬ Fecha de publicación de la solicitud: **16.04.2008**

Fecha de la concesión: **26.02.2009**

⑮ Fecha de anuncio de la concesión: **16.03.2009**

⑯ Fecha de publicación del folleto de la patente:  
**16.03.2009**

⑰ Titular/es:  
**Consejo Superior de Investigaciones Científicas  
Serrano, 117  
28006 Madrid, ES**

⑱ Inventor/es: **Lorenzo Prieto, Víctor de y  
Fernández Herrero, Luis Ángel**

⑳ Agente: **Arias Sanz, Juan**

⑳ Título: **Construcción de ADN para la producción de proteínas de fusión diméricas y sus aplicaciones.**

㉑ Resumen:

Construcción de ADN para la producción de proteínas de fusión diméricas y sus aplicaciones.

El sistema comprende una construcción de ADN que comprende: a) una *primera secuencia de ácido nucleico* que contiene la secuencia de nucleótidos que codifica para un producto de interés; b) una *segunda secuencia de ácido nucleico* que contiene la secuencia de nucleótidos que codifica para un dominio de dimerización; y c) una *tercera secuencia de ácido nucleico* que contiene la secuencia de nucleótidos que codifica para la  $\alpha$ -hemolisina (HlyA) de *Escherichia coli* o para un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de hemolisina (Hly) de *E. coli*; en donde el extremo 3' de dicha primera secuencia de ácido nucleico está unido al extremo 5' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha segunda secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico. De aplicación en la producción de proteínas recombinantes diméricas.

ES 2 296 423 B1

Aviso: Se puede realizar consulta prevista por el art. 37.3.8 LP.

## DESCRIPCIÓN

Construcción de ADN para la producción de proteínas de fusión diméricas y sus aplicaciones.

5 **Campo de la invención**

Esta invención se relaciona con la producción de proteínas recombinantes diméricas mediante el empleo de un sistema de expresión de proteínas basado en el sistema de transporte de hemolisina de *Escherichia coli*.

10 **Antecedentes de la invención**

Desde hace algún tiempo se está investigando en la producción de proteínas de fusión que comprenden fragmentos de anticuerpos recombinantes bi- o multifuncionales (minianticuerpos). Estas proteínas de fusión presentan algunas ventajas y pueden ser utilizadas con fines terapéuticos o de diagnóstico. Por este motivo, se han ido desarrollando diversos sistemas de expresión de fragmentos de anticuerpos. Algunos de estos sistemas de expresión se basan en el empleo de *Escherichia coli*.

Habitualmente se seleccionan diferentes fragmentos de anticuerpos y se producen en *E. coli* tras la donación de fragmentos de las regiones variable (V) y constante (C) de inmunoglobulinas (Ig) en vectores de fagos filamentosos o fagomídios [Hoogenboom, H. R. 1997. Designing and optimizing library selection strategies for generating high-affinity antibodies Trends in Biotechnology. **15**:62-70; Hoogenboom, H. R. 2002. Overview of antibody phage-display technology and its applications Methods Mol Biol. **178**:1-37; Winter, G., A. D. Griffiths, R. E. Hawkins, and H. R. Hoogenboom 1994. Making antibodies by phage display technology Annual Rev. Immunol. **12**:433-455]. Estos fragmentos reconstruyen el sitio de unión al antígeno del anticuerpo original que, en general, se ensambla mediante el contacto de los dominios V de las cadenas pesadas (H) y ligeras (L) [Ay, J., T. Keitel, G. Kuttner, H. Wessner, C. Scholz, M. Hahn, and W. Hohne 2000. Crystal structure of a phage library-derived single-chain Fv fragment complexed with turkey egg-white lysozyme at 2.0 Å resolution J Mol Biol. **301**:239-46]. Este es el caso de las moléculas Fab, que consisten en la asociación de dos polipéptidos que contienen los dominios V<sub>H</sub>-C<sub>H1</sub> y V<sub>L</sub>-C<sub>L</sub> y de las moléculas Fv (scFv) de cadena simple, en las que los dominios V<sub>H</sub> y V<sub>L</sub> se unen en un único polipéptido. Las moléculas de Fab y scFv tienen ventajas importantes, tales como niveles de expresión más elevados en *E. coli* y una mejor distribución y aclaramiento más rápido cuando se administran in vivo para aplicaciones diagnósticas o terapéuticas [Carter, P., and A. M. Merchant 1997. Engineering antibodies for imaging and therapy Current Opinion in Biotechnology. **8**:449-454; Marasco, W. A., and S. Dana Jones 1998. Antibodies for targeted gene therapy: extracellular gene targeting and intracellular expression Adv Drug Deliv Rev. **31**:153-170; Yokota, T., D. E. Milenic, M. Whitlow, and J. Schlom 1992. Rapid tumor penetration of a singlechain Fv and comparison with other immunoglobulin forms Cancer Res. **52**:3402-8].

También se han producido en *E. coli* fragmentos de anticuerpo basados en un único dominio de Ig [Nuttall, S. D., R. A. Irving, and P. J. Hudson 2000. Immunoglobulin VH domains and beyond: design and selection of single-domain binding and targeting reagents Curr Pharm Biotechnol. **1**:253-63; Riechmann, L., and S. Muyldermans 1999. Single domain antibodies: comparison of camel VH and camelised human VH domains J Immunol Methods. **231**:25-38; Sheriff, S., and K. L. Constantine 1996. Redefining the minimal antigen-binding fragment Nat Struct Biol. **3**:733-6]. Este progreso se realizó gracias al hallazgo de que en las especies de camélidos (por ejemplo, llamas, camellos) una proporción de sus anticuerpos naturales carece de cadena ligera, construyendo por tanto sus superficie de unión a antígenos con un único dominio V de la cadena pesada (V<sub>HH</sub>). Además de los beneficios de scFv y Fabs, los dominios V<sub>HH</sub> han demostrado una estabilidad y solubilidad superiores y una inmunogenicidad inferior [Cortez-Retamozo, V., M. Lauwereys, G. Hassanzadeh Gh, M. Gobert, K. Conrath, S. Muyldermans, P. De Baetselier, and H. Revets 2002. Efficient tumor targeting by single-domain antibody fragments of camels Int J Cancer. **98**:456-62; Nuttall, S. D., R. A. Irving, and P. J. Hudson 2000. Immunoglobulin VH domains and beyond: design and selection of single-domain binding and targeting reagents Curr Pharm Biotechnol. **1**:253-63; Riechmann, L., and S. Muyldermans 1999. Single domain antibodies: comparison of camel VH and camelised human VH domains J Immunol Methods. **231**:25-38]. Sin embargo, todos estos fragmentos de anticuerpo pierden la bivalencia de unión a antígenos (o multivalencia) mostrada por los anticuerpos completos. Su carácter monovalente se refleja por una disminución en la afinidad funcional (avidez) para sus antígenos correspondientes. Para solventar este problema, se obtuvieron por ingeniería genética cortos dominios de oligomerización (por ejemplo, hélices anfipáticas) en los extremos C terminales para producir minianticuerpos bivalentes y tetravalentes con avidez idéntica a los anticuerpos completos [Pack, P., M. Kujau, V. Schroeckh, U. Knupfer, R. Wenderoth, D. Riesenberger, and A. Plückthun 1993. Improved bivalent miniantibodies, with identical avidity as whole antibodies, produced by high cell density fermentation of *Escherichia coli* Biotechnology (N Y). **11**:1271-7; Pack, P., K. Muller, R. Zahn, and A. Plückthun 1995. Tetravalent miniantibodies with high avidity assembling in *Escherichia coli* J Mol Biol. **246**:28-34; Pack, P., and A. Plückthun 1992. Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric FV fragments with high avidity in *Escherichia coli* Biochemistry. **31**:1579-84; Plückthun, A., and P. Pack 1997. New protein engineering approaches to multivalent and bispecific antibody fragments Immunotechnology. **3**:83-105; Rheinnecker, M., C. Hardt, L. L. Ilag, P. Kufer, R. Gruber, A. Hoess, A. Lupas, C. Rottenberger, A. Plückthun, and P. Pack 1996. Multivalent antibody fragments with high functional affinity for a tumor-associated carbohydrate antigen J Immunol. **157**:2989-97].

En casi todos los casos, los fragmentos de anticuerpos monovalentes y multivalentes se han producido en el espacio periplásmico de *E. coli* fusionándoles un péptido señal en el extremo N terminal (N-SP) que es reconocido por

la maquinaria celular de la ruta de secreción general (Sec) [Plückthun, A., C. Krebber, U. Krebber, U. Horn, U. Knüpfer, R. Wenderoth, L. Nieba, K. Proba, and D. Riesenber 1996. Producing antibodies in *Escherichia coli*: from PCR to fermentation, p. 203-252. In J. McCafferty, and H. R. Hoogenboom (eds), *Antibody Engineering: A Practical Approach*. IRL Press, Oxford]. Recientemente se ha descrito un método alternativo para la producción de scFvs funcionales en el medio extracelular de cultivos de *E. coli* que emplean el transportador de la hemolisina  $\alpha$  (Hly) [Fernández, L. A., I. Sola, L. Enjuanes, and V. de Lorenzo 2000. Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system *Appl Environ Microbiol.* **66**:5024-5029]. Este sistema de secreción es independiente de los genes sec celulares y consta de dos componentes de membrana interna (IM), HlyB y HlyD, y el poro de la membrana externa (OM), TolC, que se ensamblan en un gran complejo proteico con un canal hidrófilo interno [Gentschev, I., G. Dietrich, and W. Goebel 2002. The *E. coli* alpha-hemolysin secretion system and its use in vaccine development *Trends Microbiol.* **10**:39-45; Koronakis, V., C. Andersen, and C. Hughes 2001. Channel-tunnels *Curr Opin Struct Biol.* **11**:403-7; Koronakis, V., A. Sharff, E. Koronakis, B. Luisi, and C. Hughes 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export *Nature.* **405**:914-919; Thanabalu, T., E. Koronakis, C. Hughes, and V. Koronakis 1998. Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore *EMBO J.* **17**:6487-96]. El sustrato natural de este sistema, la toxina de la hemolisina  $\alpha$  (HlyA), se expulsa a través de este canal directamente desde el citoplasma hacia el medio extracelular sin un intermediario periplásmico y de una forma dependiente de ATP. La señal reconocida por la maquinaria de secreción de Hly se localiza en el extremo C terminal de HlyA. Se ha demostrado que los híbridos scFv-HlyA, que contienen una molécula de scFv que carece del N-SP unido al último HlyA de aproximadamente 23 kDa, se secretan de una forma funcional; y oxidada por el transportador de Hly [Fernández, L. A., and V. De Lorenzo 2001. Formation of disulphide bonds during secretion of proteins through the periplasmic-independent type I pathway *Mol Microbiol.* **40**:332-46; Fernández, L. A., I. Sola, L. Enjuanes, and V. de Lorenzo 2000. Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system *Appl Environ Microbiol.* **66**:5024-5029].

Por otra parte, la dimerización es una propiedad que con frecuencia se desea conseguir por ingeniería genética en las proteínas cuando está implicada una actividad de unión (por ejemplo, en las interacciones proteína-ADN o antígeno-anticuerpo), puesto que puede intensificar su afinidad funcional (avidez) [Baxevanis, A. D., and C. R. Vinson 1993. Interactions of coiled coils in transcription factors: where is the specificity? *Curr Opin Genet Dev.* **3**:278-85; Busch, S. J., and P. Sansone-Corsi 1990. Dimers, leucine zippers and DNA-binding domains *Trends Genet.* **6**:36-40; Crothers, D. M., and H. Metzger 1972. The influence of polyvalency on the binding properties of antibodies *Immunochemistry.* **9**:341-357; Plückthun, A., and P. Pack 1997. New protein engineering approaches to multivalent and bispecific antibody fragments *Immunotechnology.* **3**:83-105]. Un procedimiento para la producción de proteínas dimericas que comprenden dos proteínas de fusión monoméricas en una interacción no covalente ha sido descrito en la patente norteamericana US 5.910.573. Las proteínas dimericas así obtenidas se acumulan dentro de la célula en el espacio periplásmico, sin secretarse al medio extracelular. Esto conduce a una mayor toxicidad de su expresión para la bacteria *E. coli*, induciendo un menor rendimiento en los cultivos (peso seco de células por litro), y dificultando la posterior purificación del anticuerpo dimerico al tener que lisar (romper) las bacterias.

Por tanto, sigue existiendo la necesidad de desarrollar sistemas alternativos para la producción de proteínas dimericas.

### Compendio de la invención

La invención proporciona una solución a la necesidad existente basada en el desarrollo de una construcción de ADN que comprende (i) una secuencia de nucleótidos que codifica para la  $\alpha$ -hemolisina (HlyA) de *Escherichia coli* o para un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de hemolisina (Hly) de *E. coli*; (ii) una secuencia de nucleótidos que codifica para un dominio de dimerización; y (iii) una secuencia de nucleótidos que codifica para un producto de interés. Mediante el empleo de dicho sistema de expresión y secreción de proteínas se obtienen proteínas dimericas en el medio. La eficacia de dicho sistema de secreción ha sido demostrada mediante la producción de minianticuerpos de alta avidez derivados de los anticuerpos de camello V<sub>HH</sub> (Ejemplo 1).

El translocador de hemolisina de *E. coli* se había usado anteriormente para la secreción de polipéptidos heterólogos, especialmente toxinas y antígenos de patógenos, así como para la secreción de anticuerpos recombinantes scFv. Sin embargo, los resultados ahora obtenidos han puesto de manifiesto que la incorporación de una hélice anfipática de autodimerización en el extremo N terminal de C-HlyA no interfiere con la secreción de Hly y permite la dimerización del polipéptido secretado. Asimismo, la dimerización intensifica la avidez de la unión del polipéptido secretado derivado de C-HlyA. Además, también puede tener otras aplicaciones, como la asociación molecular de varios antígenos y/o adyuvantes producidos por cepas bacterianas vivas, o la combinación de diversas actividades biológicas para la generación de moléculas biespecíficas (por ejemplo, la unión a un antígeno y el reclutamiento del complemento).

Por tanto, un aspecto de esta invención se relaciona con una construcción de ADN que comprende (i) una secuencia de nucleótidos que codifica para la  $\alpha$ -hemolisina (HlyA) de *Escherichia coli* o para un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de hemolisina (Hly) de *E. coli*; (ii) una secuencia de nucleótidos que codifica para un dominio de dimerización; y (iii) una secuencia de nucleótidos que codifica para un producto de interés.

En otro aspecto, la invención se relaciona con un cassette de expresión que comprende dicha construcción de ADN operativamente unida a una secuencia de control de expresión de la secuencia de nucleótidos que codifica para el producto de interés.

5 En otro aspecto, la invención se relaciona con una bacteria que comprende dicha construcción de ADN o dicho cassette de expresión.

En otro aspecto, la invención se relaciona con un método para producir un producto de interés, en forma de una proteína de fusión dimérica, que comprende crecer dicha bacteria bajo condiciones que permiten la producción y excreción al medio de cultivo de dicho producto de interés en forma de una proteína de fusión dimérica.

En otro aspecto, la invención se relaciona con una proteína de fusión dimérica obtenible por expresión de la secuencia de ácido nucleico contenida dicha construcción de ADN.

## 15 Breve descripción de las figuras

La Figura 1 ilustra la secreción del polipéptido C-HlyA que contiene el dominio ZIP. La Figura 1A muestra una representación esquemática de la estructura de los polipéptidos EHlyA y ZEHlyA que contienen la señal de secreción de 23 kDa (*'hlyA*) del transportador de Hly de *E. coli* marcado con el epítipo E. La masa de dichos polipéptidos (en kDa), deducida de su secuencia de aminoácidos, se muestra a la derecha. Se indica la composición del dominio ZIP (bisagra de Ig, cremallera de leucina, marca de 6xhis). También se muestra la secuencia de aminoácidos de la región N terminal de ambos polipéptidos EHlyA y ZEHlyA. La Figura 1B es una representación esquemática del polipéptido C-HlyA (monomérico - marcado con el epítipo E y del polipéptido C-HlyA (dimérico) marcado con el epítipo E y que contiene el dominio ZIP (bisagra de Ig, cremallera de leucina, marca de 6xhis). La Figura 1C muestra el resultado de la inmuno-transferencia desarrollada con un anticuerpo monoclonal anti-E marcado con POD de las proteínas secretadas (S) y celulares (C) producidas tras la inducción de 4 h con IPTG 0,3 mM de cultivos de células de *E. coli* HB2151, crecidas a 37°C, que contienen el plásmido pVDL9.3 (que codifica para HlyB y HlyD) y uno de los plásmidos indicados, pEHlyA o pZEHlyA. Las proteínas cargadas por carril representan las encontradas en aproximadamente 5  $\mu$ l de los sobrenadantes (S) del cultivo y las de las células de *E. coli* (C) presentes en aproximadamente 100  $\mu$ l de los mismos cultivos (DO<sub>600 nm</sub> aproximadamente 2).

La Figura 2 ilustra el entrecruzamiento de los polipéptidos C-HlyA secretados con glutarato de disuccinimidilo (DSG). Los polipéptidos EHlyA y ZEHlyA secretados (10  $\mu$ g/ml en PBS aproximadamente) se incubaron con DSG, a las concentraciones indicadas, y se sometieron a SDS-PAGE desnaturizante y a inmunotransferencia con anticuerpo monoclonal anti-E marcado con POD (véase el Ejemplo 1, apartado relativo a los Materiales y Métodos, para más detalles). Tal como se muestra, ZEHlyA se entrecruzó con DSG formando una banda proteica en SDS-PAGE de aproximadamente 66 kDa, unas dos veces el tamaño de su monómero.

La Figura 3 muestra los resultados de la cromatografía de filtración en gel de los polipéptidos de C-HlyA monoméricos y diméricos. La Figura 3A es un gráfico que representa el volumen de elución de los polipéptidos EHlyA (círculo) y ZEHlyA (triángulo) separados por cromatografía de filtración en gel (véase el Ejemplo 1, apartado relativo a los Materiales y Métodos, para más detalles) junto con patrones de proteínas de masa conocida (cuadrados). Los patrones de masa utilizados fueron la tiroglobulina (Mr 670.000), la gammaglobulina bovina (Mr 158.000), la ovoalbúmina de pollo (Mr 44.000) y la mioglobina equina (Mr 17.000). La presencia de EHlyA o ZEHlyA en las fracciones eluidas se determinó por inmunotransferencia con anticuerpo monoclonal anti-E marcado con POD. La Figura 3B muestra el resultado de la inmunotransferencia desarrollada con un anticuerpo monoclonal anti-E marcado con POD de los polipéptidos EHlyA y ZEHlyA. Una representación esquemática de EHlyA (monomérico) y de ZEHlyA (dimérico) se muestra en la parte superior.

La Figura 4 ilustra la secreción de polipéptidos V<sub>HH</sub>-HlyA. monoméricos (V<sub>amy</sub>-HlyA) y diméricos (V<sub>amy</sub>-ZHlyA). La Figura 4A es una representación esquemática de la estructura de los polipéptidos V<sub>amy</sub>-HlyA y V<sub>amy</sub>-ZHlyA que contienen la señal de secreción de 23 kDa (*'hlyA*) del transportador de Hly de *E. coli* marcado con el epítipo E. La masa de dichos polipéptidos (en kDa), deducida de su secuencia de aminoácidos, se muestra a la derecha. La Figura 4B es una representación esquemática del polipéptido V<sub>amy</sub>-ZHlyA (monomérico) marcado con el epítipo E y del polipéptido V<sub>amy</sub>-ZHlyA (dimérico) marcado con el epítipo E y que contiene el dominio ZIP (bisagra de Ig, cremallera de leucina, marca de 6xhis). La Figura 4C muestra los resultados de un Western blot en donde se pone de manifiesto la secreción de híbridos proteicos que tienen dominios V<sub>HH</sub> y -EHlyA o -ZHlyA. Las células de *E. coli* HB2151 (pVDL9.3) que llevan uno de los plásmidos indicados (pV<sub>amy</sub>HlyA o pV<sub>amy</sub>ZHlyA) se indujeron con IPTG (véase el Ejemplo 1, apartado relativo a los Materiales y Métodos, para más detalles), a la temperatura indicada, y la presencia de polipéptidos secretados marcados con el epítipo E en sobrenadantes de cultivos se determinó por inmunotransferencia con anticuerpo monoclonal anti-E marcado con POD. Los polipéptidos V<sub>amy</sub>HlyA y V<sub>amy</sub>ZHlyA de longitud completa se detectaron en los sobrenadantes de cultivos, junto con algunos fragmentos proteolíticos derivados de ellos. La Figura 4D es un gráfico del volumen de elución de los polipéptidos V<sub>amy</sub>HlyA (círculo) y V<sub>amy</sub>ZHlyA (triángulo) separados por cromatografía de filtración en gel, junto con patrones de proteínas de masa conocida (cuadrados) y detectados con inmunotransferencia con anticuerpo monoclonal anti-E marcado con POD. Los patrones de masa utilizados fueron los mismos que los utilizados en relación con la Figura 3.

## ES 2 296 423 B1

La Figura 5 es una gráfica que ilustra la actividad de unión de los polipéptidos VHH-HlyA monoméricos y diméricos. La unión de la  $\alpha$ -amilasa mediante los polipéptidos  $V_{amy}$ HlyA y  $V_{amy}$ ZhlyA, a las concentraciones indicadas, se determinó mediante ELISA (véase el Ejemplo 1, apartado relativo a los Materiales y Métodos, para más detalles). Los minianticuerpos marcados con el epítipo E unidos se detectaron con anticuerpo monoclonal anti-E marcado con POD y lectura de la D.O. a 490 nm. Como controles de la especificidad se emplearon  $V_{tox}$ HlyA y  $V_{tox}$ ZhlyA. La unión previa a un antígeno control no relacionado (ovoalbúmina) se ha sustraído ( $DO_{490\text{ nm}} \leq 0,05$ ). Los datos presentados son la media de los triplicados de cada punto. Se realizaron dos experimentos adicionales independientes, que demostraron valores similares a los mostrados en la figura.

La Figura 6 muestra el mapa del plásmido pZEHlyA.

La Figura 7 muestra el mapa del plásmido pZEHlyA2-SD.

La Figura 8 muestra el mapa del plásmido p $V_{amy}$ HlyA.

La Figura 9 muestra el mapa del plásmido p $V_{amy}$ ZhlyA.

### Descripción detallada de la invención

En un aspecto, la invención proporciona una construcción de ADN, en adelante construcción de ADN de la invención, que comprende:

- a) una *primera secuencia de ácido nucleico* que contiene la secuencia de nucleótidos que codifica para un producto de interés;
- b) una *segunda secuencia de ácido nucleico* que contiene la secuencia de nucleótidos que codifica para un dominio de dimerización; y
- c) una *tercera secuencia de ácido nucleico* que contiene la secuencia de nucleótidos que codifica para la  $\alpha$ -hemolisina (HlyA) de *E. coli* o para un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*, o una secuencia de nucleótidos que codifica para un gen homólogo, o una secuencia de nucleótidos que codifica para una variante, natural o artificial, de HlyA o de un fragmento de la misma que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*;

en donde el extremo 3' de dicha primera secuencia de ácido nucleico está unido al extremo 5' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha segunda secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico.

La *primera secuencia de ácido nucleico* contiene la secuencia de nucleótidos que codifica para un producto de interés (gen de interés). El producto de interés puede ser eucarótico, procariótico, viral, etc. Prácticamente cualquier péptido o proteína susceptible de ser expresado de forma recombinante puede ser utilizado en la construcción de ADN de la invención, por ejemplo, enzimas, inhibidores enzimáticos, hormonas, moléculas implicadas en la adhesión y/o señalización celular y compuesta por dominios, por ejemplo, inmunoglobulinas, etc. A modo ilustrativo, dicho producto de interés puede ser un antígeno inmunogénico tal como una proteína o un fragmento antigénico de la misma procedente de un patógeno, por ejemplo, de un patógeno viral, bacteriano, de un parásito, etc., que puede causar infecciones en seres humanos o animales; un agente terapéutico, por ejemplo, un antígeno específico de tumor, un antígeno de una enfermedad auto-inmune, etc.; o una molécula inmunoreguladora, por ejemplo, factores de crecimiento, citoquinas, tales como interleuquinas, interferones, etc. En una realización particular, dicho producto de interés es un minianticuerpo.

La *segunda secuencia de ácido nucleico* contiene la secuencia de nucleótidos que codifica para un dominio de dimerización. Un dominio de dimerización es una secuencia peptídica que promueve la dimerización en las proteínas que lo contienen. Prácticamente cualquier dominio de dimerización puede ser utilizado en la construcción de ADN de la invención, por ejemplo, hélices peptídicas, que contienen, al menos, una hélice, o una estructura formada por una hélice, una vuelta y otra hélice, etc., estructuras de doble arrollamiento helicoidal (coiled coil), y, en general, cualquier secuencia peptídica que promueva la dimerización en las proteínas que la contengan. En una realización particular, dicho dominio de dimerización comprende la cremallera de leucina del factor de transcripción GCN4 de levadura.

La *tercera secuencia de ácido nucleico* comprende la secuencia de nucleótidos que codifica para la  $\alpha$ -hemolisina (HlyA) de *E. coli* o para un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*, o una secuencia de nucleótidos que codifica para un gen homólogo, o una secuencia de nucleótidos que codifica para una variante, natural o artificial, de HlyA o de un fragmento de la misma que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*. La señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli* parece estar comprendida en el extremo carboxilo terminal (C-terminal), concretamente dentro de los últimos 60 aminoácidos de HlyA. La secuencia de aminoácidos y nucleótidos de la HlyA de *E. coli* puede obtenerse de GeneBank, número de acceso M10133, donde también puede obtenerse la secuencia de nucleótidos de aminoácidos de Hlyb y HlyD. En una

realización particular, dicha tercera secuencia de ácido nucleico está constituida por la secuencia de nucleótidos que codifica para la HlyA de *E. coli*. En otra realización particular, dicha tercera secuencia de ácido nucleico comprende un fragmento de la HlyA de *E. coli* que contiene la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*, tal como una secuencia de nucleótidos que codifica para los 60 últimos aminoácidos del extremo C-terminal de HlyA de *E. coli*. En este caso, dicha tercera secuencia de ácido nucleico está constituida por, o comprende, la secuencia de nucleótidos que codifica para los 60 últimos aminoácidos del extremo C-terminal de HlyA de *E. coli*.

En una realización concreta de la invención, dicha tercera secuencia de ácido nucleico contiene la secuencia de nucleótidos identificada como SEQ ID NO: 1 que codifica para un péptido de aproximadamente 23 kDa del extremo carboxilo terminal de HlyA de *E. coli* cuya secuencia de aminoácidos se muestra en la SEQ ID NO: 2.

En general, el dominio de dimerización no se fusiona directamente al gen que codifica el producto de interés sino que es ventajoso introducir un péptido espaciador (flexible) entre el extremo del gen que codifica para el producto de interés y el comienzo del dominio de dimerización. Por tanto, si se desea, la construcción de ADN de la invención también puede contener, además, una *cuarta secuencia de ácido nucleico* que codifica para un péptido espaciador situada entre dichas primera y segunda secuencias de ácido nucleico, en donde el extremo 5' de dicha cuarta secuencia de ácido nucleico está unido al extremo 3' de dicha primera secuencia de ácido nucleico y el extremo 3' de dicha cuarta secuencia de ácido nucleico está unido al extremo 5' de dicha segunda secuencia de ácido nucleico. De esta manera, la secuencia codificante del producto de interés está unida al dominio de dimerización mediante un péptido espaciador. Ventajosamente, dicho péptido espaciador es un péptido con flexibilidad estructural. Prácticamente, cualquier péptido con flexibilidad estructural puede ser utilizado. A modo ilustrativo, dicho péptido flexible puede contener repeticiones de restos de aminoácidos, tales como Gly-Gly-Gly-Ser, o cualquier otra repetición de restos de aminoácidos adecuada, o bien la región bisagra de un anticuerpo. En una realización particular, dicho péptido espaciador flexible comprende la región bisagra de un anticuerpo y la construcción de ADN de la invención contiene la secuencia codificante para dicho péptido flexible. En una realización concreta de la invención, dicha cuarta secuencia de ácido nucleico contiene la secuencia de nucleótidos identificada como SEQ ID NO: 3 que codifica para un péptido de 10 aminoácidos que comprende la región bisagra de un anticuerpo cuya secuencia de aminoácidos se muestra en la SEQ ID NO: 4.

Para facilitar el aislamiento y purificación del péptido o proteína de fusión obtenida mediante la presente invención, la construcción de ADN de la invención puede contener, si se desea, una secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de aislamiento o purificación del péptido o proteína de fusión. Por tanto, en una realización particular, la construcción de ADN de la invención contiene, si se desea, una *quinta secuencia de ácido nucleico* que codifica para un péptido susceptible de ser utilizado con fines de aislamiento o purificación.

Prácticamente cualquier péptido o secuencia peptídica que permita el aislamiento o purificación del péptido o proteína de fusión puede ser utilizada, por ejemplo, una secuencia de polihistidina, una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad, por ejemplo, péptidos etiqueta tales como c-myc, HA, E, FLAG, etc. [Using Antibodies: A laboratory manual. Ed Harlow and David Lane (1999). Cold Spring Harbor Laboratory Press. New York. Capítulo: Tagging proteins. pp. 347-377] y, en general, cualquier otra secuencia reconocida por un anticuerpo.

Dicha quinta secuencia de ácido nucleico puede estar situada en cualquier posición de la construcción de ADN de la invención excepto en la región correspondiente al extremo C-terminal de HlyA ya que, en ese caso, rompería la señal de secreción. A modo ilustrativo, dicha quinta secuencia de ácido nucleico podría estar situada entre dichas segunda y tercera secuencias de ácido nucleico, en donde el extremo 5' de dicha quinta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha quinta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico. Alternativamente, dicha quinta secuencia de ácido nucleico podría estar situada en la región correspondiente al extremo N-terminal de la proteína de fusión resultante o entre el producto de interés y el dominio de dimerización.

Para facilitar el reconocimiento de la proteína o péptido de fusión obtenido, la construcción de ADN de la invención también puede contener, si se desea, una *sexta secuencia de ácido nucleico* que codifica para un péptido susceptible de ser utilizado con fines de reconocimiento.

Prácticamente cualquier péptido o secuencia peptídica que permita el reconocimiento del péptido o proteína de fusión puede ser utilizada, por ejemplo, una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para reconocer la proteína de fusión resultante por cromatografía de inmunoafinidad, por ejemplo, péptidos etiqueta tales como c-myc, HA, E, FLAG y, en general, cualquier otra secuencia reconocida por un anticuerpo.

Dicha sexta secuencia de ácido nucleico puede estar situada en cualquier posición de la construcción de ADN de la invención excepto en la región correspondiente al extremo C-terminal de HlyA para evitar que se rompa la señal de secreción. A modo ilustrativo, dicha sexta secuencia de ácido nucleico podría estar situada entre dichas segunda y tercera secuencias de ácido nucleico, en donde el extremo 5' de dicha sexta secuencia de ácido nucleico está unida al extremo 3' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico. Alternativamente, dicha sexta secuencia de ácido nucleico podría estar situada en la región correspondiente al extremo N-terminal de la proteína de fusión resultante o entre el producto de interés y el dominio de dimerización.

Dichas quinta y sexta secuencias de ácido nucleico pueden estar separadas entre sí. Alternativamente, en una realización particular, dichas quinta y sexta secuencias de ácido nucleico pueden estar unidas entre sí. En este caso, a modo ilustrativo, dicha sexta secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de reconocimiento puede estar situada entre dichas tercera y quinta secuencias de ácido nucleico, en donde el extremo 5' de dicha sexta secuencia de ácido nucleico está unido al extremo 3' de dicha quinta secuencia de ácido nucleico y el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico. Alternativamente, dichas secuencias pueden estar unidas entre sí en el orden inverso, en cuyo caso, dicha sexta secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de reconocimiento está situada entre dichas segunda y quinta secuencias de ácido nucleico, en donde el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo 5' de dicha segunda secuencia de ácido nucleico y el extremo 5' de dicha sexta secuencia de ácido nucleico está unido al extremo 3' de dicha quinta secuencia de ácido nucleico.

Si se desea, la construcción de ADN de la invención puede contener, además, una secuencia de nucleótidos que codifica para una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos con el fin de liberar el producto de interés una vez aislada la proteína de fisión. En este caso, la construcción de ADN de la invención puede incluir, además, una séptima secuencia de ácido nucleico que comprende una secuencia de nucleótidos que codifica para una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos. Prácticamente cualquier secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimático o químicos puede ser utilizada. En una realización particular, dicha séptima secuencia de ácido nucleico comprende una secuencia de nucleótidos que codifica para un sitio de reconocimiento de una proteasa, por ejemplo, una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa, Factor de coagulación Xa y similares. En otra realización particular, dicha séptima secuencia de ácido nucleico comprende una secuencia de nucleótidos que codifica para un sitio susceptible de ser específicamente escindido por un reactivo químico tal como, por ejemplo, bromuro de cianógeno que escinde restos de metionina o cualquier otro reactivo químico apropiado.

Dicha séptima secuencia de ácido nucleico se encuentra, generalmente, a continuación del extremo 3' de dicha segunda secuencia de ácido nucleico que codifica para el dominio de dimerización de manera que por medios enzimáticos o químicos se pueda escindir la proteína dimérica.

La construcción de ADN de la invención puede obtenerse mediante el empleo de técnicas ampliamente conocidas en el estado de la técnica [Sambrook *et al.*, "Molecular cloning, a Laboratory Manual", 2<sup>nd</sup> ed., Cold Spring Harbor Laboratory Press, N.Y., 1989 Vol 1-3] Dicha construcción de ADN de la invención puede incorporar, operativamente unida, una secuencia reguladora de la expresión, constituyendo de este modo un cassette de expresión.

Por tanto, en otro aspecto, la invención proporciona un cassette de expresión que comprende la construcción de ADN de la invención operativamente unida a una secuencia de control de expresión. Las secuencias de control son secuencias que controlan y regulan la transcripción y, en su caso, la traducción del producto de interés, e incluyen secuencias promotoras (*pT7*, *plac*, *pBAD*, *ptet*, etc.), secuencias codificantes para reguladores transcripcionales (*lacI*, *tetR*, *araC*, etc.), secuencias de unión a ribosomas (RBS), y/o secuencias terminadoras de transcripción (*t12*, etc.), etc. En una realización particular, dicha secuencia de control de expresión es funcional en bacterias, en particular, en bacterias Gram negativas.

Ventajosamente, dicho cassette de expresión comprende, además, un marcador o gen que codifica para un motivo o para un fenotipo que permita la selección de la célula hospedadora transformada con dicho cassette de expresión. Ejemplos ilustrativos de dichos marcadores que podrían estar presentes en el cassette de expresión de la invención incluyen genes de resistencia a antibióticos, por ejemplo, ampicilina, tetraciclina, kanamicina, cloranfenicol, espectinomycin, etc., genes de resistencia a compuestos tóxicos (telurito, mercurio, etc.).

La construcción de ADN de la invención, o el cassette de expresión proporcionado por esta invención, pueden ser insertados en un vector apropiado. Por tanto, en otro aspecto, la invención se relaciona con un vector, tal como un vector de expresión, que comprende dicha construcción de ADN o dicho cassette de expresión. La elección del vector dependerá de la célula hospedadora en la que se va a introducir posteriormente. A modo de ejemplo, el vector donde se introduce dicha secuencia de ADN puede ser un plásmido o un vector que, cuando se introduce en una célula hospedadora, se integra o no en el genoma de dicha célula. La obtención de dicho vector puede realizarse por métodos convencionales conocidos por los técnicos en la materia [Sambrook *et al.*, 1989, citado *supra*].

En otro aspecto, la invención se relaciona con una bacteria, en particular, una bacteria Gram negativa, que comprende una construcción de ADN de la invención o un cassette de expresión de la invención, en adelante bacteria de la invención. Dicha bacteria debe tener el sistema exportador de hemolisina (Hly) de *E. coli* para lo cual, si no lo tiene de forma nativa, se debe proporcionar dicho sistema a la bacteria transformándola con un vector que contenga los genes HlyB y HlyD, por ejemplo, el plásmido pVDL9.3 [Fernández, L.A. *et al.*, Applied and Environmental Microbiology, Nov. 2000, 5024-5029]. Prácticamente cualquier bacteria Gram negativa, por ejemplo, *E. coli*, *Salmonella typhimurium*, *Pseudomonas aeruginosa*, *Pseudomonas putida*, etc., puede ser transformada con la construcción de ADN de la invención o con el cassette de expresión de la invención. Para ello, las señales promotoras, reguladoras, marcadoras y orígenes de replicación deben ser optimizados para cada especie bacteriana. En una realización particular, dicha bacteria Gram negativa es *Escherichia coli*.

## ES 2 296 423 B1

La construcción de ADN de la invención puede ser utilizada para producir productos de interés. Por tanto, en otro aspecto, la invención se relaciona con un método para producir un producto de interés, en forma de una proteína de fusión dimérica, que comprende crecer una bacteria de la invención bajo condiciones que permiten la producción y ex-  
5 para optimizar el cultivo de la bacteria de la invención dependerán de la bacteria utilizada.

Si se desea, el método para producir un producto de interés proporcionado por esta invención incluye, además, el aislamiento y purificación de dicha proteína de fusión dimérica. En este caso, la construcción de ADN de la invención incluye, además, dicha séptima secuencia de ácido nucleico previamente definida que comprende una secuencia de  
10 nucleótidos que codifica para una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos con el fin de liberar el producto de interés una vez aislada la proteína de fusión. En una realización particular, dicha secuencia de nucleótidos codifica para un sitio de reconocimiento de una proteasa, por ejemplo, una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa, Factor de coagulación Xa y similares. En otra realización particular, dicha secuencia de nucleótidos codifica para un sitio susceptible de ser  
15 específicamente escindido por un reactivo químico tal como, por ejemplo, bromuro de cianógeno que asciende restos de metionina o cualquier otro reactivo químico apropiado.

En otro aspecto, la invención se relaciona con una proteína de fusión dimérica obtenible por expresión de la secuencia de ácido nucleico contenida en la construcción de ADN de la invención, caracterizada porque cada monómero  
20 comprende:

- (i) la secuencia de aminoácidos de la HlyA de *E. coli* o de un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de hemolisina (Hly) de *E. coli*;
- (ii) una secuencia de aminoácidos correspondiente a un dominio de dimerización; y
- (iii) la secuencia de aminoácidos de un producto de interés.

De forma más concreta, cada monómero de la proteína de fusión dimérica de la invención comprende

- (i) la secuencia completa de aminoácidos de HlyA de *E. coli*, o alternativamente, un fragmento de HlyA de *E. coli* que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*
- (ii) un dominio de dimerización, tal como una hélice peptídica, una estructura de doble arrollamiento helicoidal (coiled coil), o, en general, cualquier secuencia peptídica que promueva la dimerización en las proteínas que la contengan. En una realización particular, dicho dominio de dimerización comprende la cremallera de leucina del factor de transcripción GCN4 de levadura; y
- (iii) un producto de interés, por ejemplo, una enzima, un inhibidor enzimático, una hormona, una molécula implicada en la adhesión y/o señalización celular y compuesta por dominios, por ejemplo, una inmunoglobulina, un antígeno inmunogénico, tal como una proteína o un fragmento antigénico de la misma procedente de un patógeno, por ejemplo, de un patógeno viral, bacteriano, de un parásito, etc., que puede causar infecciones en seres humanos o animales, un agente terapéutico, por ejemplo, un antígeno específico de tumor. un antígeno de una enfermedad auto-inmune, etc., o una molécula inmunoreguladora, por ejemplo, un factor de crecimiento, una citoquina, tal como una interleuquina, un interferón, etc.; en una realización particular, dicho producto de interés es un minianticuerpo susceptible de ser utilizado con fines terapéuticos, de diagnóstico o de investigación.

Cada monómero de La proteína de fusión dimérica de la invención también puede contener, si se desea, (a) un péptido espaciador entre el producto de interés y el dominio de dimerización; ventajosamente, dicho péptido espaciador es un péptido con flexibilidad estructural, por ejemplo, un péptido que contiene repeticiones de restos de aminoácidos, tales como Gly-Gly-Gly-Ser, o cualquier otra repetición de restos de aminoácidos adecuada, o bien la región bisagra de un anticuerpo; en una realización particular, dicho péptido espaciador flexible comprende la región bisagra de un anticuerpo; y/o (b) un péptido para facilitar el aislamiento o purificación del péptido o proteína de fusión, por ejemplo, una secuencia de polihistidina, o una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad, por ejemplo, péptidos etiqueta tales como c-myc, HA, E, FLAG, y, en general, cualquier otra secuencia reconocida por un anticuerpo; y/o (c) un péptido que permita el reconocimiento del péptido o proteína de fusión, por ejemplo, una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para reconocer la proteína de fusión resultante por cromatografía de inmunoafinidad, por ejemplo, péptidos etiqueta tales como c-myc, HA, E, FLAG y, en general, cualquier otra secuencia reconocida por un anticuerpo; y/o (d) una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos, por ejemplo, una secuencia de aminoácidos que constituye un sitio de reconocimiento de una proteasa, por ejemplo, una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa, Factor de coagulación Xa y similares, o una secuencia de aminoácidos susceptible de ser específicamente escindida por un reactivo químico tal como, por ejemplo, bromuro de cianógeno y similares.

El sistema de producción de proteínas de fusión diméricas proporcionado por esta invención es particularmente útil para la producción de proteínas implicadas en una actividad de unión, por ejemplo, en las interacciones proteína-ADN o antígeno-anticuerpo, puesto que puede intensificar su avidez.

Una ventaja del sistema proporcionado por esta invención radica en que permite producir proteínas tóxicas para un huésped bacteriano que, de no ser por ese sistema, no se podrían expresar en dicho huésped bacteriano debido a su toxicidad. Como es conocido, la expresión por métodos recombinantes de proteínas tóxicas para un huésped bacteriano es muy complicada o prácticamente imposible cuando dicha proteína tóxica expresada no es exportada desde la bacteria hacia el exterior. Con el método proporcionado por esta invención se puede expresar una proteína tóxica o previamente inexpressable, o expresada a bajos niveles, para producir la proteína deseada en cantidades utilizables.

El siguiente ejemplo ilustra la invención sin que deba ser considerado como limitativo del alcance de la misma.

#### Ejemplo 1

##### *Producción de minianticuerpos diméricos con elevada afinidad secretados por el sistema de transporte de hemolisina (Hly) de E. coli*

En este ejemplo se describe la secreción de minianticuerpos diméricos en sobrenadantes de cultivos de *E. coli* que emplean el sistema de transporte de hemolisina (Hly). En primer lugar, se demostró que la dimerización puede conseguirse por ingeniería genética en el sistema de transporte de la Hly. Para ello se insertó una hélice anfipática (es decir, el dominio de cremallera de leucina del factor de transcripción de levaduras GCN4) en el extremo N terminal de una versión marcada; (E-tag) del dominio C terminal de 23 kDa de la hemolisina (EHlyA). Se comprobó que el polipéptido resultante (ZEHlyA) se secretaba eficazmente por las células de *E. coli* y se acumulaba en el medio de cultivo como un dímero estable. Después se utilizaron los vectores derivados de 'EHlyA y 'ZEHlyA para la secreción de los dominios V<sub>HH</sub> de inmunoglobulinas obtenidos de anticuerpos de camello. Se secretaron los híbridos VHH-EHlyA y VHH-ZEHlyA y se encontraron en el medio extracelular como monómeros y dímeros, respectivamente. Cuando se compararon con sus homólogos monoméricos, las moléculas diméricas VHH-ZEHlyA mostraron propiedades superiores de unión a su antígeno relacionado, con un aumento de 10 veces en su afinidad funcional (avidez). Este procedimiento permite obtener fácilmente minianticuerpos VHH monoméricos y diméricos con alta avidez a partir de los sobrenadantes de cultivos de *E. coli*, facilitando así la selección y la purificación de alto rendimiento de clones de VHH a partir de grandes librerías.

#### 1. Materiales y métodos

*Cepas bacterianas, crecimiento y condiciones de inducción.* Las cepas de *E. coli* K-12 empleadas eran DH5 $\alpha$ F' (*supE44*  $\Delta$ (*lacZYA-argF*)U169  $\Phi$ 80(*lacZ* $\Delta$ M15) *hsdR17 recA1 endA1 gyrA96 thi1 relA1*; Invitrogen) para la donación y la propagación de los plásmidos y HB2151 ( $\Delta$ *lac-pro, ara, naI<sup>r</sup>, thi, F'proAB lacI<sup>a</sup> lzcZ* $\Delta$ M15) [Carter, P., H. Bedouelle, and G. Winter 1985. Improved oligonucleotide site-directed mutagenesis using M13 vectors. *Nucleic Acids Res.* **13**:4431-4443] para la expresión de proteínas. Las bacterias que contenían los plásmidos indicados en cada caso se hicieron crecer a 30°C en placas de agar LB [Miller, J. H. 1992. A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York] que contenía glucosa al 2% (p/v) (para reprimir el promotor *lac*) y los antibióticos apropiados para la selección de plásmidos. Para la inducción de los híbridos HlyA, se inocularon colonias individuales en el medio LB líquido que contenía glucosa al 2% (p/v) y se hicieron crecer a 30°C o a 37°C hasta que a la densidad óptica a 600 nm (DO<sub>600 nm</sub>) alcanzó un valor de 0,5 aproximadamente. En ese punto, las bacterias se recogieron por centrifugación, se resuspendieron a la misma densidad en LB que contenía isopropil-1-tio- $\beta$ -D-galactósido (IPTG) 0,3 mM y se incubaron (a 30°C o a 37°C) con agitación (160 r.p.m.) durante un periodo de tiempo comprendido entre 4 y 16 h. Se recogieron los sobrenadantes del cultivo tras la eliminación de las células de *E. coli* por centrifugación (10.000xg, 10 min) y se añadió 1/10 del volumen de tampón fosfato salino (PBS) 10X concentrado [PBS: Na<sub>2</sub>HPO<sub>4</sub> 8 mM, KH<sub>2</sub>PO<sub>4</sub> 1,5 mM, KCl 3 mM, NaCl 137 mM, pH 7,0]. Los sobrenadantes del cultivo se emplearon directamente para realizar inmunoensayos o se almacenaron a -80°C hasta su utilización. Los antibióticos añadidos al medio de cultivo para la selección de plásmidos fueron la ampicilina (Ap; 150  $\mu$ g/ml) y el cloranfenicol (Cm; 30  $\mu$ g/ml).

*Plásmidos y oligonucleótidos.* Se emplearon métodos estándar para la manipulación y el aislamiento del ADN, la amplificación por PCR y la secuenciación del ADN [Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl 1994. *Current Protocols in Molecular Biology*. John Wiley & Sons, New York; Sambrook, J., E. Fritsch, and T. Maniatis 1989. *Molecular cloning. A laboratory manual*. Cold Spring Harbor Laboratory Press, New York]. Los oligonucleótidos se obtuvieron de Sigma Genosys (Reino Unido) o de Isogen Bioscience BV (Países Bajos). Los plásmidos pEHlyA (Ap<sup>r</sup>), pEHlyA2-SD (Ap<sup>r</sup>) y pVDL9.3 (Cm<sup>r</sup>) ya han sido descritos [Fernández, L. A., and V. De Lorenzo 2001. Formation of disulphide bonds during secretion of proteins through the periplasmic-independent type I pathway *Mol Microbiol.* **40**:332-46; Fernández, L. A., I. Sola, L. Enjuanes, and V. de Lorenzo 2000. Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system *Appl Environ Microbiol.* **66**:5024-5029; Tzschaschel, B. D., C. A. Guzman, K. N. Timmis, and V. de Lorenzo 1996. An *Escherichia coli* hemolysin transport system-based vector for the export of polypeptides: Export of Shiga-like toxin IIeB subunit by *Salmonella typhimurium aroA*. *Nature Biotechnology.* **14**:765-769]. El plásmido pZEHlyA (Ap<sup>r</sup>) se obtuvo insertando en el sitio único *SalI* de pEHlyA un fragmento de ADN de 170 pb que codifica para el dominio ZIP amplificado por PCR y digerido con *SalI*. El mapa del plásmido

pZEHlyA se muestra en la Figura 6. La amplificación por PCR del dominio ZIP se llevó a cabo con la ADN polimerasa Vent (New England Biolabs), usando como molde 1 ng de pCLZIP (Codon Genetic Systems, GmbH), y como cebadores los oligonucleótidos identificados como SEQ ID NO: 5 y SEQ ID NO: 6, que incorporaban dos sitios *Sall* flanqueando el producto amplificado. El plásmido pZEHlyA2-SD (Ap<sup>r</sup>) se obtuvo insertando en el sitio único *Sall* de pEHlyA2-SD el fragmento de ADN de 170 pb que codifica para ZIP obtenido por digestión con *Sall* de pZEHlyA. El mapa del plásmido pZEHlyA2-SD se muestra en la Figura 7. Se seleccionó la orientación del fragmento de ADN ZIP que producía una inserción interna en el dominio C-HlyA marcado con E de pZEHlyA y pZEHlyA2-SD tras la secuenciación del ADN. Los fragmentos de ADN de aproximadamente 0,3 kb que codificaban para los dominios V<sub>HH</sub>, V<sub>amy</sub> y V<sub>tx</sub>, se amplificaron por PCR con la ADN polimerasa Vent, usando como molde 1 ng de los fagémidos A100R3A2 (anti- $\alpha$ -amilasa) o R3E5 (vacuna antitetánica), respectivamente, y como cebadores los oligonucleótidos identificados como SEQ ID NO: 7 y SEQ ID NO: 8.

Los productos amplificados de ADN que codifican para V<sub>amy</sub> y V<sub>tx</sub> contenían los sitios de restricción flanqueantes *NcoI* y *SfiI*, lo que permitía su donación en los mismos sitios de pEHlyA2-SD y pZEHlyA2-SD, generando así pV<sub>amy</sub>HlyA, pV<sub>tx</sub>HlyA, pV<sub>amy</sub>ZhlyA y pV<sub>tx</sub>ZHlyA. Los mapas de los plásmidos pV<sub>amy</sub>HlyA y pV<sub>amy</sub>ZhlyA se muestran en las Figuras 8 y 9, respectivamente. Los fagémidos A100R3A2 y R3E5 fueron proporcionados por el Dr. Hennie Hoogenboon (Dyax Co., EE. UU.). Ambos fagémidos son derivados de pCANTAB6 (Cambridge Research Biochemicals) que contienen los dominios V<sub>HH</sub> de camélidos donados entre los sitios *SfiI* y *NotI*.

*Electroforesis e inmunotransferencia de proteínas.* Se llevó a cabo la electroforesis en geles de dodecil sulfato sódico - poliacrilamida (SDS-PAGE) usando geles de apilamiento al 4% y de separación al 10% (acrilamida:bisacrilamida 29:1; Bio-Rad), utilizando el sistema de electroforesis Miniprotein<sup>®</sup> (Bio-Rat) y siguiendo los protocolos estándar [Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl 1994. Current Protocols in Molecular Biology. John Wiley & Sons, New York; Fraile, S., F. Roncal, L. A. Fernández, and V. de Lorenzo 2001. Monitoring Intracellular Levels of XylR in *Pseudomonas putida* with a Single-Chain Antibody Specific for Aromatic-Responsive Enhancer-Binding Proteins J Bacteriol. **183**:5571-9]. Para la inmunotransferencia, las proteínas se transfirieron a una membrana de difloruro de polivinilideno (Immobilon-P, Millipore) usando un aparato de electroforesis por transferencia semiseca (Bio-Rad). La membrana se bloqueó en tampón MTP (leche desnatada al 3% p/v, Tween 20 al 0,1% v/v en PEAS) y se detectaron los polipéptidos marcados con E con anticuerpo monoclonal anti-E marcado con peroxidasa (0,2  $\mu$ g/ml en tampón MTP; Amersham Bioscience). El conjugado anticuerpo-POD unido se reveló mediante quimioluminiscencia, ta; como ya se ha descrito [Fraile, S., F. Roncal, L. A. Fernández, and V. de Lorenzo 2001. Monitoring Intracellular Levels of XylR in *Pseudomonas putida* with a Single-Chain Antibody Specific for Aromatic-Responsive Enhancer-Binding Proteins J Bacteriol. **183**:5571-9]. La membrana se expuso a una película de rayos X (X-OMAT<sup>®</sup>, Kodak) o a ChemiDoc<sup>®</sup> (Bio-Rad) para la cuantificación por quimioluminiscencia (software Quantity-one<sup>®</sup>; Bio-Rad). Las concentraciones de los polipéptidos HlyA marcados con E secretados presentes en los sobrenadantes del cultivo de *E. coli* se determinaron por la intensidad de sus bandas de proteínas correspondientes en geles de SDS-poliacrilamida teñidos con plata [Ansoorge, W. 1985. Fast and sensitive detection of protein and DNA bands by treatment with potassium permanganate J. Biochem. Biophys. Methods. **11**:13-20] y mediante inmuno-transferencia usando anticuerpo monoclonal anti-E marcado con POD. Se usaron diluciones seriadas de scFvs marcados con E purificados, de concentración conocida, como patrones en estos experimentos.

*Entrecruzamiento de proteínas.* Antes de su incubación con el agente de entrecruzamiento bifuncional glutarato de disuccinimidilo (DSG, 7.7 Å spacer; Pierce), los polipéptidos de HlyA marcados con E presentes en los sobrenadantes de los cultivos se equilibraron en el mismo volumen de PBS por ultrafiltración a través de una membrana con punto de corte de 10 kDa (Microcon 10, Millipore) que eliminó los compuestos pequeños con grupos amino libres presentes en el medio de cultivo. El entrecruzamiento se llevó a cabo durante 30 minutos a temperatura ambiente añadiendo DBS 40 o 130  $\mu$ M a muestras de proteínas de 50  $\mu$ l equilibradas en PBS. Tras esta incubación, el agente de entrecruzamiento se inactivó con Tris-HCl 50 mM (pH 7,5) durante 15 minutos y se añadió un volumen de tampón de muestra SDS-PAGE 2X concentrado [Fraile, S., F. Roncal, L. A. Fernández, and V. de Lorenzo 2001. Monitoring Intracellular Levels of XylR in *Pseudomonas putida* with a Single-Chain Antibody Specific for Aromatic-Responsive Enhancer-Binding Proteins J Bacteriol. **183**:5571-9]. Tras hervir durante 5 minutos, se cargaron 10  $\mu$ l para el SDS-PAGE.

*Cromatografía por exclusión de tamaño.* Los sobrenadantes del cultivo (aproximadamente 0,2 ml) que contenían PBS 1X se mezclaron con 2 mg de proteína patrón de masa conocida (disuelta en 60  $\mu$ l de H<sub>2</sub>O) y se hicieron pasar a través de una resina Bio-Gel A de 1,5m (Bio-Rad) empaquetada en una columna de 1 m de longitud y 1,5 cm de ancho (Bio-Rad). Los patrones de filtración del gel (Bio-Rad) fueron la tiroglobulina (PM 670.000), la gammaglobulina bovina (PM 158.000), la ovoalbúmina de pollo (PM 44.000), la mioglobina equina (PM 17.000) y la vitamina B-12 (PM 1.350). La velocidad de flujo de la muestra a través de la columna se fijó a 0,2 ml/min usando una bomba peristáltica (P-1, Amersham Bioscience). El volumen de vacío de la columna se calculó mediante la elución de azul de dextrano 2000 (Amersham Bioscience). La elución de los patrones de proteínas a través de la columna se monitorizó mediante absorción de luz UV (Uvicord S II, Amersham Bioscience). Se recogieron fracciones de 1 ml (colector RediFrac, Amersham Bioscience) y se concentraron 10 veces mediante precipitación con ácido tricloroacético (TCA) al 10% (p/v) y 10  $\mu$ g de albúmina sérica bovina (BSA, Roche) que actuaba como transportador. La presencia de proteínas HlyA marcadas con E en estas fracciones se detectó por Western blot utilizando un anticuerpo monoclonal anti-E marcado con POD (véase más arriba).

*Ensayos de inmunoadsorción ligados a enzima (ELISA).* Los antígenos ( $\alpha$ -amilasa u ovoalbúmina; Sigma) se adsorbieron durante 1 h a 37°C a inmunoplasmas de microtitulación de 96 pocillos (Maxisorb, Nunc) a 200  $\mu$ g/ml en PBS. Se lavó el exceso de antígeno y las placas se bloquearon durante 16 h a 4°C en tampón MTP (véase anteriormente). Los minianticuerpos se diluyeron en tampón MTP, se añadieron a los pocillos a las concentraciones indicadas en cada caso y se incubaron durante 1 h a temperatura ambiente. A continuación, los anticuerpos no unidos se eliminaron mediante cuatro lavados de los pocillos con PBS que contenía Tween 20 al 0,1% (v/v). El conjugado anticuerpo monoclonal anti-E marcado con POD (0,2  $\mu$ g/ml en tampón MTP) se añadió a los pocillos y se incubó adicionalmente durante 1 h a temperatura ambiente para detectar los minianticuerpos unidos marcados con E. Tras lavar como antes, las placas se revelaron usando *o*-fenilenediamina (Sigma). Se dejó continuar la reacción durante 10 minutos en oscuridad, se paró con HCl 0,6 N y se determinó la DO a 490 nm de los pocillos (lector de microplacas Benchmark, Bio-Rad).

## 2. Resultados

*Dimerización por ingeniería genética de la señal de secreción de HlyA.* En primer lugar, se investigó si la dimerización de C-HlyA podría realizarse por ingeniería genética sin afectar a su secreción. Se seleccionó un corto dominio de unos 6 kDa aproximadamente, denominado ZIP, que se había utilizado para la dimerización de scFvs en el periplasma de *E. coli* [Kerschbaumer, R. J., S. Hinschl, A. Kaufmann, M. Ibl, R. Koenig, and G. Himmler 1997. Single-chain Fv fusion proteins suitable as coating and detecting reagents in a double antibody sandwich enzyme-linked immunosorbent assay Anal Biochem. 249:219-27; Pack, P., and A. Plückthun 1992. Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric FV fragments with high avidity in *Escherichia coli* Biochemistry. 31:1579-84] para su inserción en C-HlyA. El dominio ZIP consta de una hélice anfipática que forma la cremallera de leucina del factor de transcripción de levaduras GCN4, flanqueado en su extremo N terminal por una zona bisagra peptídica derivada de la IgG3 del ratón, y en su extremo C terminal por un marcador de polihistidina (Sx-his). Se insertó un fragmento de ADN que codificaba para el dominio ZIP internamente cerca del extremo N terminal de la versión de aproximadamente 27 kDa marcada con E de C-HlyA (EHlyA) presente en el plásmido pEHlyA (Figura 1A). El plásmido resultante, pZEHlyA, codifica para un polipéptido de aproximadamente 33 kDa (denominado ZEHlyA) que contiene los dominios ZIP y C-HlyA marcados con E (Figuras 1A y 1B).

La producción de ZEHlyA y EHlyA, como control sin ZIP, se indujo en cultivos de células TolC+ de tipo salvaje de *E. coli* (por ejemplo, la cepa HB2151) que llevaban pVDL9.3, que codifica para HlyB y HlyD, y que alojaban pZEHlyA o pEHlyA, respectivamente. Como se muestra en la Figura 1C, se encontraron ambos polipéptidos en niveles similares (aproximadamente 10  $\mu$ g/ml) en los sobrenadantes de los cultivos de *E. coli* crecidos a 37°C tras la inducción de 4 h con IPTG 0,3 mM. Estas proteínas se detectaron en virtud del péptido marcado con E incorporado en sus secuencias con anticuerpo monoclonal anti-E marcado con POD (Métodos). La secreción de ambos polipéptidos derivados de HlyA fue específica y dependiente de la expresión de los componentes HlyB y HlyD por *E. coli* (datos no mostrados) [Fernández, L. A., and V. De Lorenzo 2001. Formation of disulphide bonds during secretion of proteins through the periplasmic-independent type I pathway Mol Microbiol. 40:332-46]. Este resultado indicaba que la presencia del dominio de dimerización ZIP no tenía efecto en la eficacia de la exportación de la señal de C-HlyA.

A continuación, se investigó el estado de oligomerización de los polipéptidos secretados. Se incubaron muestras de alícuotas que contenía los polipéptidos secretados EHlyA o ZEHlyA con el agente de entrecruzamiento bifuncional glutarato de disuccinimidilo (DSG) y luego se sometieron a SDS-PAGE desnaturizante y a inmunotransferencia con anticuerpo monoclonal anti-E marcado con POD (Métodos). En este experimento, sólo se entrecruzaron las muestras de ZEHlyA a baja concentración de DSG (40  $\mu$ M) para formar una banda de proteínas con una masa molecular (Mr) aparente de aproximadamente 66 kDa (Figura 2, carril 5), lo que estaba de acuerdo con el tamaño esperado para un dímero de ZEHlyA. La concentración más elevada de DSG (130  $\mu$ M) intensificó la intensidad de la banda correspondiente a ZEHlyA dímérico (Figura 2, carril 6) mientras que sólo tenía una reactividad menor sobre EHlyA control (Figura. 2, carril 3).

También se demostró la dimerización de ZEHlyA mediante cromatografía por exclusión de tamaño. Se separaron muestras de alícuotas de sobrenadantes de cultivos que contenían EHlyA o ZEHlyA secretados en una columna de filtración en gel, con un límite de exclusión de 1.500 kDa, junto con proteínas de masa conocida utilizadas como patrones (Métodos). Como se muestra en la Figura 3A, ZEHlyA mostró una Mr aparente de aproximadamente 66 kDa en la cromatografía de filtración en gel, mientras que EHlyA tuvo una Mr aparente de aproximadamente 32 kDa en las mismas condiciones. Es importante que se detectara un único pico para cada proteína (Figura 3B), lo que indica que ambos polipéptidos estaban presentes como monómeros (EHlyA) y dímeros (ZEHlyA) estables en disolución. Considerados juntos, estos resultados demostraron que podría obtenerse la dimerización de proteínas mediante la incorporación de hélices anfipáticas en C-HlyA sin interferir con su secreción por el transportador de Hly.

*Secreción de minianticuerpos díméricos por el sistema de la Hly de E. coli.* A la vista de los resultados previamente obtenidos se procedió a estudiar si los fragmentos de anticuerpos díméricos podían secretarse mediante el sistema de la Hly. En primer lugar, se construyó un plásmido, denominado pZEHlyA2-SD, para generar fusiones internas entre los fragmentos de anticuerpos recombinantes que carecen del dominio N-SP y ZEHlyA. Este plásmido es un derivado de pEHlyA2-SD [Fernández, L. A., L. Sola, L. Enjuanes, and V. de Lorenzo 2000. Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system Appl Environ Microbiol. 66:5024-5029] en el que se insertó un fragmento de ADN que codifica para el dominio ZIP en un único sitio *SalI*, entre la secuencia de poliunión y el dominio C-HlyA marcado con E (Métodos). Se seleccionaron dos anticuerpos de camello V<sub>HH</sub>, frente a  $\alpha$ -amilasa (*amy*) o la vacuna antitetánica (*ttx*), para determinar su

expresión como híbridos con los restos 'EHlyA y 'ZEHlyA (Figuras 4A y 4B). El empleo de anticuerpos de camello  $V_{HH}$  como pares de fusión se debió a su pequeño tamaño (aproximadamente 15 kDa) y a su baja tendencia a formar agregados de proteínas [Muyldermans, S. 2001. Single domain camel antibodies: current status J Biotechnol. 74:277-302; Plückthun, A., and P. Pack 1997. New protein engineering approaches to multivalent and bispecific antibody fragments Immunotechnology. 3:83-105], lo que podría interferir con el análisis de la dimerización obtenida por el dominio ZIP (véase Discusión). Las células de *E. coli* HB2151 (pVDL9.3) se transformaron con un plásmido que codificaba para el híbrido  $V_{HH}$ -HlyA ( $pV_{amy}$ HlyA,  $pV_{amy}$ ZHlyA,  $pV_{tx}$ VHlyA o  $pV_{tx}$ ZHlyA) y se indujeron 4 h por la adición de IPTG 0,3 mM a los cultivos líquidos crecidos en LB a 30 ó 37°C. Los polipéptidos secretados  $V_{amy}$ HlyA y  $V_{amy}$ ZHlyA se detectaron posteriormente en los sobrenadantes de los cultivos correspondientes de *E. coli* por Western blot con anticuerpo monoclonal anti-E marcado con POD (Figura 4C). En estas condiciones, la concentración final de  $V_{amy}$ HlyA y  $V_{amy}$ ZHlyA fue de aproximadamente 2  $\mu$ g/ml a 37°C, y se redujo aproximadamente en 2 veces en los cultivos que crecieron a 30°C. Resultados similares se obtuvieron con  $V_{tx}$ HlyA y  $V_{tx}$ ZHlyA (datos no mostrados).

El estado de oligomerización de los híbridos  $V_{HH}$ -HlyA secretados se sometió a ensayo mediante cromatografía de filtración en gel (Figura 4D). Muestras de alícuotas de sobrenadantes de cultivos de *E. coli* que contenían  $V_{amy}$ HlyA o  $V_{amy}$ ZHlyA se cargaron en una columna de filtración en gel (límite de exclusión & 1.500 kDa), junto con proteínas de masa conocida. A partir de sus perfiles de elución puede deducirse (Figura 4D) que el híbrido  $V_{amy}$ HlyA tenía una  $M_r$  aparente de aproximadamente 40 kDa, lo que estaba completamente de acuerdo con la masa esperada para un monómero de este polipéptido. Por el contrario,  $V_{amy}$ ZHlyA mostró una  $M_r$  aparente de aproximadamente 95 kDa, que es aproximadamente el doble de la masa esperada de su monómero (es decir, 47 kDa). Hay que mencionar que la temperatura a la que se indujeron los cultivos de *E. coli* (30°C ó 37°C) no tuvo ninguna influencia en el comportamiento cromatográfico de estas muestras. Por lo tanto, la secreción de los anticuerpos de camello  $V_{HH}$  monoméricos o diméricos pueden producirse fusionándolos a los restos EHlyA o ZEHlyA, respectivamente.

A continuación, se probó si la dimerización mejoraba las propiedades de unión funcional de  $V_{amy}$ ZHlyA. Para este fin, se comparó la unión a la  $\alpha$ -amilasa de  $V_{amy}$ HlyA monomérico y de  $V_{amy}$ ZHlyA dimérico mediante ELISA. En estos experimentos, se incubaron diluciones seriadas de sobrenadantes de cultivos de *E. coli* que contenían cantidades idénticas de  $V_{amy}$ HlyA o  $V_{amy}$ ZHlyA con placas de ELISA recubiertas con  $\alpha$ -amilasa u ovoalbúmina (como antígeno control). Tras el lavado, los minianticuerpos unidos se detectaron con el conjugado anticuerpo monoclonal anti-E marcado con POD y se realizó la lectura a la  $DO_{490\text{ nm}}$  (Métodos). La unión específica de la  $\alpha$ -amilasa se demostró incubando estas placas con  $V_{tx}$ HlyA y  $V_{tx}$ ZHlyA. En la Figura 5 se muestra el resultado de un ELISA prototipo de estos experimentos. La unión previa a ovoalbúmina (en todos los casos la  $DO_{490\text{ nm}} \leq 0,05$ ) se ha sustraído de los valores presentados. Tal como se indica, la molécula  $V_{amy}$ ZHlyA dimérica presentó una afinidad funcional mayor frente a la  $\alpha$ -amilasa, que  $V_{amy}$ HlyA monomérica. No se observó unión a  $\alpha$ -amilasa con los derivados control de  $V_{tx}$  (Figura 5) ni con los polipéptidos EHlyA y ZEHlyA (datos no mostrados). En general, al menos se requirió una concentración diez veces mayor de  $V_{amy}$ HlyA monomérico para lograr señales de unión a la  $\alpha$ -amilasa similares a las obtenidas con  $V_{amy}$ ZHlyA. Además, en la concentración de saturación de ambos minianticuerpos (aproximadamente 0,5  $\mu$ g/ml) la unión obtenida con  $V_{amy}$ ZHlyA alcanzó un nivel meseta superior. Por tanto, la dimerización de  $V_{amy}$ ZHlyA induce un efecto de aidez sobre este minianticuerpo que se refleja en una mayor afinidad de unión funcional a su antígeno.

### 3. Discusión

La dimerización es una propiedad que con frecuencia se desea conseguir por ingeniería genética en las proteínas cuando está implicada una actividad de unión (por ejemplo, en las interacciones proteína-ADN o antígeno-anticuerpo), puesto que puede intensificar su afinidad funcional (aidez).

Este ejemplo ilustra la obtención por ingeniería genética, por primera vez, de la dimerización de las proteínas secretadas por el sistema de transporte de hemolisina de *E. coli* y se ha empleado esa tecnología para producir minianticuerpos de alta aidez derivados de los anticuerpos de camello  $V_{HH}$ .

Los resultados obtenidos demuestran que la incorporación de una hélice anfipática de autodimerización en el extremo N terminal de C-HlyA no interfiere con la secreción de Hly y permite la dimerización del polipéptido secretado. Tal como se muestra, la dimerización puede intensificar la aidez de la unión del polipéptido secretado derivado de C-HlyA. Además, también puede tener otras aplicaciones, como la asociación molecular de varios antígenos y/o adyuvantes producidos por cepas bacterianas vivas, o la combinación de diversas actividades biológicas para la generación de moléculas biespecíficas (por ejemplo, la unión a un antígeno y el reclutamiento del complemento).

El scFvs dimérico de alta aidez se ha producido en el periplasma de las células de *E. coli* insertando hélices anfipáticas en su extremo C terminal. Inicialmente parecía sencillo probar si scFv dimérico podía secretarse fusionándolos a 'ZEHlyA. Sin embargo, la tendencia reconocida de scFvs a formar dímeros y agregados de proteínas de elevado peso molecular indujo a utilizar fragmentos más pequeños de anticuerpos. De hecho, cuando se investigó el estado de oligomerización de un híbrido scFv-HlyA (6AC3HlyA) mediante cromatografía de filtración en gel, se encontró que forma un número discreto de especies moleculares que difieren en su masa aparente [es decir, a 30°C aproximadamente el 50% tenía la masa del monómero,  $M_r$  55 kDa aproximadamente, mientras que el resto tenía una  $M_r$  aparente de aproximadamente 115, aproximadamente 170 e igual o superior a 350 kDa (datos no publicados)].

## ES 2 296 423 B1

Los anticuerpos de camello  $V_{HH}$  han recibido mucha atención debido a su mejor solubilidad y a su estructura más simple, lo que facilita su amplificación y clonaje. Merece la pena destacar que los cambios efectuados no disminuyen la afinidad ni la especificidad de los anticuerpos de camello  $V_{HH}$ , debido a la presencia de regiones determinantes de la complementariedad (CDR) extremadamente variables que compensan la pérdida de diversidad provocada por la ausencia de un dominio  $V_L$ . Los anticuerpos de camello también han demostrado un potencial extraordinario como inhibidores enzimáticos puesto que sus grandes CDR pueden alcanzar los sitios activos escondidos en las enzimas. Además, la similitud entre los anticuerpos de camello  $V_{HH}$  y las secuencias de la familia VH3 humana está permitiendo la generación de librerías presentadas en fagos de los dominios  $V_H$  humanos camelizados y de los anticuerpos de camello  $V_{HH}$  humanizados.

Los beneficios expuestos anteriormente motivaron a los inventores a utilizar los dominios  $V_{HH}$  para su secreción por el translocador de hemolisina de *E. coli*. Los resultados obtenidos muestran que los anticuerpos de camello funcionales, tanto en la forma monomérica como en la dimérica, pueden recuperarse de los sobrenadantes del cultivo de *E. coli* a niveles similares a los obtenidos en su expresión periplásmica (aproximadamente 1 mg/litro de cultivo a D.O.<sub>600 nm</sub> = 1). Además, la dimerización provocada por ZEHlyA indujo un aumento de diez veces en la afinidad funcional de  $V_{amy}$ . Este valor está dentro del intervalo esperado producido por el cambio de anticuerpos monovalentes a divalentes. En conclusión, estos datos demuestran que el sistema de secreción de Hly puede emplearse para la secreción de polipéptidos y minianticuerpos diméricos de alta avidéz. La simplicidad de esta tecnología puede ser extremadamente útil para la selección de alto rendimiento de las librerías de anticuerpos.

## REIVINDICACIONES

1. Una construcción de ADN **caracterizada** porque comprende:

- 5 a) una *primera secuencia de ácido nucleico* que contiene la secuencia de nucleótidos que codifica para un péptido o proteína susceptible de ser expresado de forma recombinante;
- 10 b) una *segunda secuencia de ácido nucleico* que cor tiene la secuencia de nucleótidos que codifica para un dominio de dimerización; y
- 15 c) una *tercera secuencia de ácido nucleico* que contiene la secuencia de nucleótidos que codifica para la  $\alpha$ -hemolisina (HlyA) de *E. coli* o para un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*, o una secuencia de nucleótidos que codifica para un gen homólogo, o una secuencia de nucleótidos que codifica para una variante, natural o artificial, de HlyA o de un fragmento de la misma que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*;

20 en donde el extremo 3' de dicha primera secuencia de ácido nucleico está unido al extremo 5' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha segunda secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico.

2. Construcción de ADN según la reivindicación 1, **caracterizada** porque dicho péptido o proteína susceptible de ser expresado de forma recombinante se selecciona entre enzimas, inhibidores enzimáticos, hormonas, moléculas implicadas en la adhesión y/o señalización celular y compuesta por dominios, antígenos inmuiogénicos, agentes terapéuticos, y moléculas inmunoreguladoras.

3. Construcción de ADN según la reivindicación 2, **caracterizada** porque dicho antígeno inmunogénico o molécula inmunoreguladora se selecciona entre antígenos específicos de tumor, antígenos de enfermedades auto-inmune, factores de crecimiento, citoquinas, interleuquinas, interferones y minianticuerpos.

4. Construcción de ADN según la reivindicación 1, **caracterizada** porque dicho dominio de dimerización comprende una hélice peptídica o una estructura de doble arrollamiento helicoidal (coiled coil).

5. Construcción de ADN según la reivindicación 4, **caracterizada** porque dicho dominio de dimerización comprende la cremallera de leucina del factor de transcripción GCN4 de levadura.

6. Construcción de ADN según la reivindicación 1, **caracterizada** porque dicha tercera secuencia de ácido nucleico comprende la secuencia de nucleótidos que codifica para la  $\alpha$ -hemolisina (HlyA) de *E. coli* o para un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*, o una secuencia de nucleótidos que codifica para un gen homólogo, o una secuencia de nucleótidos que codifica para una variante, natural o artificial, de HlyA o de un fragmento de la misma que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*.

7. Construcción de ADN según la reivindicación 6, **caracterizada** porque dicha tercera secuencia de ácido nucleico se selecciona entre:

- 45 a) la secuencia de nucleótidos que codifica para la HlyA de *E. coli*;
- 50 b) una secuencia de ácido nucleico que comprende la secuencia de nucleótidos que codifica para los 60 últimos aminoácidos del extremo C-terminal de HlyA de *E. coli*;
- c) una secuencia de ácido nucleico constituida por una secuencia de nucleótidos que codifica para los 60 últimos aminoácidos del extremo C-terminal de HlyA de *E. coli*;
- 55 d) la secuencia de nucleótidos identificada como SEQ ID NO: 1; y
- e) una secuencia de nucleótidos que codifica para la secuencia de aminoácidos identificada como SEQ ID NO: 2.

8. Construcción de ADN según la reivindicación 1, **caracterizada** porque comprende, además, una *cuarta secuencia de ácido nucleico* que codifica para un péptido espaciador situada entre dichas primera y segunda secuencias de ácido nucleico, en donde el extremo 5' de dicha cuarta secuencia de ácido nucleico está unido al extremo 3' de dicha primera secuencia de ácido nucleico y el extremo 3' de dicha cuarta secuencia de ácido nucleico está unido al extremo 5' de dicha segunda secuencia de ácido nucleico.

9. Construcción de ADN según la reivindicación 8, **caracterizada** porque dicho péptido espaciador comprende repeticiones de restos de aminoácidos, preferentemente, repeticiones Gly-Gly-Gly-Ser.

## ES 2 296 423 B1

10. Construcción de ADN según la reivindicación 8, **caracterizada** porque dicho péptido espaciador es una región bisagra de un anticuerpo.

5 11. Construcción de ADN según la reivindicación 1, **caracterizada** porque comprende, además, una *quinta secuencia de ácido nucleico* que codifica para un péptido susceptible de ser utilizado con fines de aislamiento o purificación.

10 12. Construcción de ADN según la reivindicación 11, **caracterizada** porque dicho péptido susceptible de ser utilizado con fines de aislamiento o purificación comprende una secuencia de polihistidina o una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad.

15 13. Construcción de ADN según la reivindicación 11, **caracterizada** porque dicha quinta secuencia de ácido nucleico está situada entre dichas segunda y tercera secuencias de ácido nucleico, en donde el extremo 5' de dicha quinta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha quinta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico.

14. Construcción de ADN según la reivindicación 1, **caracterizada** porque comprende, además, una *sexta secuencia de ácido nucleico* que codifica para un péptido susceptible de ser utilizado con fines de reconocimiento.

20 15. Construcción de ADN según la reivindicación 14, **caracterizada** porque dicho péptido susceptible de ser utilizado con fines de reconocimiento comprende una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para reconocer la proteína de fusión resultante por cromatografía de inmunoafinidad.

25 16. Construcción de ADN según la reivindicación 15, **caracterizada** porque dicho péptido susceptible de ser utilizado con fines de reconocimiento comprende la secuencia del epítipo E.

30 17. Construcción de ADN según la reivindicación 14, **caracterizada** porque dicha sexta secuencia de ácido nucleico está situada entre dichas segunda y tercera secuencias de ácido nucleico, en donde el extremo 5' de dicha sexta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico.

18. Construcción de ADN según la reivindicación 11 ó 14, **caracterizada** porque dichas quinta y sexta secuencias de ácido nucleico están separadas entre sí.

35 19. Construcción de ADN según la reivindicación 11 ó 14, **caracterizada** porque dichas quinta y sexta secuencias de ácido nucleico están unidas entre sí.

40 20. Construcción de ADN según la reivindicación 19, **caracterizada** porque dicha sexta secuencia de ácido nucleico está situada entre dichas tercera y quinta secuencias de ácido nucleico, en donde el extremo 5' de dicha sexta secuencia de ácido nucleico está unido al extremo 3' de dicha quinta secuencia de ácido nucleico y el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico.

45 21. Construcción de ADN según la reivindicación 19, **caracterizada** porque dicha sexta secuencia de ácido nucleico está situada entre dichas segunda y quinta secuencias de ácido nucleico, en donde el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo 5' de dicha quinta secuencia de ácido nucleico y el extremo 5' de dicha sexta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico.

50 22. Construcción de ADN según la reivindicación 1, **caracterizada** porque comprende, además, una *séptima secuencia de ácido nucleico* que comprende una secuencia de nucleótidos que codifica para una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos.

23. Construcción de ADN según la reivindicación 22, **caracterizada** porque dicha séptima secuencia de ácido nucleico comprende una secuencia de nucleótidos que codifica para un sitio de reconocimiento de una proteasa.

55 24. Construcción de ADN según la reivindicación 23, **caracterizada** porque dicha proteasa se selecciona entre una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa y Factor de coagulación Xa.

60 25. Construcción de ADN según la reivindicación 22, **caracterizada** porque dicha séptima secuencia de ácido nucleico comprende una secuencia de nucleótidos que codifica para un sitio susceptible de ser específicamente escindido por un reactivo químico.

26. Construcción de ADN según la reivindicación 25, **caracterizada** porque dicho reactivo químico es bromuro de cianógeno.

65 27. Construcción de ADN según la reivindicación 22, **caracterizada** porque dicha séptima secuencia de ácido nucleico se encuentra a continuación del extremo 3' de dicha segunda secuencia de ácido nucleico.

## ES 2 296 423 B1

28. Un cassette de expresión **caracterizado** porque comprende una construcción de ADN según cualquiera de las reivindicaciones 1 a 27 operativamente unida a una secuencia de control de expresión.
29. Cassette de expresión según la reivindicación 28, **caracterizado** porque dicha secuencia de control de expresión comprende una secuencia promotora, una secuencia codificante para reguladores transcripcionales, una secuencia de unión a ribosomas (RBS) y/o una secuencia terminadora de transcripción.
30. Cassette de expresión según la reivindicación 28, **caracterizado** porque comprende, además, un marcador.
31. Cassette de expresión según la reivindicación 30, **caracterizado** porque dicho marcador comprende un gen de resistencia a antibióticos o un gen de resistencia a compuestos tóxicos.
32. Cassette de expresión según la reivindicación 28, **caracterizado** porque dicha secuencia de control de expresión es funcional en bacterias.
33. Una bacteria Gram negativa **caracterizada** porque comprende una construcción de ADN según cualquiera de las reivindicaciones 1 a 27 o un cassette de expresión según cualquiera de las reivindicaciones 28 a 32.
34. Bacteria según la reivindicación 33, **caracterizada** porque se selecciona entre *Escherichia coli*, *Salmonella typhimurium*, *Pseudomonas aeruginosa* y *Pseudomonas putida*.
35. Un método para producir un péptido o proteína susceptible de ser expresado de forma recombinante en forma de una proteína de fusión dimérica, **caracterizado** porque comprende crecer una bacteria según la reivindicación 33 ó 34 bajo condiciones que permiten la producción y excreción al medio de cultivo de dicha proteína de fusión dimérica.
36. Método según la reivindicación 35, **caracterizado** porque comprende, además, el aislamiento y purificación de dicha proteína de fusión dimérica.
37. Método según la reivindicación 35 ó 36, **caracterizado** porque el péptido o proteína susceptible de ser expresado de forma recombinante de dicha proteína de fusión dimérica se selecciona entre enzimas, inhibidores enzimáticos, hormonas, moléculas implicadas en la adhesión y/o señalización celular y compuesta por dominios, antígenos inmunogénicos, agentes terapéuticos, y moléculas inmunoreguladoras.
38. Método según la reivindicación 37, **caracterizada** porque el antígeno inmunogénico o molécula inmunoreguladora del péptido o proteína susceptible de ser expresado de forma recombinante de dicha proteína de fusión dimérica se selecciona entre antígenos específicos de tumor, antígenos de enfermedades auto-inmune, factores de crecimiento, citoquinas, interleuquinas, interferones y minianticuerpos.
39. Una proteína de fusión dimérica **caracterizada** porque se obtiene por expresión de la secuencia de ácido nucleico contenida en una construcción de ADN según cualquiera de las reivindicaciones 1 a 27.
40. Proteína de fusión dimérica según la reivindicación 39, **caracterizada** porque cada monómero comprende:
- (i) la secuencia de aminoácidos de la  $\alpha$ -hemolisina (HlyA) de *Escherichia coli* o de un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de hemolisina (Hly) de *E. coli*;
  - (ii) una secuencia de aminoácidos correspondiente a un dominio de dimerización; y
  - (iii) la secuencia de aminoácidos de un péptido o proteína susceptible de ser expresado de forma recombinante.
41. Proteína de fusión según la reivindicación 40, **caracterizada** porque cada monómero comprende:
- (i) la secuencia completa de aminoácidos de HlyA de *E. coli*, o un fragmento de HlyA de *E. coli* que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*;
  - (ii) un dominio de dimerización, seleccionado entre una hélice peptídica y una estructura de doble arrollamiento helicoidal (coiled coil); y
  - (iii) un péptido o proteína susceptible de ser expresado de forma recombinante seleccionado entre una enzima, un inhibidor enzimático, una hormona, una molécula implicada en la adhesión y/o señalización celular y compuesta por dominios, un antígeno inmunogénico, un agente terapéutico, una molécula inmunoreguladora, y un minianticuerpo.
42. Proteína de fusión según la reivindicación 40 ó 41, **caracterizada** porque cada monómero comprende, además (a) un péptido espaciador entre el péptido o proteína susceptible de ser expresado de forma recombinante y el dominio de dimerización; y/o (b) un péptido para facilitar el aislamiento o purificación del péptido o proteína de fusión; y/o

## ES 2 296 423 B1

(c) un péptido que permita el reconocimiento del péptido o proteína de fusión; y/o (d) una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos.

43. Proteína de fusión según la reivindicación 42, **caracterizada** porque cada monómero comprende, (a) un péptido que contiene repeticiones de Gly-Gly-Gly-Ser o la región bisagra de un anticuerpo; y/o (b) una secuencia de polihistidina o una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad; y/o (c) una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para reconocer la proteína de fusión resultante por cromatografía de inmunoafinidad; y/o (d) una de aminoácidos que constituye un sitio de reconocimiento d una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa o Factor de coagulación Xa o una secuencia de aminoácidos susceptible de ser específicamente escindida por un reactivo químico.

15

20

25

30

35

40

45

50

55

60

65

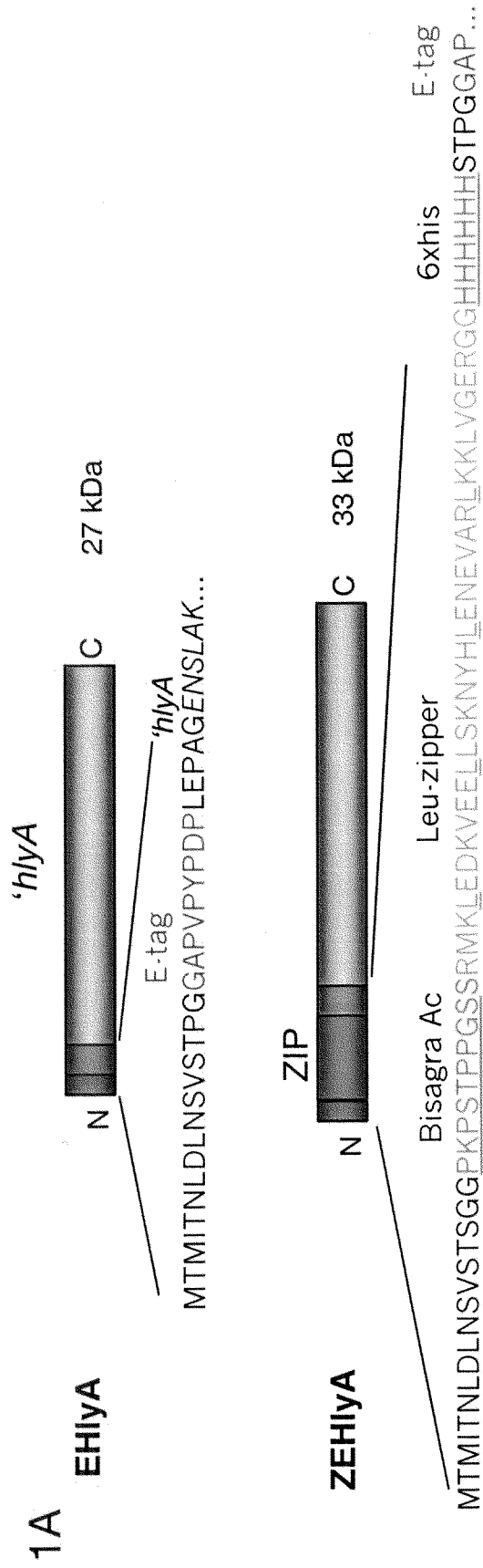


Fig. 1 (cont.)

1B

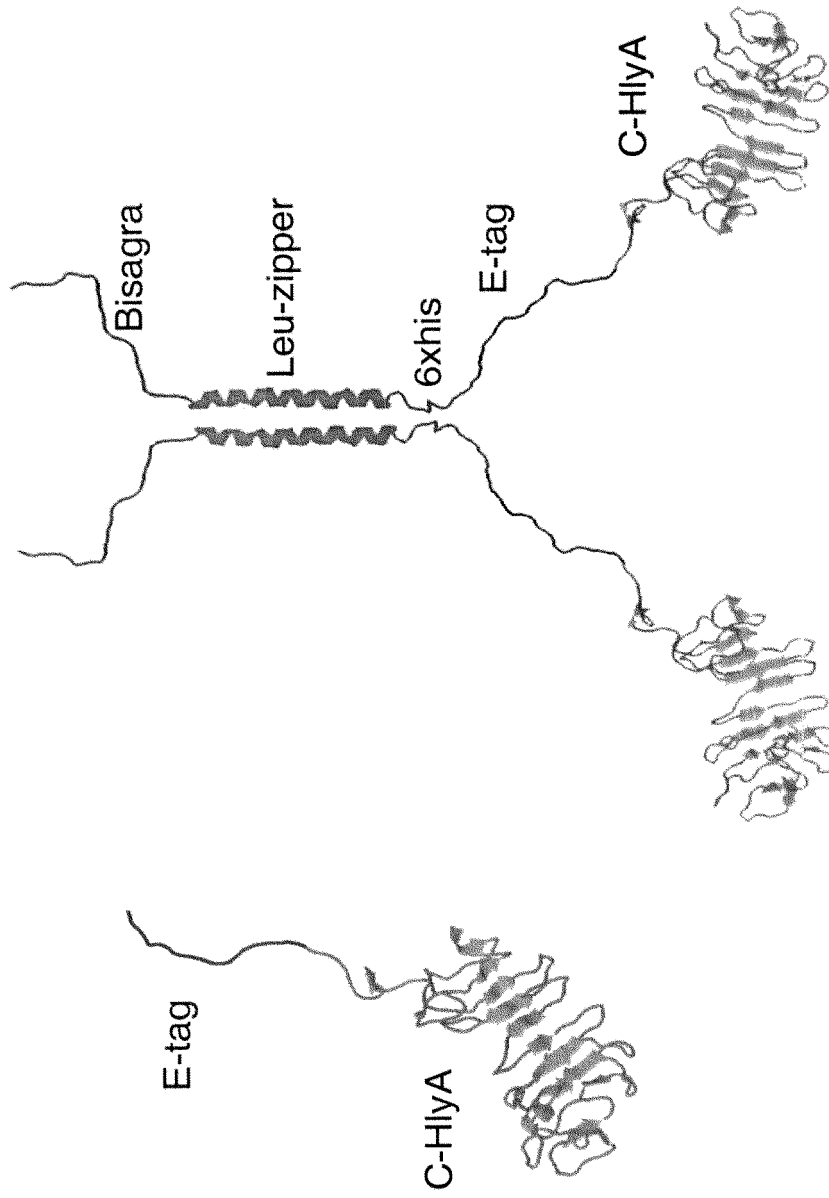


Fig. 1 (cont.)

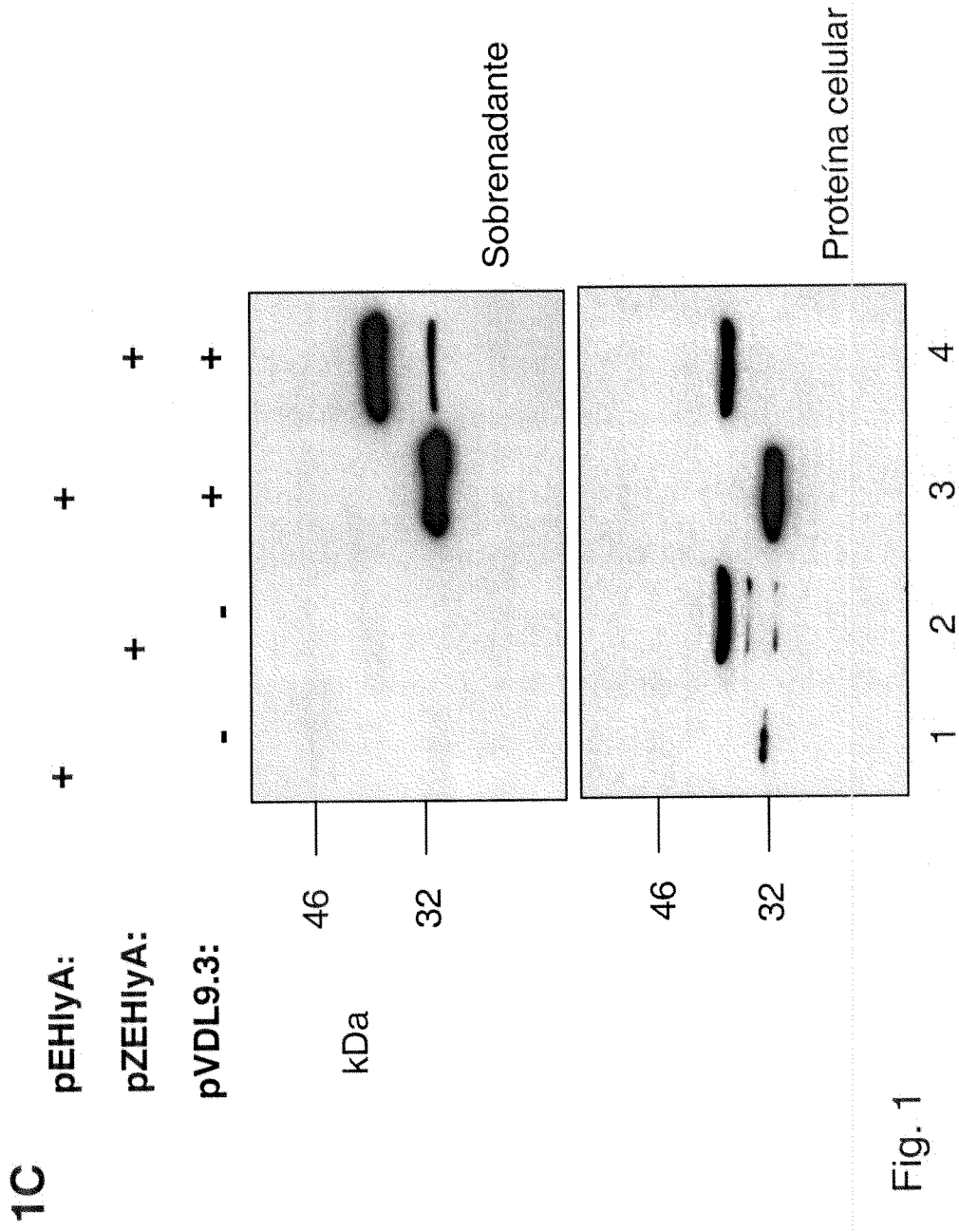


Fig. 1

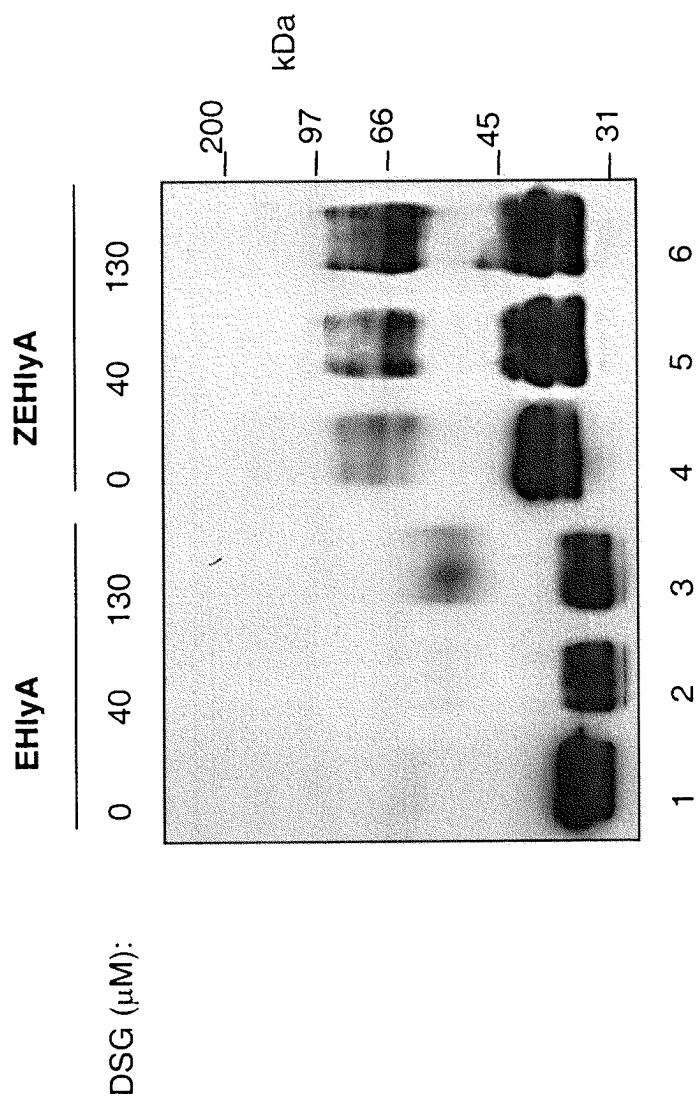


Fig. 2

3A

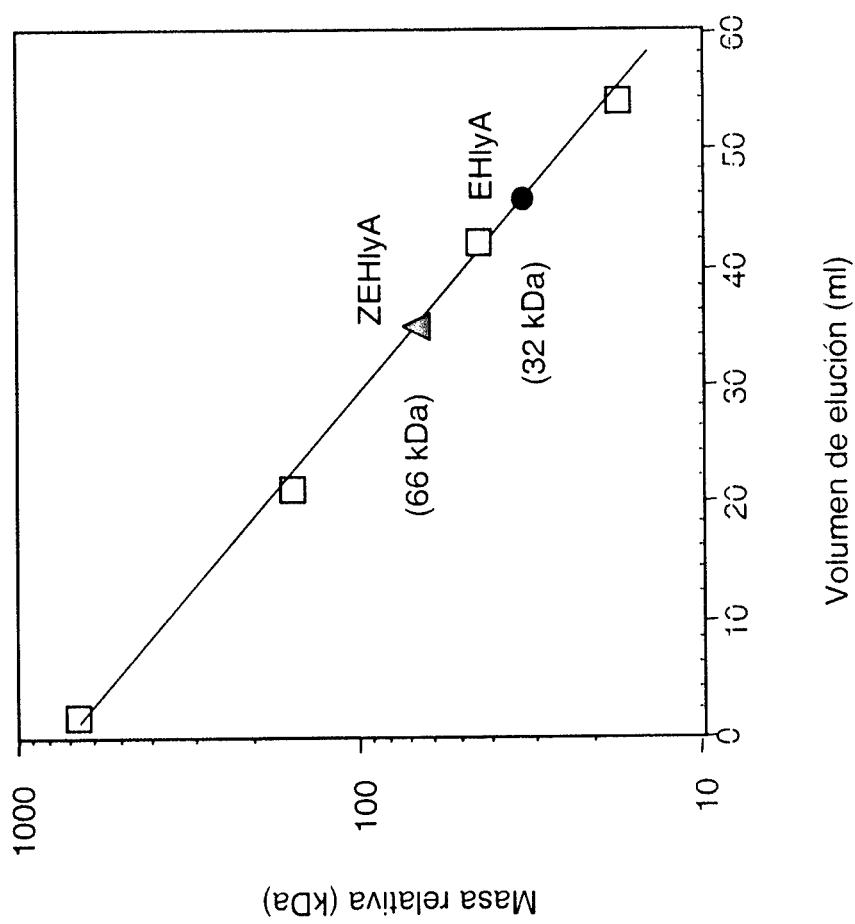
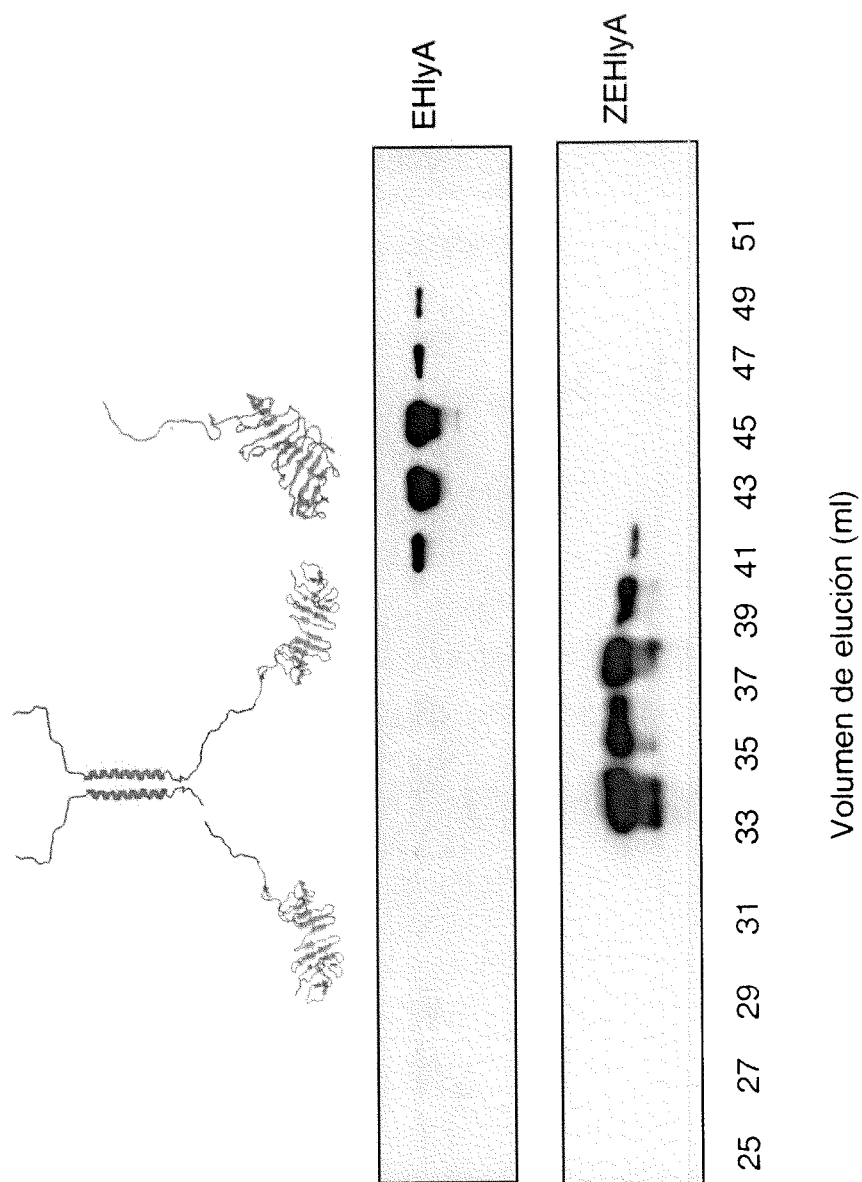


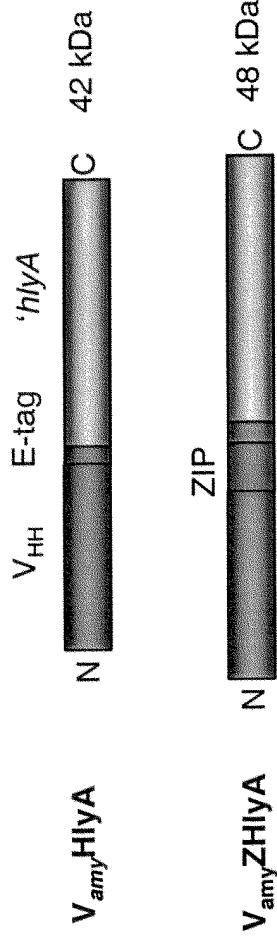
Fig. 3 (cont.)

3B

Fig. 3



4A



4B

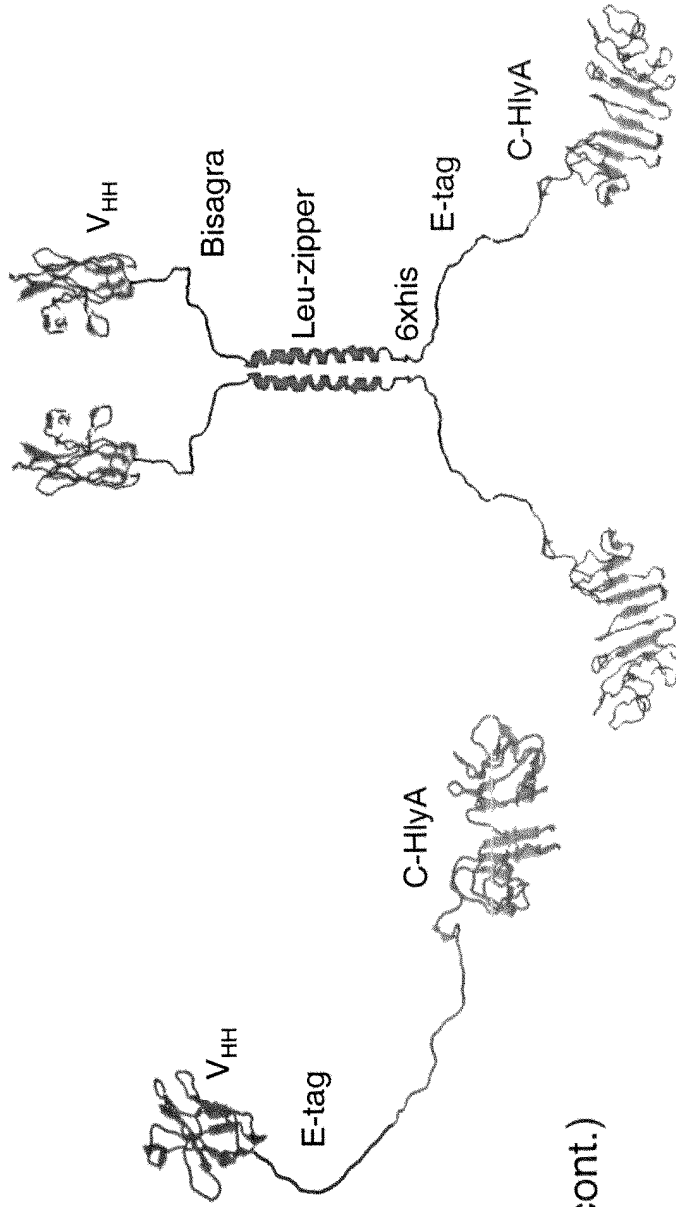
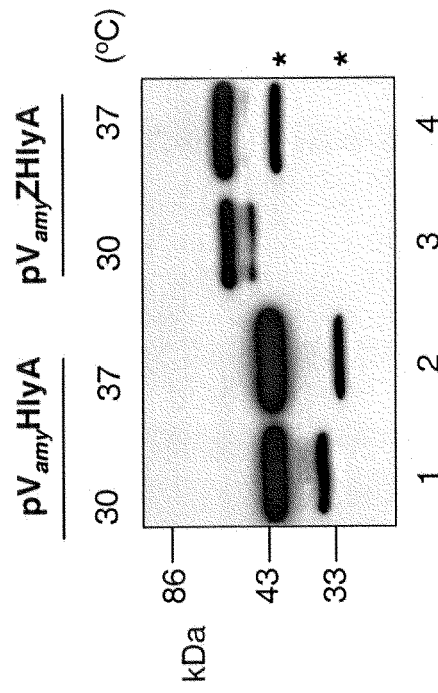


Fig. 4 (cont.)

4C



4D

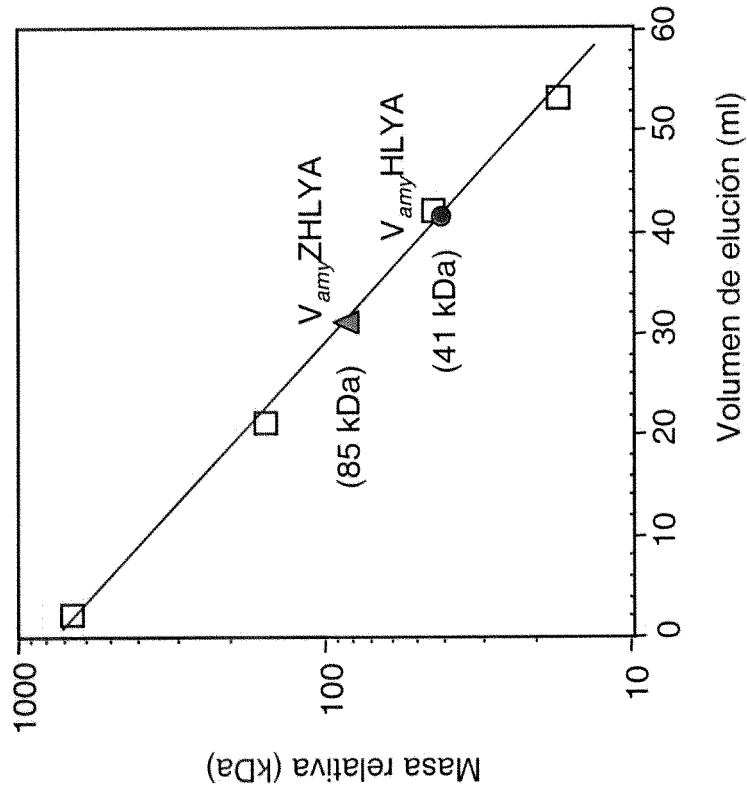
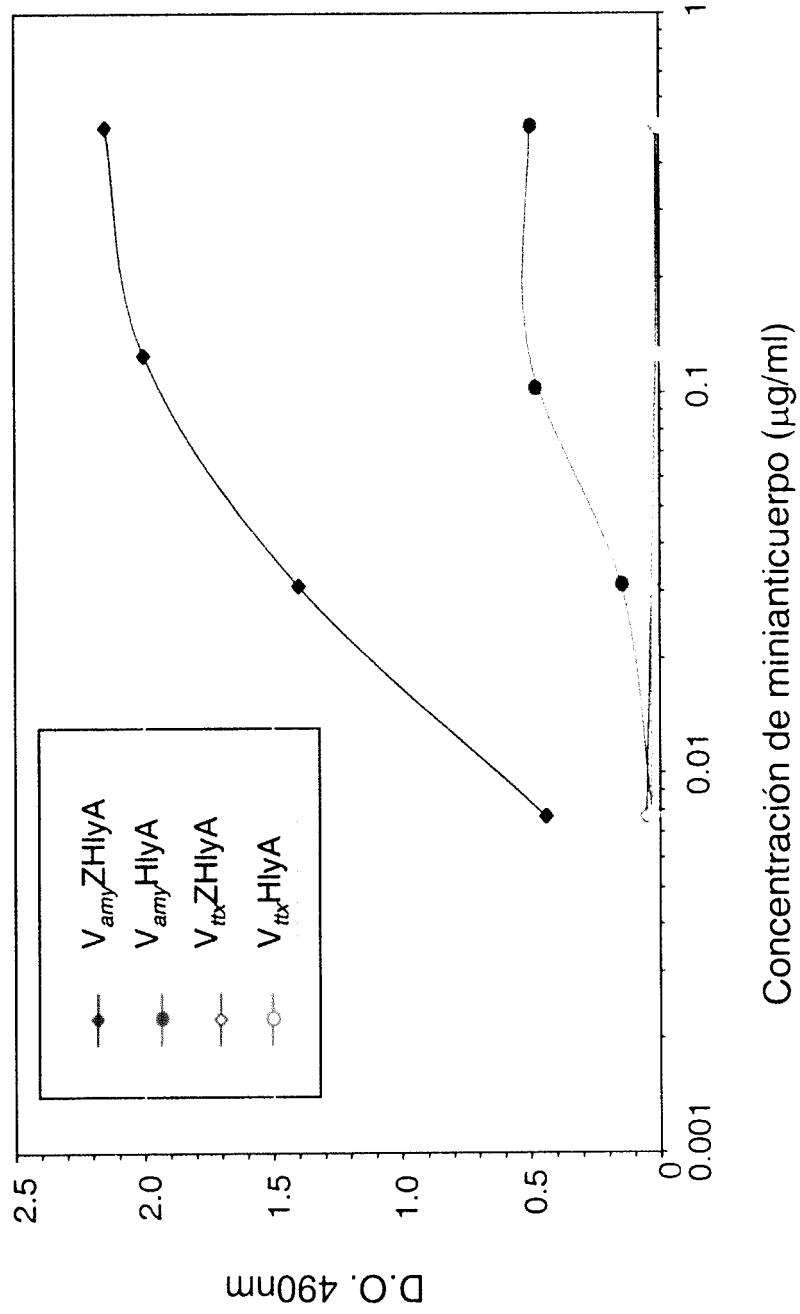


Fig. 4

Fig. 5





ES 2 296 423 B1

```

GGTAATGACAAGTTGTACGGCAGTGAGGGAGCAGACCTGCTTGATGGCGGAGAAGGGAAT
301 -----+-----+-----+-----+-----+ 360
CCATTACTGTTCAACATGCCGTCCTCCCTCGTCTGGACGAACTACCGCCTCTTCCCTTA
a      G N D K L Y G S E G A D L L D G G E G N -

GATCTTCTGAAAGGTGGATATGGTAATGATATTTATCGTTATCTTTCAGGATATGGCCAT
361 -----+-----+-----+-----+-----+ 420
CTAGAAGACTTTCACCTATAACCATTACTATAAATAGCAATAGAAAGTCCTATACCGGTA
a      D L L K G G Y G N D I Y R Y L S G Y G H -

CATATTATTGACGATGAAGGGGGGAAAGACGATAAACTCAGTTTAGCTGATATAGATTTTC
421 -----+-----+-----+-----+-----+ 480
GTATAATAACTGCTACTTCCCCCTTCTGCTATTTGAGTCAAATCGACTATATCTAAAG
a      H I I D D E G G K D D K L S L A D I D F -

CGGGACGTTGCCTTTAAGCGAGAAGGGAATGACCTCATTATGTATAAAGCTGAAGGTAAT
481 -----+-----+-----+-----+-----+ 540
GCCCTGCAACGGAAATTCGCTCTCCCTTACTGGAGTAATACATATTTTCGACTTCCATTA
a      R D V A F K R E G N D L I M Y K A E G N -

GTTCTTTCTATTGGCCACAAAAATGGTATTACATTTAAAAACTGGTTTGAAAAGAGTCA
541 -----+-----+-----+-----+-----+ 600
CAAGAAAGATAACCGGTGTTTTTACCATAATGTAAATTTTGGACCAAACCTTTTCTCAGT
a      V L S I G H K N G I T F K N W F E K E S -

GATGATCTCTCTAATCATCAGATAGAGCAGATTTTTGATAAAGACGGCAGGGTAATCACA
601 -----+-----+-----+-----+-----+ 660
CTACTAGAGAGATTAGTAGTCTATCTCGTCTAAAACTATTTCTGCCGTCCCATTAGTGT
a      D D L S N H Q I E Q I F D K D G R V I T -

CCAGATTCTCTTAAAAAGCATTGGAATATCAGCAGAGTAATAACAAGGTAAGTTATGTG
661 -----+-----+-----+-----+-----+ 720
GGTCTAAGAGAATTTTTTCGTAAACTTATAGTCGTCTCATTATTGTTCCATTCAATACAC
a      P D S L K K A F E Y Q Q S N N K V S Y V -

TATGGACATGATGCATCAACTTATGGGAGCCAGGACAATCTTAATCCATTAATTAATGAA
721 -----+-----+-----+-----+-----+ 780
ATACCTGTACTACGTAGTTGAATACCCTCGGTCCTGTTAGAATTAGGTAATTAATTACTT
a      Y G H D A S T Y G S Q D N L N P L I N E -

```

Fig. 6 (cont.)

ES 2 296 423 B1

```

                PstI
                |
ATCAGCAAAATCATTTCAGCTGCAGGTAACCTCGATGTTAAGGAGGAAAGATCTGCCGCT
781 -----+-----+-----+-----+-----+-----+-----+ 840
TAGTCGTTTTAGTAAAGTCGACGTCCATTGAAGCTACAATTCCTCCTTTCTAGACGGCGA
a      I S K I I S A A G N F D V K E E R S A A -

                NdeI
                |
TCTTTATTGCAGTTGTCCGGTAATGCCAGTGATTTTTCATATGGACGGAACCTCAATAACT
841 -----+-----+-----+-----+-----+-----+-----+ 900
AGAAATAACGTCAACAGGCCATTACGGTCACTAAAAAGTATACCTGCCTTGAATTATTGA
a      S L L Q L S G N A S D F S Y G R N S I T -

                TTGACAGCATCAGCATAA
901 -----+----- 918
                AACTGTCGTAGTCGTATT
a      L T A S A * -

```

Enzimas que cortan:

AccI    BamHI    EcoRI    HindIII    NdeI    PstI    SacI    SalI  
SmaI    XbaI

Enzimsa que no cortan:

KpnI    NcoI    NheI    NotI    SacII    SphI

Fig. 6



# ES 2 296 423 B1

```

b      H H H H H A S T P G G A P V P Y P D P L -
      6xhis tag                               |=> E-tag
      EagI
      |
361    GGAACCGGCCGGGGAAAATTCTCTTGCTAAAAATGTATTATCCGGTGGAAGGTAATGA
      -----+-----+-----+-----+-----+-----+-----+ 420
      CCTTGGCCGGCCCTTTTAAGAGAACGATTTTTACATAATAGGCCACCTTTTCCATTACT

b      E P A G E N S L A K N V L S G G K G N D -
      |=> C-hlyA
      CAAGTTGTACGGCAGTGAGGGAGCAGACCTGCTTGATGGCGGAGAAGGGAATGATCTTCT
421    -----+-----+-----+-----+-----+-----+-----+ 480
      GTTCAACATGCCGTCACCTCCCTCGTCTGGACGAACTACCGCCTCTTCCCTTACTAGAAGA

b      K L Y G S E G A D L L D G G E G N D L L -
      GAAAGGTGGATATGGTAATGATATTTATCGTTATCTTTCAGGATATGGCCATCATATTAT
481    -----+-----+-----+-----+-----+-----+-----+ 540
      CTTTCCACCTATAACCATTACTATAAATAGCAATAGAAAGTCCTATACCGGTAGTATAATA

b      K G G Y G N D I Y R Y L S G Y G H H I I -
      TGACGATGAAGGGGGAAAGACGATAAACTCAGTTTAGCTGATATAGATTTCCGGGACGT
541    -----+-----+-----+-----+-----+-----+-----+ 600
      ACTGCTACTTCCCCCTTTCTGCTATTTGAGTCAAATCGACTATATCTAAAGGCCCTGCA

b      D D E G G K D D K L S L A D I D F R D V -
      TGCCTTTAAGCGAGAAGGGAATGACCTCATTATGTATAAAGCTGAAGGTAATGTTCTTTC
601    -----+-----+-----+-----+-----+-----+-----+ 660
      ACGGAAATTCGCTCTTCCCTTACTGGAGTAATACATATTTGACTTCCATTACAAGAAAG

b      A F K R E G N D L I M Y K A E G N V L S -
      TATTGGCCACAAAAATGGTATTACATTTAAAAACTGGTTTGAAAAAGAGTCAGATGATCT
661    -----+-----+-----+-----+-----+-----+-----+ 720
      ATAACCGGTGTTTTTACCATAATGTAATTTTTGACCAAACCTTTTCTCAGTCTACTAGA

b      I G H K N G I T F K N W F E K E S D D L -
      CTCTAATCATCAGATAGAGCAGATTTTTGATAAAGACGGCAGGGTAATCACACCAGATTC
721    -----+-----+-----+-----+-----+-----+-----+ 780
      GAGATTAGTAGTCTATCTCGTCTAAAACTATTTCTGCCGTCCCATTAGTGTGGTCTAAG

b      S N H Q I E Q I F D K D G R V I T P D S -
      TCTTAAAAAAGCATTGGAATATCAGCAGAGTAATAACAAGGTAAGTTATGTGTATGGACA
781    -----+-----+-----+-----+-----+-----+-----+ 840
      AGAATTTTTTCGTAAACTTATAGTCGTCTCATTATTGTTCCATTCAATACACATACCTGT

```

**Fig. 7 (cont.)**

ES 2 296 423 B1

```

b      L K K A F E Y Q Q S N N K V S Y V W G H -
      TGATGCATCAACTTATGGGAGCCAGGACAATCTTAATCCATTAATTAATGAAATCAGCAA
841  -----+-----+-----+-----+-----+-----+-----+ 900
      ACTACGTAGTTGAATACCCTCGGTCCTGTTAGAATTAGGTAATTAATTACTTTAGTCGTT

b      D A S T Y G S Q D N L N P L I N E I S K -
                                     BglII
                                     |
      AATCATTTTCAGCTGCAGGTAACCTTCGATGTTAAGGAGGAAAAGATCTGCCGCTTCTTTATT
901  -----+-----+-----+-----+-----+-----+-----+ 960
      TTAGTAAAGTCGACGTCCATTGAAGCTACAATTCCTCCTTTCTAGACGGCGAAGAAATAA

b      I I S A A G N F D V K E E R S A A S L L -
      GCAGTTGTCCGGTAATGCCAGTGATTTTTTCATATGGACGGAACTCAATAACTTTGACAGC
961  -----+-----+-----+-----+-----+-----+-----+ 1020
      CGTCAACAGGCCATTACGGTCACTAAAAAGTATACCTGCCTTGAGTTATTGAAACTGTCC

b      Q L S G N A S D F S Y G R N S I T L T A -
      ATCAGCATAATATATTAATTTAAATGATAGCAATCTTACTGGGCTGTGCCACATAAGATT
1021 -----+-----+-----+-----+-----+-----+-----+ 1080
      TAGTCGTATTATATAATTAATTTACTATCGTTAGAATGACCCGACACGGTGTATTCTAA

b      S A * -
      GCTATTTTTTTGGAGTCATAATGGATTCTTGTCATAAAATTGATTATGGGTTATACGCC
1081 -----+-----+-----+-----+-----+-----+-----+ 1140
      CGATAAAAAAACCTCAGTATTACCTAAGAACAGTATTTTAACTAATACCCAATATGCGGG

      TGGAGATTTTAGCCCAATACCATAACGTCTCTGTAAACCCGGAAGAAATTAAACATAGAT
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
      ACCTCTAAAATCGGGTTATGGTATTGCAGAGACAATTGGGCCTTCTTTAATTGTATCTA

      TTGACACAGACGGGACTGGTCTGGGATTAACGTCATGGTTGCTTGCTGCGAAATCTTTAG
1201 -----+-----+-----+-----+-----+-----+-----+ 1260
      AACTGTGTCTGCCCTGACCAGACCCTAATTGCAGTACCAACGAACGACGCTTTAGAAATC

      AACTAAAGGTAAAACAGGTAAAAAAAACAATTGACCGATTAACTTTTATTTCTCTGCCCG
1261 -----+-----+-----+-----+-----+-----+-----+ 1320
      TTGATTTCCATTTTGTCCATTTTTTTGTTAACTGGCTAATTTGAAATAAAGAGACGGGC

      CATTAGTCTGGAGAGAGGATGGACGTCATTTTATTCTGACTAAAGTCAGTAAAGAAGCAA
1321 -----+-----+-----+-----+-----+-----+-----+ 1380
      GTAATCAGACCTCTCTCCTACCTGCAGTAAAATAAGACTGATTTTCAGTCATTTCCTCGTT

```

Fig. 7 (cont.)

ES 2 296 423 B1

```

ACAGATATCTTATTTCTGATCTGGAGCAGCGAAATCCCCGTGTTCTCGAACAGTCTGAGT
1381 -----+-----+-----+-----+-----+-----+ 1440
TGTCTATAGAATAAAGACTAGACCTCGTCGCTTTAGGGGCACAAGAGCTTGTGAGACTCA

TTGAGGCGTTATATCAGGGGCATATTATTCTTATCGCTTCCCGTTCTTCTGTGCGCCGGGA
1441 -----+-----+-----+-----+-----+-----+ 1500
AACTCCGCAATATAGTCCCCGTATAATAAGAATAGCGAAGGGCAAGAAGACAACGGCCCT

AACTGGCGAAATTTGACTTTACCTGGTTTATTCCTGCCATTATAAAATACAGGAGAATAT
1501 -----+-----+-----+-----+-----+-----+ 1560
TTGACCGCTTTAAACTGAAATGGACCAATAAGGACGGTAATATTTTATGTCCCTCTTATA

TTATTGAAACCCTTGTTGTGTCTGTTTTTTTTACAATTATTTGCATTAATAACCCCCCTTT
1561 -----+-----+-----+-----+-----+-----+ 1620
AATAACTTTGGGAACAACACAGACAAAAAATGTTAATAAACGTAATTATTGGGGGGAAA

TTTTTCAGGTGGTTATGGACAAAGTATTAGTGCACAGGGGATTTTCAACTCTTAATGTTA
1621 -----+-----+-----+-----+-----+-----+ 1680
AAAAAGTCCACCAATACCTGTTTCATAATCACGTGTCCCCTAAAAGTTGAGAATTACAAT

TTACTGTGCGATTATCTGTTGTGGTGGTGTGTTGAGATTATACTCAGCGGTTTAAGAACTT
1681 -----+-----+-----+-----+-----+-----+ 1740
AATGACAGCGTAATAGACAACACCACCACAAACTCTAATATGAGTCGCCAAATCTCTGAA

ACATTTTTGCACATAGTACAAGTCGGATTGATGTTGAGTTGGGTGCCAAACTCTTCCGGC
1741 -----+-----+-----+-----+-----+-----+ 1800
TGTA AAAACGTGTATCATGTTTCAGCCTAACTACA ACTCAACCCACGGTTTGAGAAGGCCG

ATTTACTGGCGCTACCGATCTCTTATTTTGAGAGTCGTCGTGTTGGTGATACGTGTTGCCA
1801 -----+-----+-----+-----+-----+-----+ 1860
TAAATGACCGCGATGGCTAGAGAATAAAACTCTCAGCAGCACAACTACTATGACAACGGT

GGGTAAGAGAATTAGACCAGATCCGTAATTTTCTGACAGGACAGGCATTAACATCTGTTC
1861 -----+-----+-----+-----+-----+-----+ 1920
CCCATTCTCTTAATCTGGTCTAGGCATTAAAAGACTGTCCGTGTCGTAATTGTAGACAAG

HindIII
      |
TGGACTTATTATTTTCATTCATATTTTTTGCGGTAATGTGGTATTACAGTCCAAAGCTT
1921 -----+-----+-----+-----+-----+-----+ 1979
ACCTGAATAATAAAAGTAAGTATAAAAAACGCCATTACACCATAATGTCAGGTTTCGAA

```

Fig. 7 (cont.)

## ES 2 296 423 B1

Enzimas que cortan:

BamHI	BglII	EagI	EcoRI	HindIII	NcoI	NheI	SalI
SfiI	XbaI	XmaI					

Enzimas que no cortan:

Ninguno

**Fig. 7**

# ES 2 296 423 B1

## Mapa de pVamyHLYA

Con 8 enzimas: NCOI PSTI SALI HINDIII SFII BAMHI NOTI ECORI PSTI

```

ACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGGA
1  -----+-----+-----+-----+-----+-----+-----+ 60
TGATATCCCTCTGGTGTGCCAAAGGGAGATCTTTATTAACAACAAATTGAAATTCTTCCT

      NcoI
      |
GATATATCCATGGCTCAGGTGCAGCTGGTGGAGTCTTGGGGAGGCTCGGTGCAGGCTGGG
61  -----+-----+-----+-----+-----+-----+ 120
CTATATAGGTACCGAGTCCACGTGACCACCTCAGAACCCCTCCGAGCCACGTCCGACCC

a      M A Q V Q L V E S W G G S V Q A G -
      |=> VHH amilasa

GGGTCTCTGAGACTCTCCTGCACAGCCCCTGGATTACCTCCAATAGCTGCCGCATGGAC
121 -----+-----+-----+-----+-----+-----+ 180
CCCAGAGACTCTGAGAGGACGTGTGCGGGACCTAAGTGGAGGTTATCGACCGCGTACCTG

a      G S L R L S C T A P G F T S N S C R M D -

      PstI
      |
TGGTACCGCCAGGCTGCAGGGAAGCAGCGCGAGTGGGTCTCATCTATTAGTACTGATGGT
181 -----+-----+-----+-----+-----+-----+ 240
ACCAAGGCGGTCCGACGTCCCTTCGTGCGGCTCACCCAGAGTAGATAATCATGACTACCA

a      W Y R Q A A G K Q R E W V S S I S T D G -

CGCACAAGCTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAAGACAAAGCCAAG
241 -----+-----+-----+-----+-----+-----+ 300
GCGTGTTCGATACGTCTGAGGCACTTCCCGGCTAAGTGGTAGAGGTTTCTGTTTCGGTTC

a      R T S Y A D S V K G R F T I S K D K A K -

GACACGGTGTATCTGCAAATGAACAGCCTGAAACCTGAGGACACGGCCATCTATTACTGT
301 -----+-----+-----+-----+-----+-----+ 360
CTGTGCCACATAGACGTTTACTTGTGCGGACTTTGGACTCCTGTGCCGGTACATAATGACA

a      D T V Y L Q M N S L K P E D T A I Y Y C -

GCCGTGAGGACGAATGGGTATCGTCCGCAATCTCACGAATTTGCTACTGCGGCCCCGGG
361 -----+-----+-----+-----+-----+-----+ 420
CGGCACTCCTGCTTACCCATAGCAGGCGTTAGAGTGCTTAAAGCGATGACCCCGGGCCCC

a      A V R T N G Y R P Q S H E F R Y W G P G -

```

Fig. 8 (cont.)

ES 2 296 423 B1

```

                                sfiI   SalI
                                |       |
ACCCAGGTCACCGTCTCCTCAACGGCCTCGGGGGCCGCGTCGACGCCCGGG3GTGCGCCG
421 -----+-----+-----+-----+-----+-----+ 480
TGGGTCCAGTGGCAGAGGAGTTGCCGGAGCCCCCGGCGCAGCTGCGGGCCCCACGCGGC

a   T   Q   V   T   V   S   S   T   A   S   G   A   A   S   T   P   G   G   A   P   -
                                |=> E-tag

                                BamHI
                                |
GTGCCGTATCCGGATCCGCTGGAACCGGCCGGGgAAAAATtCTCTTGCTAAAAATGTATTA
481 -----+-----+-----+-----+-----+-----+ 540
CACGGCATAGGCCTAGGCGACCTTGCCCGGCCcTTTTAAgGAGAACGATTTTTACATAAT

a   V   P   Y   P   D   P   L   E   P   A   G   E   N   S   L   A   K   N   V   L   -
                                |=> C-hlyA

TCCGGTGGAAAAGGTAATGACAAGTTGTACGGCAGTGAGGGAGCAGACCTGCTTGATGGC
541 -----+-----+-----+-----+-----+-----+ 600
AGGCCACCTTTTCCATTACTGTTCAACATGCCGTCACTCCCTCGTCTGGACGAACTACCG

a   S   G   G   K   G   N   D   K   L   Y   G   S   E   G   A   D   L   L   D   G   -

GGAGAAGGGAATGATCTTCTGAAAGGTGGATATGGTAATGATATTTATCGITATCTTTCA
601 -----+-----+-----+-----+-----+-----+ 660
CCTCTTCCCTTACTAGAAGACTTTCCACCTATACCATTACTATAAATAGCAATAGAAAGT

a   G   E   G   N   D   L   L   K   G   G   Y   G   N   D   I   Y   R   Y   L   S   -

GGATATGGCCATCATATTATTGACGATGAAGGGGGAAAGACGATAAACTCAGTTTAGCT
661 -----+-----+-----+-----+-----+-----+ 720
CCTATACCGGTAGTATAATAACTGCTACTTCCCCCTTTCTGCTATTTGACTCAAATCGA

a   G   Y   G   H   H   I   I   D   D   E   G   G   K   D   D   K   L   S   L   A   -

GATATAGATTTCCGGGACGTTGCCTTTAAGCGAGAAGGGAATGACCTCATIATGTATAAA
721 -----+-----+-----+-----+-----+-----+ 780
CTATATCTAAAGGCCCTGCAACGGAAATTCGCTCTTCCCTTACTGGAGTAAATACATATTT

a   D   I   D   F   R   D   V   A   F   K   R   E   G   N   D   L   I   M   Y   K   -

GCTGAAGGTAATGTTCTTTCTATTGGCCACAAAATGGTATTACATTTAAAAACTGGTTT
781 -----+-----+-----+-----+-----+-----+ 840
CGACTTCCATTACAAGAAAGATAACCGGTGTTTTTACCATAATGTAAATTTTGGACCAA

a   A   E   G   N   V   L   S   I   G   H   K   N   G   I   T   F   K   N   W   F   -

GAAAAAGAGTCAGATGATCTCTCTAATCATCAGATAGAGCAGATTTTGTATAAAGACGGC
841 -----+-----+-----+-----+-----+-----+ 900
CTTTTTCTCAGTCTACTAGAGAGATTAGTAGTCTATCTCGTCTAAAACTATTTCTGCCG

a   E   K   E   S   D   D   L   S   N   H   Q   I   E   Q   I   F   D   K   D   G   -

AGGGTAATCACACCAGATTCTCTTAAAAAAGCATTGGAATATCAGCAGAGT'AATAACAAG
901 -----+-----+-----+-----+-----+-----+ 960
TCCCATTAGTGTGGTCTAAGAGAATTTTTTCGTAAACTTATAGTCGTCTCATATTGTTC

```

Fig. 8 (cont.)

# ES 2 296 423 B1

```

a      R V I T P D S L K K A F E Y Q Q S N N K -
      GTAAGTTATGTGTATGGACATGATGCATCAACTTATGGGAGCCAGGACAATCTTAATCCA
961  -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1020
      CATTCAATACACATACCTGTACTACGTAGTTGAATACCCTCGGTCCTGTTAASAATTAGGT

a      V S Y V Y G H D A S T Y G S Q D N L N P -
      PstI
      |
      TTAATTAATGAAATCAGCAAAATCATTTCAGCTGCAGGTAACCTTCGATGTTAAGGAGGAA
1021 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080
      AATTAATTACTTTAGTCGTTTTAGTAAAGTCGACGTCCATTGAAGCTACAAATCCTCCTT

a      L I N E I S K I I S A A G N F D V K E E -
      AGATCTGCCGCTTCTTTATTGCAGTTGTCCGGTAATGCCAGTGATTTTTTCATATGGACGG
1081 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1140
      TCTAGACGGCGAAGAAATAACGTCAACAGGCCATTACGGTCACTAAAAAGTATACCTGCC

a      R S A A S L L Q L S G N A S D F S Y G R -
      AACTCAATAACTTTGACAGCATCAGCATAATATATTAATTTAAATGATAGCAATCTTACT
1141 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200
      TTGAGTTATTGAAACTGTCGTAGTCGTATTATATAATTAATTTACTATCGTTAGAATGA

a      N S I T L T A S A * -
      GGGCTGTGCCACATAAGATTGCTATTTTTTTGGAGTCATAATGGATTCTTCTCATAAAAT
1201 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1260
      CCCGACACGGTGTATTCTAACGATAAAAAAACCTCAGTATTACCTAAGAACAGTATTTTA

      TGATTATGGGTTATACGCCCTGGAGATTTTAGCCCAATACCATAACGTCTCTGTTAACCC
1261 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1320
      ACTAATACCCAATATGCGGGACCTCTAAAATCGGGTTATGGTATTGCAGACACAATTGGG

      GGAAGAAATTAACATAGATTTGACACAGACGGGACTGGTCTGGGATTAACGTCATGGTT
1321 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1380
      CCTTCTTTAATTTGTATCTAAACTGTGTCTGCCCTGACCAGACCCTAATTCAGTACCAA

      GCTTGCTGCGAAATCTTTAGAACTAAAGGTA AAAACAGGTAAAAAAAACAATGACCGATT
1381 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1440
      CGAACGACGCTTTAGAAATCTTGATTTCATTTTGTCCATTTTTTTTTGTTACTGGCTAA

      AAACCTTTATTTCTCTGCCCGCATTAGTCTGGAGAGAGGATGGACGTCATTTATTCTGAC
1441 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500
      TTTGAAATAAAGAGACGGGCGTAATCAGACCTCTCTCCTACCTGCAGTAAATAAGACTG

```

**Fig. 8 (cont.)**

ES 2 296 423 B1

TAAAGTCAGTAAAGAAGCAAACAGATATCTTATTTCTGATCTGGAGCAGCGAAATCCCCG  
 1501 -----+-----+-----+-----+-----+-----+ 1560  
 ATTTTCAGTCATTTCTTCGTTTGTCTATAGAATAAAGACTAGACCTCGTCGC TTTAGGGGC  
  
 TGTTCCTCGAACAGTCTGAGTTTGAGGCGTTATATCAGGGGCATATTATTCTTATCGCTTC  
 1561 -----+-----+-----+-----+-----+-----+ 1620  
 ACAAGAGCTTGTTCAGACTCAAACCTCCGCAATATAGTCCCCGTATAATAAGAAATAGCGAAG  
  
 CCGTTCTTCTGTTGCCGGGAAACTGGCGAAATTTGACTTTACCTGGTTTATTCCTGCCAT  
 1621 -----+-----+-----+-----+-----+-----+ 1680  
 GGCAAGAAGACAACGGCCCTTTGACCGCTTTAAACTGAAATGGACCAAATAAGGACGGTA  
  
 TATAAAATACAGGAGAATATTTATTGAAACCCTTGTGTGTCTGTTTTTTTTTACAATTATT  
 1681 -----+-----+-----+-----+-----+-----+ 1740  
 ATATTTTATGTCCTCTTATAAAATAACTTTGGGAACAACACAGACAAAAAAATGTTAATAA  
  
 TGCATTAATAACCCCCCTTTTTTTTTCAGGTGGTTATGGACAAAGTATTAGTGCACAGGGG  
 1741 -----+-----+-----+-----+-----+-----+ 1800  
 ACGTAATTATTGGGGGGAAAAAAAAGTCCACCAATACCTGTTTCATAATCACGTGTCCCC  
  
 ATTTTCAACTCTTAATGTTATTACTGTGCGATTATCTGTTGTGGTGGTGTITGAGATTAT  
 1801 -----+-----+-----+-----+-----+-----+ 1860  
 TAAAAGTTGAGAATTACAATAATGACAGCGTAATAGACAACACCACCACAAACTCTAATA  
  
 ACTCAGCGGTTTAAGAACTTACATTTTTTGCACATAGTACAAGTCGGATTGATGTTGAGTT  
 1861 -----+-----+-----+-----+-----+-----+ 1920  
 TGAGTCGCCAAATTCTTGAATGTA AAAACGTGTATCATGTTTCAGCCTAACTACAAC TCAA  
  
 GGGTGCCAAACTCTTCCGGCATTACTGGCGCTACCGATCTCTTATTTTGAGAGTCGTCCG  
 1921 -----+-----+-----+-----+-----+-----+ 1980  
 CCCACGGTTTGAGAAGGCCGTAAATGACCGGATGGCTAGAGAATAAAAC TCTCAGCAGC  
  
 TGTTGGTGATACTGTTGCCAGGGTAAGAGAATTAGACCAGATCCGTAATTTCTGACAGG  
 1981 -----+-----+-----+-----+-----+-----+ 2040  
 ACAACC ACTATGACAACGGTCCCATTCTCTTAATCTGGTCTAGGCATTAAPAGACTGTCC  
  
 ACAGGCATTAACATCTGTTCTGGACTTATTATTTTCATTCATATTTTTTGC GGTAATGTG  
 2041 -----+-----+-----+-----+-----+-----+ 2100  
 TGTCCGTAATTGTAGACAAGACCTGAATAATAAAAAGTAAGTATAAAAAACCCATTACAC

Fig. 8 (cont.)

ES 2 296 423 B1

```

                HindIII
                |
2101  GTATTACAGTCCAAAGCTTACTCTGGTGATCTTATTTTCGCTGCCTTGTTAIGCTGCATG
-----+-----+-----+-----+-----+-----+-----+ 2160
      CATAATGTCAGGTTTCGAATGAGACCACTAGAATAAAAAGCGACGGAACAATACGACGTAC

2161  GTCTGTTTTTATTAGCCCCATTTTGGCGACGTCGCCTTGATGATAAGTTTTACGGAATGC
-----+-----+-----+-----+-----+-----+-----+ 2220
      CAGACAAAATAATCGGGGTAAAACGCTGCAGCGGAACTACTATTCAAAGTGCCTTACG

2221  GGATAATCAATCTTTCCTGGTGGAAATCAGTCACGGCGATTAACACTATAAAAAGCTATGGC
-----+-----+-----+-----+-----+-----+-----+ 2280
      CCTATTAGTTAGAAAGGACCACCTTAGTCAGTGCCGCTAATTGTGATATTTTCGATACCG

                Pst I
                |
2281  AGTCTCACCTCAGATGACGAACATATGGGACAAACAATTGGCAGGATATGTTGCTGCAGG
-----+-----+-----+-----+-----+-----+-----+ 2340
      TCAGAGTGGAGTCTACTGCTTGTATACCCTGTTTGTAAACCGTCTATAACAACGACGTCC

2341  CTTCAAAGTGACAGTATTAGCAACCATTGGTCAACAAGGAATACAGTTAATACAAAAGAC
-----+-----+-----+-----+-----+-----+-----+ 2400
      GAAGTTTCACTGTCAATAATCGTTGGTAACCAGTTGTTCCCTTATGTCAATTATGTTTTCTG

2401  TGTTATGATCATCAACCTGTGGTTGGGAGCACACCTGGTTATTTCCGGGGATTTAAGTAT
-----+-----+-----+-----+-----+-----+-----+ 2460
      ACAATACTAGTAGTTGGACACCAACCCTCGTGTGGACCAATAAAGGCCCTAAATTCATA

2461  TGGTCAGTTAATTGCTTTTAATATGCTTGCTGGTCAGATTGTTGCACCGGITATTTCGCCT
-----+-----+-----+-----+-----+-----+-----+ 2520
      ACCAGTCAATTAACGAAAATTATACGAACGACCAGTCTAACAACGTGGCCAATAAGCGGA

2521  TGCACAAATCTGGCAGGATTTCCAGCAGGTTGGTATATCAGTTACCCGCCTTGGTGATGT
-----+-----+-----+-----+-----+-----+-----+ 2580
      ACGTGTTTAGACCGTCTAAAGGTCTCAACCATATAGTCAATGGGCGGAACCACTACA

2581  GCTTAACTCTCCTAACTGAAAGTTATCATGGGAACTGGCATTACCGGAAATTAATGGTGA
-----+-----+-----+-----+-----+-----+-----+ 2640
      CGAATTGAGAGGTTGACTTTCAATAGTACCCTTTGACCGTAATGGCCTTTAATTACCACT

2641  TATCACTTTTCGTAATATCCGGTTTCGCTATAAGCCTGACTCTCCGGTTATTTAGATAA
-----+-----+-----+-----+-----+-----+-----+ 2700
      ATAGTGAAAAGCATTATAGGCCAAAGCGATATTCGGACTGAGAGGCCAATAAATCTATT

2701  TATCAATCTCAGTATTAAGCAGGGGGAGGTTATTGGTATTGTTCGGACGTTCTGGTTTCAGG
-----+-----+-----+-----+-----+-----+-----+ 2760
      ATAGTTAGAGTCATAATTCGTCCCCCTCCAATAACCATAACAGCCTGCAAGACCAAGTCC

```

Fig. 8 (cont.)

## ES 2 296 423 B1

```
AAAAAGCACATTAACATAAATTAATTCAACGTT
2761 -----+-----+-----+--- 2792
TTTTTCGTGTAATTGATTTAATTAAGTTGCAA
```

Enzimas que cortan:

BamHI HindIII NcoI PstI SalI SfiI

Enzimas que no cortan:

EcoRI NotI

**Fig. 8**

# ES 2 296 423 B1

## Mapa de pVamyZHLYA

Con 11 enzimas: ECORI SPHI PSTI NCOI NHEI NDEI BAMHI HINDIII  
SALI SFII NOTI

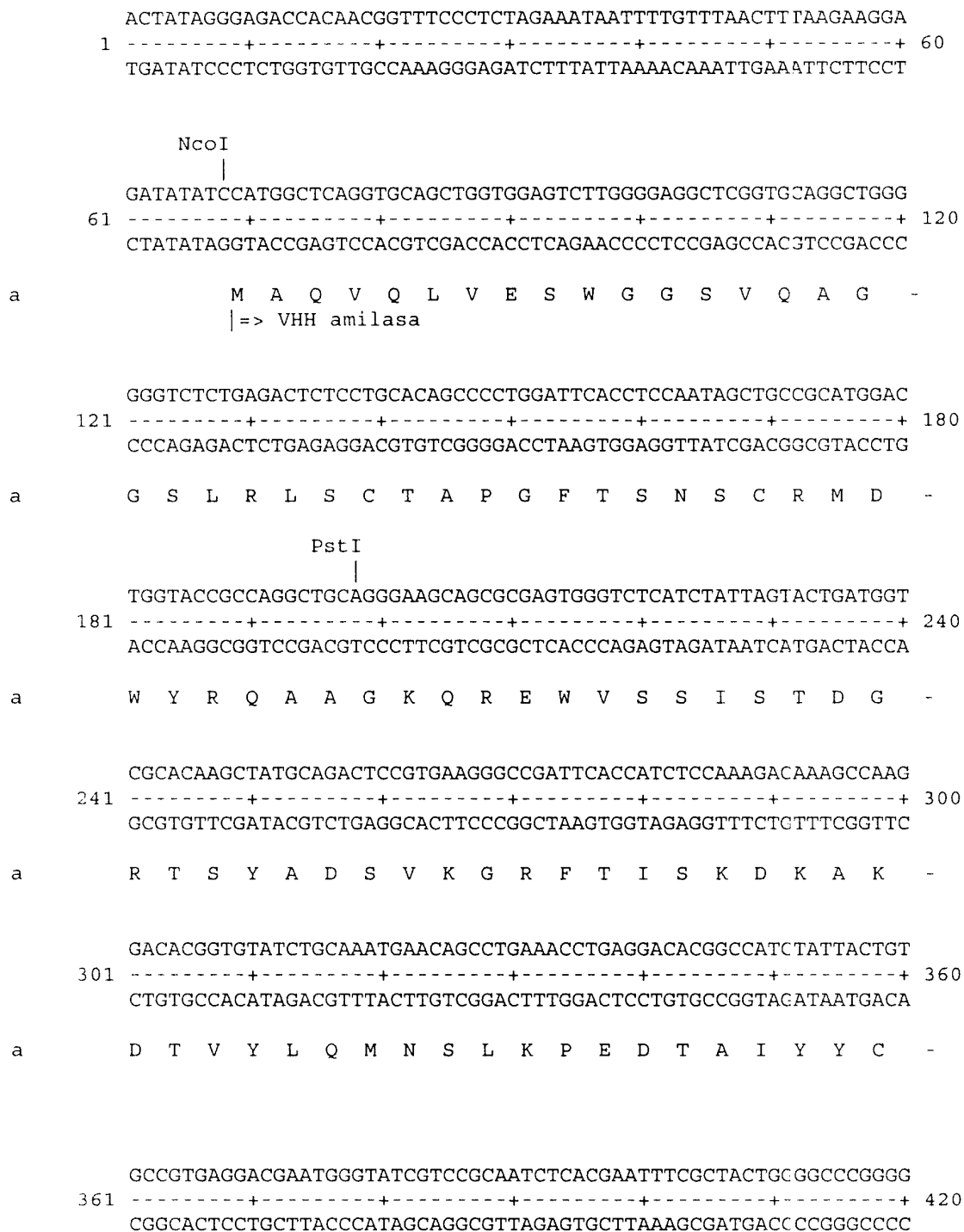


Fig. 9 (cont.)

ES 2 296 423 B1

```

a   A V R T N G Y R P Q S H E F R Y W G P G -
                                     SfiI   SalI
                                     |     |
ACCCAGGTCACCGTCTCCTCAACGGCCTCGGGGGCCGCGTCGACGTCCGGC3GTCCGAAG
421 -----+-----+-----+-----+-----+-----+ 480
TGGGTCCAGTGGCAGAGGAGTTGCCGGAGCCCCCGGCGCAGCTGCAGGCCGCCAGGCTTC

a   T Q V T V S S T A S G A A S T S G G P K -
                                     |=> Ig hinge

CCTTCCACTCCGCCCCGGTCTTCCCGTATGAAACAGCTGGAAGACAAAGTA3AgGAGCTC
481 -----+-----+-----+-----+-----+-----+ 540
GGAAGGTGAGGCCGGGCCAGAAGGGCATACTTTGTGACCTTCTGTTTCATCTTCTCGAG

a   P S T P P G S S R M K Q L E D K V E E L -
                                     |=> Leucine Zipper GCN4

                                     HindIII
                                     |
CTTAGCAAGAACTACCATCTAGAAAACGAGGTAGCTCGTCTGAAAAAGCTTGTGGTGAA
541 -----+-----+-----+-----+-----+-----+ 600
GAATCGTTCTTGATGGTAGATCTTTTGTCTCCATCGAGCAGACTTTTTCGAACCAACCACTT

a   L S K N Y H L E N E V A R L K K L V G E -

                                     SalI
                                     |
CGTGGTGGTCACCATCACCATCACCATGCGTTCGACGCCCCGGGGTGCGCCGGTGCCGTAT
601 -----+-----+-----+-----+-----+-----+ 660
GCACCACCAGTGGTAGTGGTAGTGGTACGCAGCTGCGGGCCCCCACGCGCCACGGCATA

a   R G G H H H H H A S T P G G A P V P Y -
       |=> 6xhis tag                               |=> E-tag

BamHI
|
CCGGATCCGCTGGAACCGGCCGGGgAAAATtCTCTTGCTAAAAATGTATTATCCGGTGGA
661 -----+-----+-----+-----+-----+-----+ 720
GGCCTAGGCACCTTGGCCGGCCCcTTTAAgGAGAACGATTTTTACATAATAGGCCACCT

a   P D P L E P A G E N S L A K N V L S G G -
                                     |=> C-hlyA

AAAGGTAATGACAAGTTGTACGGCAGTGAGGGAGCAGACCTGCTTGATGGCGGAGAAGGG
721 -----+-----+-----+-----+-----+-----+ 780
TTTCCATTACTGTTCAACATGCCGTCACTCCCTCGTCTGGACGAACCTACCCCTCTTCCC

a   K G N D K L Y G S E G A D L L D G G E G -

```

Fig. 9 (cont.)

ES 2 296 423 B1

```

AATGATCTTCTGAAAGGTGGATATGGTAATGATATTTATCGTTATCTTTCAAGATATGGC
781 -----+-----+-----+-----+-----+-----+-----+ 840
TTACTAGAAGACTTTCCACCTATACCATTACTATAAATAGCAATAGAAAGTCTATACCG

a      N D L L K G G Y G N D I Y R Y L S G Y G -

CATCATATTATTGACGATGAAGGGGGAAAGACGATAAACTCAGTTTAGCTGATATAGAT
841 -----+-----+-----+-----+-----+-----+-----+ 900
GTAGTATAATAACTGCTACTTCCCCCTTCTGCTATTTGAGTCAAATCGACTATATCTA

a      H H I I D D E G G K D D K L S L A D I D -

TTCCGGGACGTTGCCTTTAAGCGAGAAGGGAATGACCTCATTATGTATAAAGCTGAAGGT
901 -----+-----+-----+-----+-----+-----+-----+ 960
AAGGCCCTGCAACGGAATTCGCTCTTCCCTTACTGGAGTAATACATATTTGACTTCCA

a      F R D V A F K R E G N D L I M Y K A E G -

AATGTTCTTTCTATTGGCCACAAAATGGTATTACATTTAAAAACTGGTTTGAAAAAGAG
961 -----+-----+-----+-----+-----+-----+-----+ 1020
TTACAAGAAAAGATAACCGGTGTTTTTACCATAATGTAAATTTTTGACCAAACFTTTTCTC

a      N V L S I G H K N G I T F K N W F E K E -

TCAGATGATCTCTCTAATCATCAGATAGAGCAGATTTTTGATAAAGACGGCAGGGTAATC
1021 -----+-----+-----+-----+-----+-----+-----+ 1080
AGTCTACTAGAGAGATTAGTAGTCTATCTCGTCTAAAAACTATTTCTGCCGTCCCATTAG

a      S D D L S N H Q I E Q I F D K D G R V I -

ACACCAGATTCTCTTAAAAAAGCATTGGAATATCAGCAGAGTAATAACAAGGTAAGTTAT
1081 -----+-----+-----+-----+-----+-----+-----+ 1140
TGTGGTCTAAGAGAATTTTTTCGTAACTTATAGTCGTCTCATTATTGTTCCATTCAATA

a      T P D S L K K A F E Y Q Q S N N K V S Y -

GTGTATGGACATGATGCATCAACTTATGGGAGCCAGGACAATCTTAATCCATTAATTAAT
1141 -----+-----+-----+-----+-----+-----+-----+ 1200
CACATACCTGTACTACGTAGTTGAATACCCTCGGTCTGTTAGAATTAGGIAATTAATTA

a      V Y G H D A S T Y G S Q D N L N P L I N -

```

Fig. 9 (cont.)

ES 2 296 423 B1

PstI  
|

1201 GAAATCAGCAAAATCATTTCAGCTGCAGGTAACCTTCGATGTTAAGGAGGAAAGATCTGCC  
-----+-----+-----+-----+-----+-----+ 1260  
CTTTAGTCGTTTTAGTAAAGTCGACGTCCATTGAAGCTACAATTCCTCCTTICTAGACGG

a E I S K I I S A A G N F D V K E E R S A -

NdeI  
|

1261 GCTTCTTTATTGCAGTTGTCCGGTAATGCCAGTGATTTTTCATATGGACGGAACTCAATA  
-----+-----+-----+-----+-----+-----+ 1320  
CGAAGAAATAACGTCAACAGGCCATTACGGTCACTAAAAAGTATACCTGCCITGAGTTAT

a A S L L Q L S G N A S D F S Y G R N S I -

1321 ACTTTGACAGCATCAGCATAATATATTAATTTAAATGATAGCAATCTTACTGGGCTGTGC  
-----+-----+-----+-----+-----+-----+ 1380  
TGAAACTGTCTAGTCGTATTATATAATTAATTTACTATCGTTAGAATGACCCGACACG

a T L T A S A \* -

1381 CACATAAGATTGCTATTTTTTTGGAGTCATAATGGATTCTTGTCAAAAATTGATTATGG  
-----+-----+-----+-----+-----+-----+ 1440  
GTGTATTCTAACGATAAAAAAACCTCAGTATTACCTAAGAACAGTATTTTAACTAATACC

1441 GTTATACGCCCTGGAGATTTTAGCCCAATACCATAACGTCTCTGTTAACCCGGAAGAAAT  
-----+-----+-----+-----+-----+-----+ 1500  
CAATATGCGGGACCTCTAAAATCGGGTTATGGTATTGCAGAGACAATTGGGCCTTCTTTA

1501 TAAACATAGATTTGACACAGACGGGACTGGTCTGGGATTAACGTCATGGTIGCTTGCTGC  
-----+-----+-----+-----+-----+-----+ 1560  
ATTTGTATCTAAACTGTGTCTGCCCTGACCAGACCCTAATTGCAGTACCAACGAACGACG

1561 GAAATCTTTAGAACTAAAGGTA AAAACAGGTAAAAAAAACAATTGACCGATTAACTTTAT  
-----+-----+-----+-----+-----+-----+ 1620  
CTTTAGAAATCTTGATTTCCATTTTGTCCATTTTTTTTGTAACTGGCTAATTTGAAATA

1621 TTCTCTGCCCGCATTAGTCTGGAGAGAGGATGGACGTCATTTTATTCTGACTAAAGTCAG  
-----+-----+-----+-----+-----+-----+ 1680  
AAGAGACGGGCGTAATCAGACCTCTCTCCTACCTGCAGTAAAAATAAGACTCATTTCAGTC

1681 TAAAGAAGCAAACAGATATCTTATTTCTGATCTGGAGCAGCGAAATCCCCGTTTCTCGA  
-----+-----+-----+-----+-----+-----+ 1740  
ATTTCTTCGTTTGTCTATAGAATAAAGACTAGACCTCGTCGCTTTAGGGGCACAAGAGCT

Fig. 9 (cont.)

ES 2 296 423 B1

```

ACAGTCTGAGTTTGAGGCGTTATATCAGGGGCATATTATTCTTATCGCTTCCCGTTCTTC
1741 -----+-----+-----+-----+-----+-----+-----+ 1800
TGTCAGACTCAAACCTCCGCAATATAGTCCCCGTATAATAAGAATAGCGAAGGGCAAGAAG

TGTTGCCGGGAAACTGGCGAAATTTGACTTTACCTGGTTTATTCCTGCCATATATAAAATA
1801 -----+-----+-----+-----+-----+-----+-----+ 1860
ACAACGGCCCTTTGACCGCTTTAAACTGAAATGGACCAAATAAGGACGGTAATATTTTAT

CAGGAGAATATTTATTGAAACCCCTTGTTGTGTCTGTTTTTTTTACAATTATTTCATTAAAT
1861 -----+-----+-----+-----+-----+-----+-----+ 1920
GTCCTCTTATAAATAACTTTGGGAACAACACAGACAAAAAATGTTAATAAACGTAATTA

AACCCCCCTTTTTTTTTCAGGTGGTTATGGACAAAGTATTAGTGCACAGGGGATTTTCAAC
1921 -----+-----+-----+-----+-----+-----+-----+ 1980
TTGGGGGGAAAAAAGTCCACCAATACCTGTTTCATAATCACGTGTCCCTAAAAGTTG

TCTTAATGTTATTACTGTGCGCATTATCTGTTGTGGTGGTGGTTTGAGATTATACTCAGCGG
1981 -----+-----+-----+-----+-----+-----+-----+ 2040
AGAATTACAATAATGACAGCGTAATAGACAACACCACCACAACTCTAATAIGAGTCGCC

TTTAAGAACTTACATTTTTGCACATAGTACAAGTCGGATTGATGTTGAGTTGGGTGCCAA
2041 -----+-----+-----+-----+-----+-----+-----+ 2100
AAATTCTTGAATGTAAAAACGTGTATCATGTTTCAGCCTAACTACAACCAACCCACGGTT

ACTCTTCCGGCATTACTGCGCTACCGATCTCTTATTTTGAGAGTCGTCGTGTTGGTGA
2101 -----+-----+-----+-----+-----+-----+-----+ 2160
TGAGAAGGCCGTAAATGACCGCGATGGCTAGAGAATAAACTCTCAGCAGCACAACTACT

TACTGTTGCCAGGGTAAGAGAATTAGACCAGATCCGTAATTTTCTGACAGGACAGGCATT
2161 -----+-----+-----+-----+-----+-----+-----+ 2220
ATGACAACGGTCCCATTCTCTTAATCTGGTCTAGGCATTAAGACTGTCTGTCCGTAA

AACATCTGTTCTGGACTTATTATTTTCATTCATATTTTTTTCGGTAATGTGGTATTACAG
2221 -----+-----+-----+-----+-----+-----+-----+ 2280
TTGTAGACAAGACCTGAATAATAAAAGTAAGTATAAAAAACGCCATTACACCATAATGTC

HindIII
|
TCCAAAGCTTACTCTGGTGATCTTATTTTCGCTGCCTTGTTATGCTGCATCGTCTGTTTT
2281 -----+-----+-----+-----+-----+-----+-----+ 2340
AGGTTTCGAATGAGACCACTAGAATAAAAGCGACGGAACAATACGACGTAC CAGACAAAA

TATTAGCCCCATTTTTCGACGTCGCCTTGATGATAAGTTTTTTCACGGAATGCGGATAATCA
2341 -----+-----+-----+-----+-----+-----+-----+ 2400
ATAATCGGGGTAAAACGCTGCAGCGGAACACTATTCAAAGTGCCTTACCCCTATTAGT

```

Fig. 9 (cont.)

ES 2 296 423 B1

```

ATCTTTCCTGGTGAATCAGTCACGGCGATTAACACTATAAAAAGCTATGGCAGTCTCACC
2401 -----+-----+-----+-----+-----+-----+-----+ 2460
TAGAAAGGACCACCTTAGTCAGTGCCGCTAATTGTGATATTTTCGATACCGTCAGAGTGG

                NdeI                               PstI
                |                                   |
TCAGATGACGAACATATGGGACAAACAATTGGCAGGATATGTTGCTGCAGGCTTCAAAGT
2461 -----+-----+-----+-----+-----+-----+-----+ 2520
AGTCTACTGCTTGTATACCCTGTTTGTAAACCGTCCTATACAACGACGTCCAAAGTTTCA

GACAGTATTAGCAACCATTGGTCAACAAGGAATACAGTTAATACAAAAGACIGTTATGAT
2521 -----+-----+-----+-----+-----+-----+-----+ 2580
CTGTCATAATCGTTGGTAACCAGTTGTTCCCTTATGTCAATTATGTTTCTGACAATACTA

CATCAACCTGTGGTTGGGAGCACACCTGGTTATTTCCGGGGATTTAAGTATGGTCAGTT
2581 -----+-----+-----+-----+-----+-----+-----+ 2640
GTAGTTGGACACCAACCCTCGTGTGGACCAATAAAGGCCCTAAATTCATAACCAGTCAA

AATTGCTTTTAAATATGCTTGCTGGTCAGATTGTTGCACCGGTTATTCGCCTTGACAAAAT
2641 -----+-----+-----+-----+-----+-----+-----+ 2700
TTAACGAAAATTATACGAACGACCAGTCTAACAACGTGGCCAATAAGCGGAACGTGTTTA

CTGGCAGGATTTCCAGCAGGTTGGTATATCAGTTACCCGCCTTGGTGATGIGCTTAACTC
2701 -----+-----+-----+-----+-----+-----+-----+ 2760
GACCGTCCTAAAGGTCGTCCAACCATATAGTCAATGGGCGGAACCACTACACGAATTGAG

TCCAAGTAAAGTTATCATGGGAACTGGCATTACCGGAAATTAATGGTGATATCACTTT
2761 -----+-----+-----+-----+-----+-----+-----+ 2820
AGGTTGACTTTCAATAGTACCCTTTGACCGTAATGGCCTTTAATTACCACIATAGTGAAA

TCGTAATATCCGGTTTCGCTATAAGCCTGACTCTCCGGTTATTTTAGATAATATCAATCT
2821 -----+-----+-----+-----+-----+-----+-----+ 2880
AGCATTATAGGCCAAAGCGATATTCGACTGAGAGGCCAATAAAATCTATTATAGTTAGA

CAGTATTAAGCAGGGGGAGGTTATTGGTATTGTCCGACGTTCTGGTTCAGCAAAAAGCAC
2881 -----+-----+-----+-----+-----+-----+-----+ 2940
GTCATAATTCGTCCCCCTCCAATAACCATAACAGCCTGCAAGACCAAGTCC'TTTTTCGTG

ATTAAGTAAATTAATTCAACGTT
2941 -----+-----+-----+-----+-----+-----+-----+ 2963
TAATTGATTTAATTAAGTTGCAA

```

Fig. 9 (cont.)

## ES 2 296 423 B1

Enzimas que cortan:

BamHI HindIII NcoI NdeI PstI SalI SfiI

Enzimas que no cortan:

EcoRI NheI NotI SphI

**Fig. 9**

# ES 2 296 423 B1

## LISTA DE SECUENCIAS

<110> Consejo Superior de Investigaciones Científicas

5 <120> Sistema para la producción de proteínas diméricas basado en el sistema de transporte de hemolisina de *Escherichia coli*

<160> 8

10

<170> PatentIn version 2.0

<210> 1

15 <211> 654

<212> DNA

<213> *Escherichia coli*

20 <400> 1

```
ggaaaattct cttgctaaaa atgtattatc cggtggaana ggtaatgaca agttgtacgg 60
cagtgagggg gcagacctgc ttgatggcgg agaaggggat gatcttctga aaggtggata 120
tggtaatgat atttatcggt atctttcagg atatggccat catattattg acgatgaagg 180
25 ggggaaagac gataaactca gtttagctga tatagatttc cgggacgttg cctttaagcg 240
agaaggggat gacctcatta tgtataaagc tgaaggtaat gttctttcta ttggccacaa 300
aaatggtatt acatttaaaa actggtttga aaaagagtca gatgatctct ctaatcatca 360
gatagagcag atttttgata aagacggcag ggtaatcaca ccagattctc ttaaaaaagc 420
30 atttgaatat cagcagagta ataacaaggt aagttatgtg tatggacatg atgcatcaac 480
ttatgggagc caggacaatc ttaatccatt aattaatgaa atcagcaaaa tcatttcagc 540
tgcaggtaac ttcgatgtta aggaggaaaag atctgccgct tctttattgc agttgtccgg 600
taatgccagt gatttttcat atggacggaa ctcaataact ttgacagcat cagc 654
```

35 <210> 2

<211> 218

<212> PRT

<213> *Escherichia coli*

40

<400> 2

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

# ES 2 296 423 B1

	115	120	125
5	Gly Arg Val Ile Thr Pro Asp Ser Leu Lys Lys Ala Phe Glu Tyr Gln 130 135 140		
	Gln Ser Asn Asn Lys Val Ser Tyr Val Tyr Gly His Asp Ala Ser Thr 145 150 155 160		
10	Tyr Gly Ser Gln Asp Asn Leu Asn Pro Leu Ile Asn Glu Ile Ser Lys 165 170 175		
	Ile Ile Ser Ala Ala Gly Asn Phe Asp Val Lys Glu Glu Arg Ser Ala 180 185 190		
15	Ala Ser Leu Leu Gln Leu Ser Gly Asn Ala Ser Asp Phe Ser Tyr Gly 195 200 205		
20	Arg Asn Ser Ile Thr Leu Thr Ala Ser Ala 210 215		

<210> 3

<211> 30

<212> DNA

<213> Secuencia artificial

<223> Región bisagra de anticuerpo

30

<400> 3

cgtccgaag cttccactc cgcccgggtc

30

35

<210> 4

<211> 10

<212> PRT

40 <213> Región bisagra de anticuerpo

<400> 4

45	Gly Pro Lys Pro Ser Thr Pro Pro Gly Ser 1 5 10
----	---

<210> 5

<211> 33

<212> DNA

<213> Secuencia artificial

<223> Oligonucleótido iniciador BACKHINGE

55

<400> 5

gcgtcgacgt cggcgggtcc gaagccttc act

33

60

<210> 6

<211> 44

<212> DNA

65 <213> Secuencia artificial

<223> Oligonucleótido iniciador FORHIS

# ES 2 296 423 B1

<400> 6

gcgtcgacgc atggtgatgg tgatggtgac cacc

44

5

<210> 7

<211> 47

<212> DNA

10 <213> Secuencia artificial

<223> Oligonucleótido iniciador VHHA1

<400> 7

15

ctatgcggcc cagccggcca tggctcaggt gcagctggtg gagtctt

47

<210> 8

20

<211>47

<212> DNA

<213> Secuencia artificial

<223> Oligonucleótido iniciador VHHASfi1

25

<400> 8

cgtcgacgcg gccccgagg cgttgagga gacggtgacc tgggtccc

30

48

35

40

45

50

55

60

65



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

① ES 2 296 423

② Nº de solicitud: 200301830

③ Fecha de presentación de la solicitud: 31.07.2003

④ Fecha de prioridad:

## INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤ Int. Cl.: C12N 15/67 (2006.01)  
C07K 16/18 (2006.01)

### DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
A	PACK, P. et al. "Miniantibodies: use of amphipatic helices to produce functional, flexibly linked dimeric FV fragments with high avidity in Escherichia coli". BIOCHEMISTRY. 1992. Vol. 31, páginas 1579-1584, todo el documento.	1-43
A	FERNANDEZ, L.A. et al. "Specific secretion of active single-chain Fv antibodies into the supernatants of Escherichia coli cultures by use of the hemolysin system". APPLIED AND ENVIRONMENTAL MICROBIOLOGY. Noviembre 2000. Vol.66, nº 11, páginas 5024-5029, todo el documento.	1-43
A	WO 8706953 A1 (HOLLAND, I.B. et al) 19.11.1987, todo el documento.	1-43

#### Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

#### El presente informe ha sido realizado

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe  
29.11.2004

Examinador  
M. Novoa Sanjurjo

Página  
1/1