



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

① Número de publicación: **2 327 309**

② Número de solicitud: 200800909

⑤ Int. Cl.:  
**C12N 15/62** (2006.01)

**C07K 14/42** (2006.01)

⑫

PATENTE DE INVENCION

B1

⑫ Fecha de presentación: **02.04.2008**

⑬ Fecha de publicación de la solicitud: **27.10.2009**

Fecha de la concesión: **09.07.2010**

⑮ Fecha de anuncio de la concesión: **26.07.2010**

⑯ Fecha de publicación del folleto de la patente:  
**26.07.2010**

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

⑱ Inventor/es: **Mancheño Gómez, José María y**  
**Angulo Herrera, Iván**

⑳ Agente: **Pons Ariño, Ángel**

㉑ Título: **Proteínas de fusión con un dominio lectina de tipo  $\beta$ -trébol, procedimiento de obtención y sus aplicaciones.**

㉒ Resumen:

Proteínas de fusión con un dominio lectina de tipo  $\beta$ -trébol, procedimiento de obtención y sus aplicaciones. La presente invención describe un nuevo método general de expresión y purificación de proteínas de fusión conteniendo un dominio peptídico lectina beta-trébol como etiqueta de afinidad y solubilidad. Esta etiqueta permite producir diferentes proteínas de fusión solubles con elevados rendimientos y purificarlas mediante un protocolo de purificación eficaz, sencillo y de bajo coste. Dicho protocolo está basado en la capacidad del péptido para unir azúcares derivados de galactosa y consta de una sola etapa de cromatografía de afinidad que emplea matrices derivadas de agarosa y lactosa como eluyente.

ES 2 327 309 B1

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

## DESCRIPCIÓN

Proteínas de fusión con un dominio lectina de tipo  $\beta$ -trébol, procedimiento de obtención y sus aplicaciones.

## 5 Sector de la técnica

La invención se puede encuadrar de un modo genérico en el sector de la *biotecnología*, con evidentes aplicaciones en el campo farmacéutico y, actualmente, en el desarrollo de proyectos de genómica estructural de alto rendimiento. De un modo más concreto, la invención se encuadraría en los procesos de producción y purificación de proteínas de fusión.

## Estado de la técnica

El exitoso desarrollo de proyectos de genómica ocurrido en los últimos años ha permitido obtener una vasta cantidad de información genética, especialmente de organismos procariotas, en donde el número de genomas secuenciados se contará por miles en pocos años (Abby, S. & Daubin, V. (2007). *Trends Microbiol.* **15**, 135-141). Un factor decisivo de este éxito ha sido, sin duda alguna, la homogeneidad química y estructural de los ácidos nucleicos, lo cual ha permitido desarrollar protocolos de secuenciación de alto rendimiento (“*high throughput*”) de carácter universal. Actualmente, en la era post-genómica, el interés se dirige principalmente al análisis funcional y estructural de las proteínas codificadas por dichos genomas. En este sentido, y a diferencia de los ácidos nucleicos, las proteínas son química y estructuralmente mucho más heterogéneas, lo cual supone de *facto* la imposibilidad de desarrollar protocolos de alto rendimiento universales para la purificación y caracterización estructural de estas moléculas (Waugh, D.S. (2005). *Trends Biotechnol.* **23**, 316-320). Otros problemas relevantes asociados al diseño de este tipo de procesos de alto rendimiento es el bajo rendimiento en la producción y la escasa solubilidad de muchas de las proteínas sobre-expresadas (Esposito, D. & Chatterjee, D.K. (2006). *Curr. Opin. Biotechnol.* **17**, 353-358). Esto último ha conducido a la búsqueda de sistemas de sobre-expresión de proteínas recombinantes alternativos al que puede considerarse el sistema de expresión por excelencia: la bacteria *Escherichia coli*. Algunos sistemas considerados están basados en células eucariotas de mamífero, levaduras o de insectos (Greene, J.J. (2004). *Methods Mol. Biol.* **267**, 3-14; Daly, R. & Hearn, M.T. (2005). *J. Mol. Recognit.* **18**, 119-138; Kost, T.A., Condreay, J.P. & Jarvis, D.L. (2005). *Nat. Biotechnol.* **23**, 567-575) o bien sistemas libres de células (Murthy, T.V., Wu, W., Qiu, Q.Q., Shi, Z. Labaer, J. & Brizuela, L. (2004). *Prot. Expr. Purif.* **36**, 217-225). No obstante, el bajo coste, la facilidad de manejo y de escalado de cultivos siguen haciendo de *E. coli* el sistema de sobre-expresión prácticamente universal (Esposito, D. & Chatterjee, D.K. (2006). *Curr. Opin. Biotechnol.* **17**, 353-358).

Una aproximación empleada actualmente en el desarrollo de protocolos de alto rendimiento universales de producción/purificación de proteínas en *E. coli* es el uso de etiquetas de afinidad (“*affinity tag*”) (Waugh, D.S. (2005). *Trends Biotechnol.* **23**, 316-320; Arnau, J., Lauritzen, C., Petersen, G.E. & Petersen, J. (2006). *Prot. Express. Purif.* **48**, 1-13). Estas etiquetas son proteínas o péptidos que, fusionados a la proteína recombinante de interés, idealmente permitirían aumentar el rendimiento de su expresión, incrementar su solubilidad y desarrollar un protocolo simple y eficaz de purificación. Se puede afirmar que actualmente no existe una etiqueta de afinidad que reúna estas características de un modo general o universal; de hecho, se considera que el uso combinado de varias etiquetas podría ser una posible solución al diseño de estrategias generales (Waugh, D.S. (2005). *Trends Biotechnol.* **23**, 316-320). Obviamente, el empleo sistemático de esta aproximación combinatorial requiere de un conocimiento preciso de las ventajas e inconvenientes de las correspondientes etiquetas individuales.

En la Tabla I y Tabla III se muestran las principales etiquetas de afinidad (“*affinity tag*”) pues permiten desarrollar protocolos de purificación genéricos basados en cromatografías de afinidad, mientras que en la Tabla II, aparecen etiquetas que en algunos casos se ha comprobado actúan como “agentes solubilizantes” (“*solubility enhancers*”).

(Tabla pasa a página siguiente)

# ES 2 327 309 B1

TABLA I  
Principales etiquetas de afinidad

5

Etiqueta	Proteína	Matriz	Referencia
His6	Etiqueta de seis histidinas	Metales inmovilizados	1
GST	Glutathion-S-transferase	Glutation	2
MBP	Maltose-binding protein	Amilosa	3
FLAG	Péptido FLAG	Anticuerpo anti-FLAG	4
BAP	Péptido aceptor de biotina	Avidita	5
Strep II	Péptido de unión a streptavidina	Streptavidina	6
CBP	Péptido de unión a calmodulina	Calmodulina	7

10

15

20

25

30

35

40

45

50

55

60

65

1. Gaberc-Porekar, V. & Menart, V. (2001). *J. Biochem. Biophys. Methods* **49**, 335-360.
2. Smith, D. B. & Johnson, K. S. (1988). *Gene* **67**, 31-40.
3. Di Guan, C., Li, P., Riggs, P. D. & Inoue, H. (1998). *Gene* **67**, 21-30.
4. Einhauer, A. & Jungbauer, A. (2001). *J. Biochem. Biophys. Methods* **49**, 455-465.
5. Schatz, P. J. (1993). *Biotechnology (NY)* **11**, 1138-1143.
6. Voss, S. & Skerra, A. (1997). *Protein Eng.* **10**, 975-982.
7. Vaillancourt, P., Zheng, C. F., Hoang, D. Q. & Breister, L. (2000) *Methods Enzymol.* **326**, 340-362.

TABLA II

*Principales etiquetas solubilizantes*

5

<b>Etiqueta</b>	<b>Proteína</b>	<b>Origen</b>	<b>Referencia</b>
MBP	Maltose-binding protein	<i>Escherichia coli</i>	8
GST	Glutathione-S-transferase	<i>Schischosoma japonicum</i>	9
Trx	Thioredoxin	<i>Escherichia coli</i>	10
NusA	N-Utilization substance	<i>Escherichia coli</i>	11
SUMO	Small Ubiquitin-modifier	<i>Homo sapiens</i>	12
SET	Solubility-enhancing tag	Sintético	13

30

8. Nallamsetty, S. & Waugh, D. S. (2006). *Prot. Expr. Purif.* **45**, 175-182.

35

9. Nygren, P. A., Stahl, S. & Uhlen, M. (1993). *Trends Biotechnol.* **12**, 184-188.

10. LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F. & McCoy, J. M. (1993). *Biotechnology (NY)*. **11**, 187-193.

40

11. Davis, G. D., Elisee, C., Newham, D. M. & Harrison, R. G. (1999). *Biotechnol. Bioeng.* **65**, 382-388.

12. Marblestone, J. G., Edavettal., S. C., Lim, Y., Lim, P., Zuo, X. & Butt, T. R. (2006). *Prot. Sci.* **15**, 182-189.

45

13. Zhang, Y. B., Howitt, J., McCorkle, S., Lawrence, P., Springer, K. & Freimuth, P. (2004). *Prot. Expr. Purif.* **36**, 207-216.

50

55

60

65

Hoy día el interés se centra principalmente en MBP, GST, NusA y Trx. No obstante y teniendo en cuenta, por un lado, que estudios recientes demuestran una baja eficacia de GST (Dyson, M.R., Shadbolt, S.P., Vincent, K.J., Perera, R.L. & McCafferty, J. (2004). *BMC Biotechnol.* **4**, 32) y Trx (Marblestone, J.G., Edavettal., S.C., Lim, Y., Lim, P., Zuo, X. & Butt, T.R. (2006). *Prot. Sci.* **15**, 182-189) como agentes solubilizantes en *E. coli* y, por otro, que NusA, aun con una eficacia similar a MBP como agente solubilizante, no permite desarrollar un protocolo de purificación inmediato (Davis, G.D., Elisee, C., Newham, D.M. & Harrison, R.G. (1999). *Biotechnol. Bioeng.* **65**, 382-388; Dummler, A., Lawrence, A.M. & de Marco, A. (2005). *Microb. Cell Fact.* **4**, 34), MBP ha devenido una de las etiquetas de afinidad sobre la que más interés se tiene y una de las más ampliamente estudiadas (Dyson, M.R., Shadbolt, S.P., Vincent, K.J., Perera, R.L. & McCafferty, J. (2004). *BMC Biotechnol.* **4**, 32; Braud, S., Moutiez, M., Belin, P., Abello, N., Drevet, P., Zinn-Justin, S., Courcon, M., Masson, C., Dassa, J. & Charbonnier, J.B. (2005). *J. Proteome Res.* **4**, 2137-2147). Asimismo, puede afirmarse que la eficacia de estas etiquetas es función de la proteína “acompañante” (Dyson, M.R., Shadbolt, S.P., Vincent, K.J., Perera, R.L. & McCafferty, J. (2004). *BMC Biotechnol.* **4**, 32), lo cual no es sino reflejo de la diversidad química y estructural de las proteínas. Indudablemente, una validación cualitativa de las etiquetas de afinidad requerirá de una aproximación a gran escala o “proteómica” (Esposito, D. & Chatterjee, D.K. (2006). *Curr. Opin. Biotechnol.* **17**, 353-358).

TABLA III

*Principales etiquetas de afinidad*

Etiqueta	Proteína	Matriz	Referencia
His6	Etiqueta de seis histidinas	Metales inmovilizados	7
GST	Glutathion-S-transferase	Glutation	8
MBP	Maltose-binding protein	Amilosa	9
FLAG	Péptido FLAG	Anticuerpo anti-FLAG	10
BAP	Péptido aceptor de biotina	Avidina	11
Strep II	Péptido de unión a streptavidina	Streptavidina	12
CBP	Péptido de unión a calmodulina	Calmodulina	13

14. Gaberc-Porekar, V. & Menart, V. (2001). *J. Biochem. Biophys. Methods* **49**, 335-360.

15. Smith, D. B. & Johnson, K. S. (1988). *Gene* **67**, 31-40.

16. Di Guan, C., Li, P., Riggs, P. D. & Inoue, H. (1998). *Gene* **67**, 21-30.

17. Einhauer, A. & Jungbauer, A. (2001). *J. Biochem. Biophys. Methods* **49**, 455-465.

18. Schatz, P. J. (1993). *Biotechnology (NY)* **11**, 1138-1143.

19. Voss, S. & Skerra, A. (1997). *Protein Eng.* **10**, 975-982.

20. Vaillancourt, P., Zheng, C. F., Hoang, D. Q. & Breister, L. (2000) *Methods Enzymol.* **326**, 340-362.

Una etiqueta de afinidad ideal aumenta el rendimiento en la producción de proteínas (Waugh, D.S. (2005). *Trends Biotechnol.* **23**, 316-320). Este efecto está asociado normalmente con etiquetas proteicas (no péptidos pequeños) situadas en el extremo N-terminal de la fusión, en donde la etiqueta proporcionaría un entorno eficaz para iniciar el proceso de traducción, evitando que el ARNm adoptara estructuras secundarias que pudiesen interferir en la unión a los ribosomas (Hartz, D., McPheeters, D.S., Green, L. & Gold, L. (1991). *J. Mol. Biol.* **218**, 99-105). Por otro lado, también deben permitir desarrollar un protocolo de purificación eficaz, potencialmente escalable a procesos de alto rendimiento con costes mínimos (Waugh, D.S. (2005). *Trends Biotechnol.* **23**, 316-320). En este sentido, la matriz cromatográfica de afinidad cobra especial relevancia (ver Tabla III), distinguiéndose dos grandes grupos: matrices con proteínas inmovilizadas, que suelen ser caras, con baja capacidad de unión y vidas medias cortas (FLAG, BAP, StrepII, CBP), y matrices con ligandos de pequeño tamaño (MBP, GST), más baratas y estables. Por otro lado, la disociación de la proteína inmovilizada en la matriz debe producirse en condiciones no "agresivas" para el estado nativo de la proteína.

Claramente, cualquier estrategia basada en el uso de etiquetas de afinidad debería proporcionar la posibilidad de eliminarla de un modo eficaz (Waugh, D.S. (2005). *Trends Biotechnol.* **23**, 316-320), aspecto éste especialmente relevante cuando se trata de proteínas con potenciales aplicaciones en terapia humana (Arnau, J., Lauritzen, C., Petersen, G.E. & Petersen, J. (2006). *Prot. Express. Purif.* **48**, 1-13). La disponibilidad de endoproteasas con una especificidad exquisita como AcTEV® (tobacco etch virus; Invitrogen) o PreScission® (Amersham Biotech) han mitigado notablemente este problema. Puesto que los determinantes de la especificidad de estas endoproteasas se localizan en el lado C-terminal del enlace peptídico diana (TEV: ENLYFQ/G; PreScission: LEVLFQ/GP), tras la hidrólisis, estos permanecerán en las etiquetas si están situadas en el extremo N-terminal de la proteína de fusión.

Finalmente, de acuerdo a las expectativas de aumento de inversiones en el mercado farmacéutico en fármacos de naturaleza proteica para el periodo 2004-2010 (de 34000 a 52000 millones de dólares; Pavlou, A.K. & Reichert, J.M. (2004). *Nat. Biotechnol.* **22**, 1513-1519), se espera la aparición de nuevas proteínas de interés terapéutico, de ahí que la búsqueda de nuevas etiquetas de afinidad sea actualmente un campo de intensa investigación (Arnaud, J., Lauritzen, C., Petersen, G.E. & Petersen, J. (2006). *Prot. Express. Purif.* **48**, 1-13).

En resumen, la puesta a punto de un procedimiento de alto rendimiento de carácter general de producción y/o purificación de proteínas se enfrenta al problema de la heterogeneidad química y estructural de estas moléculas. Concretamente, los tres principales desafíos a los que se intenta aportar soluciones mediante el uso de las etiquetas de afinidad y solubilidad son:

- Aumento del rendimiento de la producción de las proteínas heterólogas,
- Aumento de la solubilidad de la proteína de interés, y
- Aportar un protocolo de purificación eficaz y fácilmente escalable a procesos de alto rendimiento.

## Descripción de la invención

### Breve descripción

Un aspecto de la invención lo constituye una proteína de fusión recombinante, en adelante proteína de fusión de la invención, que comprende, al menos, una secuencia del siguiente grupo:

- a) la secuencia de aminoácidos del péptido LSL<sub>150</sub> del hongo *Laetiporus sulphureus* (SEQ ID NO2), y
- b) una secuencia análoga a la SEQ ID NO2 de a).

En el sentido utilizado en esta descripción, el término “análoga” pretende incluir a cualquier secuencia de aminoácidos que pueda ser aislada o construida en base a la secuencia de aminoácidos mostrada en la SEQ. ID NO2, por ejemplo, mediante la introducción de sustituciones de aminoácidos conservativas o no conservativas, incluyendo la inserción de uno o más aminoácidos, la adición de uno o más aminoácidos en cualquiera de los extremos de la molécula o la delección de uno o más aminoácidos en cualquier extremo o en el interior de la secuencia.

En general, una secuencia de aminoácidos análoga es sustancialmente homóloga a la secuencia de aminoácidos identificada como la SEQ ID NO2. En el sentido utilizado en esta descripción, la expresión “sustancialmente homóloga” significa que las secuencias de aminoácidos en cuestión tienen un grado de identidad, a nivel de aminoácidos, de, al menos, un 60%, preferentemente de, al menos un 85%, o más preferentemente de, al menos, un 95%.

Otro aspecto de la invención lo constituye el procedimiento de obtención de la proteína de fusión, en adelante procedimiento de la invención, que comprende las siguientes etapas:

- i) Obtención de una construcción génica que comprende la secuencia de nucleótidos de LSL<sub>150</sub> (SEQ ID NO1) y la secuencia que codifica para la proteína o péptido de interés, y del vector de expresión que lo comprende,
- ii) Transformación de una célula huésped, preferentemente procariota, más preferentemente de *Escherichia coli*, en la que la construcción génica o el vector de expresión de i) bajo control de los promotores adecuados, permite la expresión de la proteína o péptido de fusión de interés, y purificación del extracto crudo de las células,
- iii) Inmovilización de la proteína o péptido de fusión en matrices derivadas de agarosa mediante el paso de extracto crudo de ii) a través de la matriz,
- iv) Disociación de la proteína o péptido de fusión de la matriz cromatográfica mediante lavado de la matriz con un medio conteniendo uno o varios azúcares competidores,

y, opcionalmente

- v) Separación del péptido SEQ ID NO2 de la proteína de fusión mediante tratamiento proteolítico adecuado.

Para llevar a cabo el paso i) del procedimiento de la invención es necesario elaborar una construcción genética que comprenda la secuencia de ADN codificante del péptido SEQ ID NO2, la proteína o péptido de interés y opcionalmente, una secuencia de reconocimiento de endoproteasas, por ejemplo, TEV (tobacco etch virus endoprotease) con objeto de separar la etiqueta de afinidad de la proteína de interés. Así, otro aspecto de la invención lo constituye

una construcción genética codificante de la secuencia de ADN codificante de la proteína de fusión de la invención, *en adelante construcción genética de la invención*, que comprende, al menos, una secuencia del siguiente grupo:

- a) la secuencia de ADN SEQ ID NO1, y
- b) una secuencia homóloga a la SEQ ID NO1.

La construcción genética de la invención permite la expresión de una proteína de interés unida al péptido de SEQ ID NO2. Esta construcción genética de la invención, también puede comprender, en caso necesario y para permitir un mejor aislamiento, detección o secreción al exterior de la célula del péptido de interés expresado, una secuencia de nucleótidos que codifica para un péptido susceptible de ser utilizado con fines de aislamiento, detección o secreción de dicho péptido. Por tanto, un objeto particular de la presente invención lo constituye una construcción genética de la invención que comprende cualquier otra secuencia de nucleótidos codificante de un péptido o secuencia peptídica que permita el aislamiento, la detección o la secreción al exterior de la célula del péptido expresado, por ejemplo, a título ilustrativo y sin que limite el alcance de la invención, una secuencia de polihistidina (6xHis), una secuencia peptídica reconocible por un anticuerpo monoclonal (por ejemplo, para su identificación, o cualquier otra que sirva para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad: péptidos etiqueta tales como c-myc, HA, E-tag) (Using antibodies: a laboratory manual. Ed. Harlow and David Lane (1999). Cold Spring Harbor Laboratory Press. New York. Capítulo: Tagging proteins. Pp. 347-377).

La construcción genética de la invención descrita previamente puede aislarse y obtenerse por un experto mediante el empleo de técnicas ampliamente conocidas en el estado de la técnica (Sambrook *et al.* "Molecular cloning", a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory Press, N.Y., 1989 vol 1-3). Dichas secuencias de nucleótidos pueden estar integradas en un vector de expresión genética que permite la regulación de la expresión de la misma en condiciones adecuadas en el interior de las células.

Por tanto, otro aspecto de la presente invención lo constituye un vector de expresión, *en adelante vector de expresión de la invención*, que comprende una construcción genética de la invención y que permite la expresión de una proteína o péptido. Ejemplos de una realización particular se pueden encontrar en los ejemplos de la presente invención.

En general, un vector de expresión comprende, además de la construcción genética descrita en la invención, un promotor que dirige su transcripción (por ejemplo, pT7, plac, ptrc, ptac, pBAD, ret, etc.), al que está operativamente enlazado, y otras secuencias necesarias o apropiadas que controlan y regulan dicha transcripción y, en su caso, la traducción del producto de interés, por ejemplo, señales de inicio y terminación de transcripción (tlt2, etc.), señal de poliadenilación, origen de replicación, secuencias de unión a ribosomas (RBS), secuencias codificantes de reguladores transcripcionales, (enhancers), silenciadores transcripcionales (silencers), represores, etc. Ejemplos de vectores de expresión apropiados pueden seleccionarse de acuerdo con las condiciones y necesidades de cada caso concreto entre plásmidos de expresión, vectores virales (DNA o RNA), cósmidos, cromosomas artificiales, etc. que pueden contener, además, marcadores utilizables para seleccionar las células transfectadas o transformadas con el gen o genes de interés. La elección del vector dependerá de la célula huésped y del tipo de uso que se quiera realizar. Por tanto, según un modo de realización particular de la presente invención dicho vector es un plásmido o un vector viral. La obtención de dicho vector puede realizarse por métodos convencionales conocidos por los técnicos en la materia al igual que para la transformación de microorganismos y células eucariotas se pueden utilizar diferentes métodos ampliamente conocidas - transformación química, electroporación, microinyección, etc. - descritos en diversos manuales [Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.].

Otro aspecto de la presente invención lo constituye una célula huésped, *en adelante célula de la invención*, que comprende la construcción genética o el vector de expresión de la invención, preferentemente una célula procarionota.

Una realización particular de la invención lo constituye una célula huésped de la invención constituida por una célula *E. coli.*, preferentemente células *E. coli* BL21 DE3.

### Descripción detallada

La presente invención se refiere a un nuevo método general o universal de expresión y purificación de proteínas de fusión conteniendo un dominio lectina  $\beta$ -trébol como etiqueta de afinidad.

La presente invención se basa en que los inventores han observado que el módulo lectina de la proteína hemolítica del hongo *Laetiporus sulphureus* (los primeros 150 aminoácidos: LSL<sub>150</sub>) (SEQ ID NO2) es funcional en tanto que parte constituyente de numerosas proteínas recombinantes de fusión diseñadas *ad hoc*. Es decir, que por un lado, la SEQ ID NO2 es versátil desde una perspectiva de plegamiento de proteínas pues se pliega adecuadamente estando acompañada de diferentes proteínas en su extremo C-terminal y, por otro, conserva su capacidad de unión de azúcares. Sorprendentemente, las proteínas recombinantes de interés fusionadas al extremo C-terminal de SEQ ID NO2 también se pliegan adecuadamente, según se deduce de los resultados de actividad llevados a cabo con las

proteínas puras. Además, estas proteínas de interés mantienen su estado nativo incluso una vez separadas de SEQ ID NO2 mediante tratamiento proteolítico (TEV proteasa; ver más abajo). Todo ello ha permitido purificar diversas proteínas de fusión conteniendo la SEQ ID NO2 en su extremo N-terminal empleando un único protocolo de purificación.

5 En este sentido, los inventores han observado que la SEQ ID NO2 (de origen eucariota) se produce en *E. coli* (procariota) con un elevado rendimiento (aprox. 80 mg por litro de cultivo), plegándose adecuadamente. De hecho, la estructura tridimensional de SEQ ID NO2 producida en *E. coli* que han determinado es esencialmente igual a la del correspondiente módulo lectina de LSL aislada directamente del hongo *L. sulphureus* (Figura 3). Esta estructura es garantía del plegamiento adecuado de SEQ ID NO2 expresada heterológicamente en un organismo procariota, lo cual no es siempre obvio. Sorprendentemente, en todos los ejemplos indicados más abajo, el rendimiento con el que se obtienen las proteínas de fusión construidas (ver ejemplos más abajo) es tal que permite afrontar proyectos de cristalización de las proteínas puras, una vez separadas de SEQ ID NO2, con un coste bajo. En el caso concreto de la proteína PLD de *A. haemolyticum* (ver ejemplo 3), su expresión con una etiqueta de histidinas (His6 tag) en las mismas condiciones que LSL<sub>150</sub>-PLD, no permitía obtener la proteína de fusión soluble. Es decir, la secuencia SEQ ID NO2 reúne propiedades de etiqueta de afinidad y de solubilidad. En este sentido, es de destacar que actualmente se desconocen las bases moleculares por las cuales una proteína actúa como etiqueta de solubilidad y que, por ello, no disponemos de criterios objetivos con los cuales realizar una búsqueda sistemática de proteínas que pudiesen actuar como tales. En nuestro caso particular, podemos afirmar que *del análisis estructural y biofísico de la secuencia SEQ ID NO2 no se pueden deducir sus propiedades como etiqueta de solubilidad*. La idea inicial de que cualquier proteína muy soluble puede actuar como tal se ha comprobado incorrecta (Waugh, D.S. (2005). *Trends Biotechnol.* **23**, 316-320). Aunque la solubilidad es una propiedad necesaria para una potencial etiqueta de solubilidad, es obvio que no es un requisito suficiente.

25 En tanto que etiqueta de afinidad y de solubilidad, la SEQ ID NO2 permite producir diferentes proteínas de fusión con elevados rendimientos y establecer un protocolo de purificación de las mismas de carácter general, eficaz, sencillo y de bajo coste. Estas características se deben a factores de distinta naturaleza. En primer lugar, se debe a que el método de purificación de la proteína de fusión implica una única etapa cromatográfica de afinidad. En segundo lugar, a que el costo de los principales componentes empleados en dicho protocolo, esto es, las matrices cromatográficas derivadas de agarosa (Sephacrose™ 4B, 6B, CL-4B, CL-6B) y eluyentes (D(+)-lactosa), es sorprendentemente bajo. Así, y a modo de comparación, mientras que el precio medio de estas resinas está en torno a unos 420-580 €/litro (en el caso de Sephacrose™ 6B se emplea una columna con unos 25 ml de resina) y el de D(+)-lactosa como eluyente unos 28 €/kilo, el precio de los componentes empleados en el sistema descrito en (Tielker, D., Rosenau, F., Bartels, K.M., Rosenbaum, T. & Jaeger, K.E. (2006). *Bio-Techniques* **41**, 327-332), esto es, manosa-agarosa es de 33120 €/litro y el de D-manosa de unos 900 €/kilo. Y, finalmente, en tercer lugar, las matrices cromatográficas empleadas, por una parte, no requieren de ningún tratamiento químico adicional para actuar como matrices de afinidad frente a la secuencia SEQ ID NO2 y, por otra, al ser esencialmente inertes a los homogeneizados celulares y no requerir regeneraciones más allá de su lavado con buffer, pueden reutilizarse de un modo esencialmente continuo.

40 Una vez comprobada la capacidad de la secuencia SEQ ID NO2 para actuar como etiqueta de afinidad y solubilidad, se puede afirmar *a posteriori* que ésta reúne características ventajosas adicionales para ser empleada biotecnológicamente en procesos de producción y purificación de proteínas. En primer lugar, al ser un dominio monomérico no impone restricciones estéricas sobre la proteína C-terminal “acompañante”; en segundo lugar, al carecer de residuos de cisteína la secuencia SEQ ID NO2 no demanda el uso de agentes moduladores del estado redox del medio y, finalmente, en tercer lugar, esta secuencia es LSL<sub>150</sub> es desde una perspectiva evolutiva un módulo proteico, por lo que está adaptada estructural y funcionalmente a ser parte constituyente de proteínas modulares, como son *de facto* las proteínas recombinantes de fusión. En este sentido es destacable el hecho de que ninguna de las etiquetas de afinidad (proteínas) actualmente relevantes, como MBP, GST o Trx, proceden de proteínas modulares.

50 Por tanto, un aspecto de la invención lo constituye una proteína de fusión recombinante, en adelante proteína de fusión de la invención, que comprende, al menos, una secuencia del siguiente grupo:

- a) la secuencia de aminoácidos del péptido LSL<sub>150</sub> del hongo *Laetiporus sulphureus* (SEQ ID NO2), y
- b) una secuencia análoga a la SEQ ID NO2 de a).

60 En el sentido utilizado en esta descripción, el término “análoga” pretende incluir a cualquier secuencia de aminoácidos que pueda ser aislada o construida en base a la secuencia de aminoácidos mostrada en la SEQ. ID NO2, por ejemplo, mediante la introducción de sustituciones de aminoácidos conservativos o no conservativos, incluyendo la inserción de uno o más aminoácidos, la adición de uno o más aminoácidos en cualquiera de los extremos de la molécula o la delección de uno o más aminoácidos en cualquier extremo o en el interior de la secuencia.

65 En general, una secuencia de aminoácidos análoga es sustancialmente homóloga a la secuencia de aminoácidos identificada como la SEQ ID NO2. En el sentido utilizado en esta descripción, la expresión “sustancialmente homóloga” significa que las secuencias de aminoácidos en cuestión tienen un grado de identidad, a nivel de aminoácidos, de, al menos, un 60%, preferentemente de, al menos un 85%, o más preferentemente de, al menos, un 95%.



## ES 2 327 309 B1

Otro aspecto de la invención lo constituye el procedimiento de obtención de la proteína de fusión, en adelante procedimiento de la invención, que comprende las siguientes etapas:

- 5 i) Obtención de una construcción génica que comprende la secuencia de nucleótidos de **LSL**<sub>150</sub> (SEQ ID NO1) y la secuencia que codifica para la proteína o péptido de interés, y del vector de expresión que lo comprende,
- 10 ii) Transformación de una célula huésped, preferentemente procariota, más preferentemente de *Escherichia coli*, en la que la construcción génica o el vector de expresión de i) bajo control de los promotores adecuados, permite la expresión de la proteína o péptido de fusión de interés, y purificación del extracto crudo de las células,
- 15 iii) Inmovilización de la proteína o péptido de fusión en matrices derivadas de agarosa mediante el paso de extracto crudo de ii) a través de la matriz,
- iv) Disociación de la proteína o péptido de fusión de la matriz cromatográfica mediante lavado de la matriz con un medio conteniendo uno o varios azúcares competidores,

y, opcionalmente

- 20 v) Separación del péptido SEQ ID NO2 de la proteína de fusión mediante tratamiento proteolítico adecuado.

Para llevar a cabo el paso i) del procedimiento de la invención es necesario elaborar una construcción genética que comprenda la secuencia de ADN codificante del péptido SEQ ID NO2, la proteína o péptido de interés y opcionalmente, una secuencia de reconocimiento de endoproteasas, por ejemplo, TEV (*tobacco etch virus endoprotease*) con objeto de separar la etiqueta de afinidad de la proteína de interés. Así, otro aspecto de la invención lo constituye una construcción genética codificante de la secuencia de ADN codificante de la proteína de fusión de la invención, en adelante construcción genética de la invención, que comprende, al menos, una secuencia del siguiente grupo:

- 30 a) la secuencia de ADN SEQ ID NO1, y
- b) una secuencia homóloga a la SEQ ID NO1.

La construcción genética de la invención permite la expresión de una proteína de interés unida al péptido de SEQ ID NO2. Esta construcción genética de la invención, también puede comprender, en caso necesario y para permitir un mejor aislamiento, detección o secreción al exterior de la célula del péptido de interés expresado, una secuencia de nucleótidos que codifica para un péptido susceptible de ser utilizado con fines de aislamiento, detección o secreción de dicho péptido. Por tanto, un objeto particular de la presente invención lo constituye una construcción genética de la invención que comprende cualquier otra secuencia de nucleótidos codificante de un péptido o secuencia peptídica que permita el aislamiento, la detección o la secreción al exterior de la célula del péptido expresado, por ejemplo, a título ilustrativo y sin que limite el alcance de la invención, una secuencia de polihistidina (6xHis), una secuencia peptídica reconocible por un anticuerpo monoclonal (por ejemplo, para su identificación, o cualquier otra que sirva para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad: péptidos etiqueta tales como c-myc, HA, E-tag) (Using antibodies: a laboratory manual. Ed. Harlow and David Lane (1999). Cold Spring Harbor Laboratory Press. New York. Capítulo: Tagging proteins. Pp. 347-377).

La construcción genética de la invención descrita previamente puede aislarse y obtenerse por un experto mediante el empleo de técnicas ampliamente conocidas en el estado de la técnica (Sambrook *et al.* "Molecular cloning", a Laboratory Manual 2nd ed., Cold Spring Harbor Laboratory Press, N.Y., 1989 vol 1-3). Dichas secuencias de nucleótidos pueden estar integradas en un vector de expresión genética que permite la regulación de la expresión de la misma en condiciones adecuadas en el interior de las células.

Por tanto, otro aspecto de la presente invención lo constituye un vector de expresión, en adelante vector de expresión de la invención, que comprende una construcción genética de la invención y que permite la expresión de una proteína o péptido. Ejemplos de una realización particular se pueden encontrar en los ejemplos de la presente invención.

En general, un vector de expresión comprende, además de la construcción genética descrita en la invención, un promotor que dirige su transcripción (por ejemplo, pT7, plac, ptrc, ptac, pBAD, ret, etc.), al que está operativamente enlazado, y otras secuencias necesarias o apropiadas que controlan y regulan dicha transcripción y, en su caso, la traducción del producto de interés, por ejemplo, señales de inicio y terminación de transcripción (tlt2, etc.), señal de poliadenilación, origen de replicación, secuencias de unión a ribosomas (RBS), secuencias codificantes de reguladores transcripcionales, (enhancers), silenciadores transcripcionales (silencers), represores, etc. Ejemplos de vectores de expresión apropiados pueden seleccionarse de acuerdo con las condiciones y necesidades de cada caso concreto entre plásmidos de expresión, vectores virales (DNA o RNA), cósmidos, cromosomas artificiales, etc. que pueden contener, además, marcadores utilizables para seleccionar las células transfectadas o transformadas con el gen o genes de interés. La elección del vector dependerá de la célula huésped y del tipo de uso que se quiera realizar. Por tanto, según un

modo de realización particular de la presente invención dicho vector es un plásmido o un vector viral. La obtención de dicho vector puede realizarse por métodos convencionales conocidos por los técnicos en la materia al igual que para la transformación de microorganismos y células eucariotas se pueden utilizar diferentes métodos ampliamente conocidas - transformación química, electroporación, microinyección, etc. - descritos en diversos manuales [Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.].

Otro aspecto de la presente invención lo constituye una célula huésped, en adelante célula de la invención, que comprende la construcción genética o el vector de expresión de la invención, preferentemente una célula procariota.

Una realización particular de la invención lo constituye una célula huésped de la invención constituida por una célula *E. coli*., preferentemente células *E. coli* BL21 DE3.

### Descripción de las figuras

Figura 1.- *Plásmido pKLSL150 y secuencia de la proteína LSL<sub>150</sub>*. Parte izquierda: esquema del vector de expresión pKLSL150 que contiene la secuencia de 450 nt que codifica para la proteína LSL<sub>150</sub> bajo el control del promotor de T7 y un regulador inducible por IPTG. Este plásmido contiene el gen de resistencia a la Kanamicina (*Kan*), el gen del represor *lacI*, y un sitio múltiple de clonaje (*NcoI/XhoI*). Parte derecha: secuencia de nucleótidos y aminoácidos de la proteína LSL<sub>150</sub> flanqueada por los sitios de restricción *NcoI* y *EcoRI* utilizados en el clonaje.

Figura 2.- *Análisis mediante SDS-PAGE de la purificación de la proteína de fusión LSL<sub>150</sub>*. En el gel aparecen muestras de diferentes fracciones del perfil de elución de LSL<sub>150</sub> al lavar la columna con buffer conteniendo 0.2 M lactosa. A la izquierda se muestran patrones de masa molecular con sus correspondientes masas en kDa.

Figura 3.- *Estructura tridimensional de la proteína LSL<sub>150</sub>*. Parte izquierda: representación del esqueleto peptídico de LSL<sub>150</sub> (negro: hebras  $\beta$ ; gris: bucles). Parte derecha: esqueleto peptídico del módulo lectina de la proteína completa aislada directamente del hongo *Laetiporus sulphureus* (negro: hebras  $\beta$  gris: bucles). N: extremo amino terminal; C: extremo carboxilo terminal. La estructura de ambas proteínas se ha determinado mediante cristalografía de proteínas. La superposición revela una rmsd para 148 carbonos  $\alpha$  de 1.1 Å, esto es, son casi perfectamente superponibles.

Figura 4.- *Plásmido pKLSLt y secuencia de reconocimiento de la TEV proteasa*. Parte superior: esquema del vector de expresión pKLSLt derivado del pKLSL150 (Figura 1). En él se han insertado, en el extremo 3' de la secuencia que codifica para la proteína LSL<sub>150</sub>, las secuencias que codifican para el *linker* y para el sitio de reconocimiento por la TEV proteasa. Se indica asimismo, la situación de un sitio múltiple de clonaje (*EcoRI/XhoI*). Parte inferior: secuencia de nucleótidos y aminoácidos del extremo C-terminal de LSL<sub>150</sub> así como del linker, de la secuencia de corte TEV y de los sitios de restricción.

Figura 5.- *Plásmido pKLSLt-EGFP*. Parte superior: esquema del vector de expresión pKLSLt-EGFP derivado del pKLSLt (Figura 4) en el cual se ha insertado la secuencia que codifica para la *Enhancer Green Fluorescence Protein* (EGFP) de *Aequorea victoria* (Clontech) de 720 pb entre los sitios de restricción *EcoRI* y *Not I*. Parte inferior: secuencia de nucleótidos y aminoácidos del sitio de corte TEV y parte de la proteína clonada.

Figura 6.- *Análisis mediante SDS-PAGE de la expresión y purificación de la proteína de fusión LSL<sub>150</sub>-EGFP y del tratamiento de ésta con la TEV proteasa*. Calles: 1. marcadores de masa molecular (kDa); 2. fracción soluble del homogeneizado celular; 3. proteína LSL<sub>150</sub>-EGFP tras Sepharose™ 4B; 4. tratamiento de LSL<sub>150</sub>-EGFP con TEV proteasa; 5. EGFP tras filtración en gel con Superdex 75; 6. LSL<sub>150</sub> tras filtración en gel con Superdex 75. A la izquierda se muestra la masa molecular en kDa.

Figura 7.- *Plásmido pKLSLt-PLD*. Parte superior: esquema del vector de expresión pKLSLt-PLD derivado del pKLSLt (Figura 4) en el cual se ha insertado la secuencia que codifica para la proteína PLD (fosfolipasa dependiente de esfingomielina del microorganismo *Arcanobacterium haemolyticum*) de 846 pb entre los sitios de restricción *EcoRI* y *Hind III*. Parte inferior: secuencia de nucleótidos y aminoácidos del sitio de corte TEV y parte de la proteína clonada.

Figura 8.- *Análisis mediante SDS-PAGE de la expresión y purificación de la proteína de fusión LSL<sub>150</sub>-PLD y del tratamiento de ésta con la TEV proteasa*. Calles: 1. marcadores de masa molecular (kDa); 2. fracción soluble del homogeneizado celular; 3. proteína LSL<sub>150</sub>-PLD tras Sepharose™ 4B; 4. tratamiento de LSL<sub>150</sub>-PLD con TEV proteasa; 5. PLD tras filtración en gel con Superdex 75; 6. LSL<sub>150</sub> tras filtración en gel con Superdex 75. A la izquierda se muestra la masa molecular en kDa.

Figura 9.- *Plásmido pKLSLt-SMD*. Parte superior: esquema del vector de expresión pKLSLt-SMD derivado del pKLSLt (figura 4) en el cual se ha insertado la secuencia que codifica para la proteína SMD (fosfolipasa dependiente de esfingomielina del microorganismo *Corynebacterium pseudotuberculosis*) de 802 pb entre los sitios de restricción *EcoRI* y *Hind III*. Parte inferior: secuencia de nucleótidos y aminoácidos del sitio de corte TEV y parte de la proteína clonada.

Figura 10.- *Análisis mediante SDS-PAGE de la expresión y purificación de la proteína de fusión LSL<sub>150</sub>-SMD y del tratamiento de ésta con la TEV proteasa*. Calles: 1. marcadores de masa molecular (kDa); 2. fracción soluble

del homogeneizado celular; 3. proteína LSL<sub>150</sub>-SMD tras Sepharose™ 4B; 4. tratamiento de LSL<sub>150</sub>-SMD con TEV proteasa; 5. SMD tras filtración en gel con Superdex 75; 6. LSL<sub>150</sub> tras filtración en gel con Superdex 75. A la izquierda se muestra la masa molecular en kDa.

5 Figura 11.- *Plásmido pKLSLt-CPL7*. Parte superior: esquema del vector de expresión pKLSLt-CPL7 derivado del pKLSLt (figura 4) en el cual se ha insertado la secuencia que codifica para la proteína CPL7 (mureín hidrolasa del fago Cp-7 del microorganismo *Streptococcus pneumoniae*) de 1003 pb entre los sitios de restricción *EcoRI* y *Hind III*. Parte inferior: secuencia de nucleótidos y aminoácidos del sitio de corte TEV y parte de la proteína clonada.

10 Figura 12.- *Análisis mediante SDS-PAGE* de la expresión y purificación de la proteína de fusión LSL<sub>150</sub>-CPL7 y del tratamiento de ésta con la TEV proteasa. Calles: 1. marcadores de masa molecular (kDa); 2. fracción soluble del homogeneizado celular; 3. proteína LSL<sub>150</sub>-CPL7 tras Sepharose™ 4B; 4. tratamiento de LSL<sub>150</sub>-CPL7 con TEV proteasa; 5. CPL7 tras filtración en gel con Superdex 75; 6. LSL<sub>150</sub> tras filtración en gel con Superdex 75. A la izquierda se muestra la masa molecular en kDa.

15 Figura 13.- *Plásmido pKLSLt-IP5\_2K*. Parte superior: esquema del vector de expresión pKLSLt-IP5\_2K derivado del pKLSLt (figura 4) en el cual se ha insertado la secuencia que codifica para la proteína IP5\_2K (Inositol (1,3,4,5,6)-pentakisfosfato quinasa de *Arabidopsis thaliana*) de 1353 pb entre los sitios de restricción *EcoRI* y *Hind III*. Parte inferior: secuencia de nucleótidos y aminoácidos del sitio de corte TEV y parte de la proteína clonada.

20 Figura 14.- *Análisis mediante SDS-PAGE* de la expresión y purificación de la proteína de fusión LSL<sub>150</sub>-IP5\_2K y del tratamiento de ésta con la TEV proteasa. Calles: 1. fracción soluble del homogeneizado celular; 2. marcadores de masa molecular (kDa); 3. proteína LSL<sub>150</sub>-IP5\_2K tras Sepharose™ 4B; 4. tratamiento de LSL<sub>150</sub>-IP5\_2K con TEV proteasa; 5. IP5\_2K tras filtración en gel con Superdex 75. A la izquierda se muestra la masa molecular en kDa.

## 25 Ejemplos de la invención

### Ejemplo 1

30 *Desarrollo del vector con el dominio lectina β-trébol del hongo Laetiporus sulphureus*

Con objeto de caracterizar estructuralmente el dominio lectina β-trébol base de la invención, el gen que codifica para este péptido LSL<sub>150</sub> fue expresado heterológicamente en *E. coli*. Concretamente, el péptido LSL<sub>150</sub> de la invención (SEQ ID NO2) es el módulo lectina de la proteína modular LSL (Mr 35 kDa) producida por el hongo *Laetiporus sulphureus* (Tateno, H. & Goldstein, I.J. (2003). *J. Biol. Chem.* **278**, 40455-40463). LSL<sub>150</sub> está constituido por los 35 150 primeros aminoácidos de LSL (17 kDa), que adoptan una estructura denominada de β-trébol (Mancheño, J.M., Tateno, H., Goldstein, I.J., Martínez-Ripoll, M. & Hermoso, J.A. (2005). *J. Biol. Chem.* **280**, 17251-17259). Este se puede describir como un pequeño barril β formado por 6 hebras β, en uno de cuyos extremos se encuentra el denominado triplete de bucles. La estructura en su conjunto resulta de la triplicación de una unidad estructural básica de unos 50 residuos, cada una de las cuales poseería capacidad de unión de azúcares derivados de galactosa. Por ello, LSL<sub>150</sub> tiene tres potenciales sitios de unión de azúcares, dos de los cuales son competentes para unir disacáridos conteniendo galactosa según se ha demostrado (Mancheño, J.M., Tateno, H., Goldstein, I.J., Martínez-Ripoll, M. & Hermoso, J.A. (2005). *J. Biol. Chem.* **280**, 17251-17259).

45 Previamente, el vector plasmídico pKLSL150 (Figura 1) fue construido insertando la fase de lectura abierta de los primeros 450 nt del gen de la proteína LSL (*LSL*<sub>450</sub>) en el vector pET28a+ (Novagen). El inserto se obtuvo mediante reacción en cadena de la polimerasa (PCR) utilizando como molde el pET43-LSL (Tateno, H. & Goldstein, I.J. (2003). *J. Biol. Chem.* **278**, 40455-40463), así como oligonucleótidos específicos que hibridan en la región inicial y 50 final del gen *LSL*<sub>450</sub> y presentan en sus extremos las secuencias diana de las enzimas de restricción *NcoI* y *EcoRI* respectivamente. Tras la reacción, el fragmento obtenido y el vector pET28a+, fueron digeridos con las correspondientes enzimas de restricción, purificados y ligados en presencia de la enzima T4 ligasa (Roche). Posteriormente, se transformaron bacterias competentes con el producto de la ligación y se seleccionó el plásmido recombinante pKLSL150 siguiendo procedimientos estándares de clonaje (Sambrook, J. & Russell, D.W. (2001). Cold Spring Harbour Lab. Press. 3rd Ed).

55 Para la producción de la proteína LSL<sub>150</sub> recombinante, se transformaron bacterias *E. coli* (BL21 DE3) con el vector de expresión pKLSL150, en donde la expresión del gen que codifica para ella se encuentra bajo el control transcripcional del promotor de T7. Las células de *E. coli* fueron crecidas en 1 L de medio Luria Bertani LB conteniendo kanamicina (50 µg/mL) durante 3-4 horas a 37°C, hasta alcanzarse una OD<sub>600 nm</sub> de 0.6, momento en donde se induce la expresión de la proteína mediante la adición de 0.3 mM IPTG y posterior incubación a 16°C durante 20 horas. Los cultivos celulares se separaron del medio de cultivo mediante centrifugación a 4000 x g durante 15 min. El sedimento así obtenido se resuspende en 25 ml de buffer 20 mM Tris, pH 8.0 conteniendo 0.1 M NaCl y 0.04% azida sódica. Con objeto de romper las células y obtener un homogeneizado celular, el sedimento resuspendido se pasó por una French press. La fracción soluble del homogeneizado celular se obtiene mediante centrifugación a 20000 rpms en un rotor SS34 durante 30 minutos. Esta fracción soluble se aplicó en una columna de Sepharose™ 4B, equilibrada previamente 65 en buffer 20 mM Tris-HCl, pH 8.0 con 0.1 M NaCl y 0.04% azida sódica y mantenida a 4°C. Tras lavado exhaustivo de la columna con el mismo buffer (se alcanza línea base en mediciones de Abs 280 nm), se lleva a cabo la elución de la proteína LSL<sub>150</sub> mediante lavado con el mismo buffer anterior conteniendo lactosa 0.2 M. La columna puede ser

empleada de nuevo sin necesidad de etapas de regeneración adicionales, simplemente tras equilibrarla de nuevo en el buffer de lavado sin azúcar. Debido a que la matriz cromatográfica resulta ser inerte a los homogeneizados celulares su uso continuado está garantizado. La proteína así purificada se ha conseguido con un rendimiento de aproximadamente 50 mg/litro de cultivo. El protocolo de purificación proporciona proteína LSL<sub>150</sub> electroforéticamente pura (Figura 2). Todas las etapas del proceso de purificación se realizan a 4°C. Se ha determinado la estructura tridimensional de LSL<sub>150</sub> recombinante mediante cristalografía de proteínas, encontrándose que es perfectamente superponible (rmsd 1.1 Å para 148 carbonos) al dominio lectina de LSL aislada directamente del hongo *Laetiporus sulphureus* (Figura 3), lo cual es garantía del plegamiento adecuado de esta proteína expresada heterológicamente. El extraordinario rendimiento de la producción de LSL<sub>150</sub>, que apunta a unas muy eficaces etapas de transcripción y traducción, así como la eficacia, sencillez y costes perfectamente asumibles del protocolo de purificación condujeron al planteamiento de la hipótesis de trabajo de que LSL<sub>150</sub> pudiese actuar como etiqueta de afinidad y de solubilidad en la producción de proteínas de fusión. Además, esta molécula posee características muy atractivas como etiqueta de fusión: en primer lugar, es extraordinariamente soluble (se han preparado soluciones de hasta 60-80 mg/ml). Aunque esta característica no es un requisito *suficiente* para considerarse un agente solubilizante, sí que es un requisito *necesario* para serlo; en segundo lugar, es monomérica, por lo que no va a contribuir al establecimiento de interacciones intermoleculares entre las proteínas de fusión de las que forme parte, y en tercer lugar, carece de residuos de cisteína, lo que implica la ausencia de problemas de oxidaciones.

Con objeto de validar esta hipótesis se han construido una serie de fusiones génicas entre la secuencia que codifica para LSL<sub>150</sub> y las secuencias que codifican para distintas proteínas, entre las cuales se ha incorporado la secuencia de reconocimiento de la proteasa TEV. Para ello se construyó el vector plasmídico pKLSLt insertando en el vector pET28a+ (Novagen) la fase de lectura abierta de los primeros 450 nt del gen de la proteína LSL (*LSL*<sub>450</sub>) seguida de la secuencias de nucleótidos que codifican para un linker de unión y el sitio de reconocimiento de la TEV proteasa (*LTEV*). Dicho inserto se obtuvo mediante reacción en cadena de la polimerasa (PCR) utilizando como molde el pET43-LSLa (Tateno, H. & Goldstein, I.J. (2003). *J. Biol. Chem.* 278, 40455-40463) y oligonucleótidos que hibridan en la región inicial y final del gen LSL<sub>450</sub> y presentan en sus extremos las secuencias diana de las enzimas de restricción *NcoI* y la secuencia (*LTEV*) seguida de la secuencia diana para la enzima de restricción *EcoRI*, respectivamente (Figura 4). Tras la reacción, el fragmento obtenido y el vector pET28a+, fueron digeridos con las correspondientes enzimas de restricción, purificados y ligados en presencia de la enzima T4 ligasa (Roche). Posteriormente, se transformaron bacterias competentes con el producto de la ligación y se seleccionó el plásmido recombinante pKLSLt siguiendo procedimientos estándares de clonaje (Sambrook, J. & Russell, D.W. (2001). Cold Spring Harbour Lab. Press. 3rd Ed).

## 35 Ejemplo 2

### *Expresión de la proteína de fusión LSL<sub>150</sub>-EGFP*

#### 40 2.1.- Construcción del vector plasmídico pKLSLt

Con objeto de validar las aplicaciones del dominio SEQ ID NO2 se han construido una serie de fusiones génicas entre la secuencia de ADN que codifica para el dominio LSL<sub>150</sub> (SEQ ID NO1) y las secuencias que codifican para distintas proteínas, entre las cuales se ha incorporado la secuencia de reconocimiento de la proteasa TEV. Para ello se construyó el vector plasmídico pKLSLt insertando en el vector pET28a+ (Novagen) la fase de lectura abierta de los primeros 450 nt del gen de la proteína LSL (*LSL*<sub>450</sub>) seguida de la secuencias de nucleótidos que codifican para un linker de unión y el sitio de reconocimiento de la TEV proteasa (*LTEV*). Dicho inserto se obtuvo mediante reacción en cadena de la polimerasa (PCR) utilizando como molde el pET43-LSLa (Tateno & Goldstein, 2003) y oligonucleótidos que hibridan en la región inicial y final del gen LSL<sub>450</sub> y presentan en sus extremos las secuencias diana de las enzimas de restricción *NcoI* y la secuencia (*LTEV*) seguida de la secuencia diana para la enzima de restricción *EcoRI*, respectivamente (Figura 4). Tras la reacción, el fragmento obtenido y el vector pET28a+, fueron digeridos con las correspondientes enzimas de restricción, purificados y ligados en presencia de la enzima T4 ligasa (Roche). Posteriormente, se transformaron bacterias competentes con el producto de la ligación y se seleccionó el plásmido recombinante pKLSLt siguiendo procedimientos estándares de clonaje (Sambrook y Russell; 2001).

#### 55 2.2.- Construcción del vector plasmídico pKLSLt-EGFP

El vector plasmídico pKLSLt-EGFP (Figura 5) fue construido insertando una fase de lectura abierta de 720 nt del gen de la proteína EGFP (*Enhancer Green Fluorescence Protein*) en el vector pKLSLt (Figura 4). El inserto se obtuvo mediante reacción en cadena de la polimerasa (PCR) utilizando como molde el plásmido pCMV-EGFP (Clontech), así como oligonucleótidos específicos que hibridan en la región inicial y final del gen EGFP y presentan en sus extremos las secuencias diana de las enzimas de restricción *EcoRI* y *NotI*, respectivamente. Tras la reacción, el fragmento obtenido y el vector pKLSLt, fueron digeridos con las correspondientes enzimas de restricción, purificados y ligados en presencia de la enzima T4 ligasa (Roche). Posteriormente, se transformaron bacterias competentes con el producto de la ligación y se seleccionó el plásmido recombinante pKLSLt-EGFP siguiendo procedimientos estándares de clonaje (Sambrook y Russell; 2001).

## ES 2 327 309 B1

### 2.3.- Producción y purificación de la proteína de fusión LSL<sub>150</sub>-EGFP

Se incubaron células de *E. coli* (BL21 DE3) conteniendo el plásmido pKLSLt-EGFP (Figura 5) durante 3-4 horas a 37°C en 1 L de medio LB conteniendo kanamicina (50 µg/mL), hasta alcanzarse una OD<sub>600 nm</sub> de 0.6. En ese momento se induce la expresión de la proteína LSL<sub>150</sub>-EGFP mediante la adición de 0.3 mM IPTG, tras lo cual se mantiene el cultivo a 16°C durante 20 horas. Tras centrifugación a 4000 x g durante 15 min, las células se separaron del medio de cultivo. El sedimento así obtenido se resuspende en 25 ml de buffer 20 mM Tris, pH 8.0 conteniendo 0.1M NaCl y 0.04% azida sódica. Con objeto de romper las células y obtener un homogeneizado celular, el sedimento resuspendido se pasó por una French press. La fracción soluble del homogeneizado celular se obtiene mediante centrifugación a 20000 rpms en un rotor SS34 durante 30 minutos. Esta fracción soluble se aplicó en una columna de Sepharose™ 4B, equilibrada previamente en buffer 20 mM Tris-HCl, pH 8.0 con 0.1M NaCl y 0.04% azida sódica y mantenida a 4°C. Tras lavado exhaustivo de la columna con el mismo buffer (se alcanza línea base en mediciones de Abs 280 nm), se lleva a cabo la elución de LSL<sub>150</sub>-EGFP mediante lavado de la matriz cromatográfica con el mismo buffer anterior conteniendo lactosa 0.2 M (Figura 6). Las fracciones conteniendo LSL<sub>150</sub>-EGFP se juntaron (Figura 6, calle 3), concentrándose la proteína mediante ultrafiltración con membranas Centriprep YM-10 (Amicon). El tratamiento de esta proteína con TEV proteasa en buffer 50 mM Tris, pH 8.0, con 200 mM NaCl, 0.04% azida sódica (80:1; w/w), rinde EGFP y LSL<sub>150</sub> (Figura 6, calle 4) las cuales se resolvieron posteriormente mediante cromatografía de penetrabilidad en Superdex 75 (Figura 6, calles 5 y 6, respectivamente).

### 20 Ejemplo 3

#### Expresión de la proteína de fusión LSL<sub>150</sub>-PLD

##### 25 3.1.- Construcción de los plásmidos bacterianos

El vector plasmídico pKLSLt-PLD (Figura 7) fue construido insertando la fase de lectura abierta de los primeros 846 nt del gen de la fosfolipasa dependiente de esfingomielina (PLD) de la bacteria patógena *Arcanobacterium haemolyticum* en el vector pKLSLt (Figura 4). El inserto se obtuvo mediante reacción en cadena de la polimerasa (PCR) utilizando como molde el plásmido pJB31 (proporcionado por el Dr. Stephen Billington del Department of Veterinary Science and Microbiology, University of Arizona), así como oligonucleótidos específicos que hibridan en la región inicial y final del gen PLD y presentan en sus extremos las secuencias diana de las enzimas de restricción *EcoRI* y *HindIII* respectivamente. Tras la reacción, el fragmento obtenido y el vector pKLSLt, fueron digeridos con las correspondientes enzimas de restricción, purificados y ligados en presencia de la enzima T4 ligasa (Roche). Posteriormente, se transformaron bacterias competentes con el producto de la ligación y se seleccionó el plásmido recombinante pKLSLt-PLD siguiendo procedimientos estándares de clonaje (Sambrook y Russell; 2001).

##### 40 3.2.- Producción y purificación de la proteína de fusión LSL<sub>150</sub>-PLD

Se incubaron células de *E. coli* (BL21 DE3) conteniendo el plásmido pKLSLt-PLD (Figura 7) durante 3-4 horas a 37°C en 1 L de medio LB conteniendo kanamicina (50 µg/mL), hasta alcanzarse una OD<sub>600 nm</sub> de 0.6. En ese momento se induce la expresión de la proteína LSL<sub>150</sub>-PLD mediante la adición de 0.3 mM IPTG, tras lo cual se mantiene el cultivo a 16°C durante 20 horas. Tras centrifugación a 4000 x g durante 15 min, las células se separaron del medio de cultivo. El sedimento así obtenido se resuspende en 25 ml de buffer 20 mM Tris, pH 8.0 conteniendo 0.1 M NaCl y 0.04% azida sódica. Con objeto de romper las células y obtener un homogeneizado celular, el sedimento resuspendido se pasó por una French press. La fracción soluble del homogeneizado celular se obtiene mediante centrifugación a 20000 rpms en un rotor SS34 durante 30 minutos. Esta fracción soluble se aplicó en una columna de Sepharose™ 4B, equilibrada previamente en buffer 20 mM Tris-HCl, pH 8.0 con 0.1 M NaCl y 0.04% azida sódica y mantenida a 4°C. Tras lavado exhaustivo de la columna con el mismo buffer (se alcanza línea base en mediciones de Abs 280 nm), se lleva a cabo la elución de LSL<sub>150</sub>-PLD mediante lavado de la matriz cromatográfica con el mismo buffer anterior conteniendo lactosa 0.2 M (Figura 8). Las fracciones conteniendo LSL<sub>150</sub>-PLD se juntaron (Figura 8, calle 3), concentrándose la proteína mediante ultrafiltración con membranas YM-10 (Amicon). El tratamiento de esta proteína con TEV proteasa en buffer 50 mM Tris, pH 8.0, con 200 mM NaCl, 0.04% azida sódica (80:1; w/w), rinde PLD y LSL<sub>150</sub> (Figura 8, calle 4) las cuales se resolvieron posteriormente mediante cromatografía de penetrabilidad en Superdex 75 (Figura 8, calles 5 y 6, respectivamente).

### 60 Ejemplo 4

#### Expresión de la proteína de fusión LSL<sub>150</sub>-SMD

##### 65 4.1.- Construcción de los plásmidos bacterianos

El vector plasmídico pKLSLt-SMD (Figura 9) fue construido insertando la fase de lectura abierta de 802 nt del gen de la fosfolipasa dependiente de esfingomielina del microorganismo patógeno *Corynebacterium pseudotuberculosis* (SMD) en el vector pKLSLt (Figura 4). El inserto se obtuvo mediante reacción en cadena de la polimerasa

(PCR) utilizando como molde el plásmido pJG90 (Tambourgi *et al.*, 2002), así como oligonucleótidos específicos que hibridan en la región inicial y final del gen EGFP y presentan en sus extremos las secuencias diana de las enzimas de restricción *EcoRI* y *HindIII* respectivamente. Tras la reacción, el fragmento obtenido y el vector pKLSLt, fueron digeridos con las correspondientes enzimas de restricción, purificados y ligados en presencia de la enzima T4 ligasa (Roche). Posteriormente, se transformaron bacterias competentes con el producto de la ligación y se seleccionó el plásmido recombinante pKLSLt-SMD siguiendo procedimientos estándares de clonaje (Sambrook y Russell; 2001).

#### 4.2.- Purificación de la proteína de fusión LSL<sub>150</sub>-SMD

Se incubaron células de *E. coli* (BL21 DE3) conteniendo el plásmido pKLSLt-SMD (Figura 9) durante 3-4 horas a 37°C en 1 L de medio LB conteniendo kanamicina (50 µg/mL), hasta alcanzarse una OD<sub>600 nm</sub> de 0.6. En ese momento se induce la expresión de la proteína LSL<sub>150</sub>-SMD mediante la adición de 0.3 mM IPTG, tras lo cual se mantiene el cultivo a 16°C durante 20 horas. Tras centrifugación a 4000 x g durante 15 min, las células se separaron del medio de cultivo. El sedimento así obtenido se resuspende en 25 ml de buffer 20 mM Tris, pH 8.0 conteniendo 0.1 M NaCl y 0.04% azida sódica. Con objeto de romper las células y obtener un homogeneizado celular, el sedimento resuspendido se pasó por una French press. La fracción soluble del homogeneizado celular se obtiene mediante centrifugación a 20000 rpms en un rotor SS34 durante 30 minutos. Esta fracción soluble se aplicó en una columna de Sepharose™ 4B, equilibrada previamente en buffer 20 mM Tris-HCl, pH 8.0 con 0.1 M NaCl y 0.04% azida sódica y mantenida a 4°C. Tras lavado exhaustivo de la columna con el mismo buffer (se alcanza línea base en mediciones de Abs 280 nm), se lleva a cabo la elución de LSL<sub>150</sub>-SMD mediante lavado de la matriz cromatográfica con el mismo buffer anterior conteniendo lactosa 0.2 M (Figura 10). Las fracciones conteniendo LSL<sub>150</sub>-SMD se juntaron (Figura 10, calle 3), concentrándose la proteína mediante ultrafiltración con membranas YM-10 (Amicon). El tratamiento de esta proteína con TEV proteasa en buffer 50 mM Tris, pH 8.0, con 200 mM NaCl, 0.04% azida sódica (80:1; w/w), rinde SMD y LSL<sub>150</sub> (Figura 10, calle 4) las cuales se resolvieron posteriormente mediante cromatografía de penetrabilidad en Superdex 75 (Figura 10, calles 5 y 6, respectivamente).

#### Ejemplo 5

##### Expresión de la proteína de fusión LSL<sub>150</sub>-CPL7

##### 5.1.- Construcción de los plásmidos bacterianos

El vector plasmídico pKLSLt-CPL7 (Figura 11) fue construido insertando la fase de lectura abierta de los primeros 1003 nt del gen de la proteína mureín hidrolasa CPL7 del bacteriófago Cp-7 de *Streptococcus pneumoniae* en el vector pKLSLt (Figura 4). El inserto se obtuvo mediante reacción en cadena de la polimerasa (PCR) utilizando como molde el plásmido pCP700 (García, P., García, J.L., García, E., Sánchez-Puelles, J.M., López, R. (1990) *Gene* 86, 81-88), así como oligonucleótidos específicos que hibridan en la región inicial y final del gen CPL7 y presentan en sus extremos las secuencias diana de las enzimas de restricción *EcoRI* y *HindIII* respectivamente. Tras la reacción, el fragmento obtenido y el vector pKLSLt, fueron digeridos con las correspondientes enzimas de restricción, purificados y ligados en presencia de la enzima T4 ligasa (Roche). Posteriormente, se transformaron bacterias competentes con el producto de la ligación y se seleccionándose el plásmido recombinante pKLSLt-CPL7 siguiendo procedimientos estándares de clonaje (Sambrook y Russell; 2001).

##### 5.2.- Purificación de la proteína de fusión LSL<sub>150</sub>-CPL7

Se incubaron células de *E. coli* (BL21 DE3) conteniendo el plásmido pKLSLt-CPL7 (Figura 11) durante 3-4 horas a 37°C en 1 L de medio LB conteniendo kanamicina (50 µg/mL), hasta alcanzarse una OD<sub>600 nm</sub> de 0.6. En ese momento se induce la expresión de la proteína LSL<sub>150</sub>-CPL7 mediante la adición de 0.3 mM IPTG, tras lo cual se mantiene el cultivo a 16°C durante 20 horas. Tras centrifugación a 4000 x g durante 15 min, las células se separaron del medio de cultivo. El sedimento así obtenido se resuspende en 25 ml de buffer 20 mM Tris, pH 8.0 conteniendo 0.1 M NaCl y 0.04% azida sódica. Con objeto de romper las células y obtener un homogeneizado celular, el sedimento resuspendido se pasó por una French press. La fracción soluble del homogeneizado celular se obtiene mediante centrifugación a 20000 rpms en un rotor SS34 durante 30 minutos. Esta fracción soluble se aplicó en una columna de Sepharose™ 4B, equilibrada previamente en buffer 20 mM Tris-HCl, pH 8.0 con 0.1 M NaCl y 0.04% azida sódica y mantenida a 4°C. Tras lavado exhaustivo de la columna con el mismo buffer (se alcanza línea base en mediciones de Abs 280 nm), se lleva a cabo la elución de LSL<sub>150</sub>-CPL7 mediante lavado de la matriz cromatográfica con el mismo buffer anterior conteniendo lactosa 0.2 M (Figura 12). Las fracciones conteniendo LSL<sub>150</sub>-CPL7 se juntaron (Figura 12, calle 3), concentrándose la proteína mediante ultrafiltración con membranas YM-10 (Amicon). El tratamiento de esta proteína con TEV proteasa en buffer 50 mM Tris, pH 8.0, con 200 mM NaCl, 0.04% azida sódica (80:1; w/w), rinde CPL7 y LSL<sub>150</sub> (Figura 12, calle 4) las cuales se resolvieron posteriormente mediante cromatografía de penetrabilidad en Superdex 75 (Figura 12, calles 5 y 6, respectivamente).

## Ejemplo 6

*Expresión de la proteína de fusión LSL<sub>150</sub>-IP5\_2K*5 6.1.- *Construcción de los plásmidos bacterianos*

El vector plasmídico pKLSLt-IP5\_2K (Figura 13) fue construido insertando la fase de lectura abierta de los primeros 1353 nt del gen de la proteína Inositol (1,3,4,5,6)-pentakisfosfato quinasa de *Arabidopsis thaliana* en el vector pKLSLt (Figura 4). El inserto se obtuvo mediante reacción en cadena de la polimerasa (PCR) utilizando como molde el plásmido pET28c conteniendo el cDNA de IP5\_2K (proporcionado por la Dr. Beatriz González, Grupo de Cristalografía, Insto. Rocasolano, CSIC), así como oligonucleótidos específicos que hibridan en la región inicial y final del gen IP5\_2K y presentan en sus extremos las secuencias diana de las enzimas de restricción *SacI* y *XhoI* respectivamente. Tras la reacción, el fragmento obtenido y el vector pKLSLt, fueron digeridos con las correspondientes enzimas de restricción, purificados y ligados en presencia de la enzima T4 ligasa (Roche). Posteriormente, se transformaron bacterias competentes con el producto de la ligación, seleccionándose el plásmido recombinante pKLSLt-IP5\_2K siguiendo procedimientos estándares de clonaje (Sambrook y Russell; 2001).

20 6.2.- *Purificación de la proteína de fusión LSL<sub>150</sub>-IP5\_2K*

Se incubaron células de *E. coli* (BL21 DE3) conteniendo el plásmido pKLSLt-IP5\_2K (Figura 13) durante 3-4 horas a 37°C en 1 L de medio LB conteniendo kanamicina (50 µg/mL), hasta alcanzarse una OD<sub>600 nm</sub> de 1. En ese momento se induce la expresión de la proteína LSL<sub>150</sub>-IP5\_2K mediante la adición de 0.3 mM IPTG, tras lo cual se mantiene el cultivo a 16°C durante 16 horas. Tras centrifugación a 4000 x g durante 15 min, las células se separaron del medio de cultivo. El sedimento así obtenido se resuspende en 25 ml de buffer 20 mM Tris, pH 8.0 conteniendo 0.1 M NaCl y 0.04% azida sódica. Con objeto de romper las células y obtener un homogeneizado celular, el sedimento resuspendido se pasó por una French press. La fracción soluble del homogeneizado celular se obtiene mediante centrifugación a 20000 rpm en un rotor SS34 durante 30 minutos. Esta fracción soluble se aplicó en una columna de Sepharose<sup>TM</sup> 4B, equilibrada previamente en buffer 20 mM Tris-HCl, pH 8.0 con 0.1 M NaCl y 0.04% azida sódica y mantenida a 4°C. Tras lavado exhaustivo de la columna con el mismo buffer (se alcanza línea base en mediciones de Abs 280 nm), se lleva a cabo la elución de LSL<sub>150</sub>-IP5\_2K mediante lavado de la matriz cromatográfica con el mismo buffer anterior conteniendo lactosa 0.2 M (Figura 14). Las fracciones conteniendo LSL<sub>150</sub>-IP5\_2K se juntaron (Figura 14, calle 3), concentrándose la proteína mediante ultrafiltración con membranas YM-10 (Amicon). El tratamiento de esta proteína con TEV proteasa en buffer 50 mM Tris, pH 8.0, con 200 mM NaCl, 0.04% azida sódica (80:1; w/w), rinde IP5\_2K y LSL<sub>150</sub> (Figura 14, calle 4) las cuales se resolvieron posteriormente mediante cromatografía de penetrabilidad en Superdex 75 (Figura 14, calles 5).

A modo de comparación, IP5\_2K se ha producido como proteína de fusión con GST (GST-IP5\_2K) y purificado siguiendo protocolos estándar para esta etiqueta afinidad/solubilidad (resultados no mostrados). Al igual que LSL-IP5\_2K, GST-IP5\_2K fue tratada con TEV proteasa, y posteriormente IP5\_2K purificada mediante cromatografía de penetrabilidad en Superdex 75. Los rendimientos finales de ambas purificaciones *llevadas a cabo en paralelo, en condiciones experimentales esencialmente idénticas*, fueron: 23 mg de IP5\_2K partiendo de 4 litros de cultivo con el sistema basado en LSL y de 3.2 mg con el basado en GST.

45

50

55

60

65

REIVINDICACIONES

1. Proteína de fusión recombinante **caracterizada** porque comprende, al menos, una secuencia del siguiente grupo:

- a) la secuencia de aminoácidos del péptido LSL<sub>150</sub> del hongo *Laetiporus sulphureus* (SEQ ID NO2), y
- b) una secuencia análoga a la SEQ ID NO2 de a).

2. Procedimiento de obtención de la proteína de fusión según reivindicación 1 **caracterizado** porque comprende las siguientes etapas:

- i) Obtención de una construcción génica que comprende la secuencia de nucleótidos de LSL<sub>150</sub> (SEQ ID NO1) y la secuencia que codifica para la proteína o péptido de interés, y del vector de expresión que lo comprende,
- ii) Transformación de una célula huésped, preferentemente procariota, más preferentemente de *Escherichia coli*, en la que la construcción génica o el vector de expresión de i) bajo control de los promotores adecuados, permite la expresión de la proteína o péptido de fusión de interés, y purificación del extracto crudo de las células,
- iii) Inmovilización de la proteína o péptido de fusión en matrices derivadas de agarosa mediante el paso de extracto crudo de ii) a través de la matriz,
- iv) Disociación de la proteína o péptido de fusión de la matriz cromatográfica mediante lavado de la matriz con un medio conteniendo uno o varios azúcares competidores,

y, opcionalmente

- v) Separación del péptido SEQ ID NO2 de la proteína de fusión mediante tratamiento proteolítico adecuado.

3. Construcción genética codificante de la secuencia de ADN codificante de la proteína de fusión según reivindicación 1 **caracterizada** porque comprende, al menos, una secuencia del siguiente grupo:

- a) la secuencia de ADN SEQ ID NO1, y
- b) una secuencia homóloga a la SEQ ID NO1.

4. Construcción genética según reivindicación 3 **caracterizada** porque comprende cualquier otra secuencia de nucleótidos codificante de un péptido o secuencia peptídica que permita el aislamiento, la detección o la secreción al exterior de la célula del péptido expresado, por ejemplo, a título ilustrativo y sin que limite el alcance de la invención, una secuencia de polihistidina (6xHis), una secuencia peptídica reconocible por un anticuerpo monoclonal (por ejemplo, para su identificación, o cualquier otra que sirva para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad: péptidos etiqueta tales como c-myc, HA, E-tag).

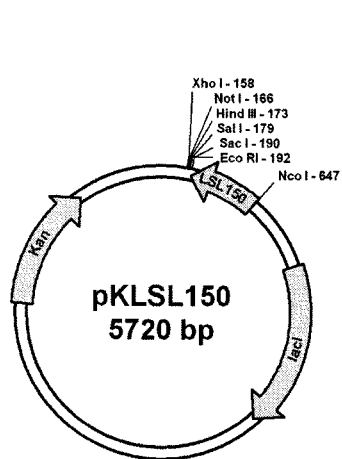
5. Vector de expresión **caracterizado** porque comprende una construcción genética según reivindicaciones 3 y 4 que permite la expresión de una proteína o péptido.

6. Vector de expresión según reivindicación 5 **caracterizado** porque el vector es un plásmido o un vector viral.

7. Célula huésped **caracterizada** porque comprende la construcción genética según reivindicaciones 3 y 4 o el vector de expresión según reivindicaciones 5 y 6, preferentemente una célula procariota.

8. Célula huésped según reivindicación 7 **caracterizada** porque está constituida por una célula *E. coli.*, preferentemente células *E. coli* BL21 DE3.





NcoI  
 ccattggcaatgaccgacatctacaatcccgccggagggtctctacttccgctccttggc  
 M A M T D I Y I P P E G L Y F R L L G  
 ttgtccagtcggcaggtgatcttccgcccgaactctccctctcccgatgttggtctgtct  
 F A S R Q V I F A R N S P S P D V G L S  
 ccggtcaacgaccaggtaccgaccagtacttctcgtctcatctacggcactggagaacac  
 P V N D Q A T D Q Y F S L I Y G T G E H  
 gcggctctctacggcgaataagagcaaaagcgaagggaaggtgctcttctcggctaggcct  
 A G L Y A I K S K A T G K V L F S R R P  
 gggaaacggtatgtggccaaaatcgatggcgacggcggttatcccgacaactggttcaag  
 A E P Y V G Q I D G D G R Y P D N W F K  
 attgagccaggaagacctatctctccaaatatttccggctcgttcagccgtcgaactggc  
 I E P G K T Y L S K Y F R L V Q P S T G  
 accgcgcttgtctcggcaccgcatcttgcagccatctcttggaatcaccctcagactgaa  
 T A L V S R T H L Q P Y F W N H P Q T E  
EcoRI  
 gtcttcgacgaccaatacttcaccttctctctcggaggatcgagaattc  
 V F D D Q Y F T F L F E D - E F

FIG. 1

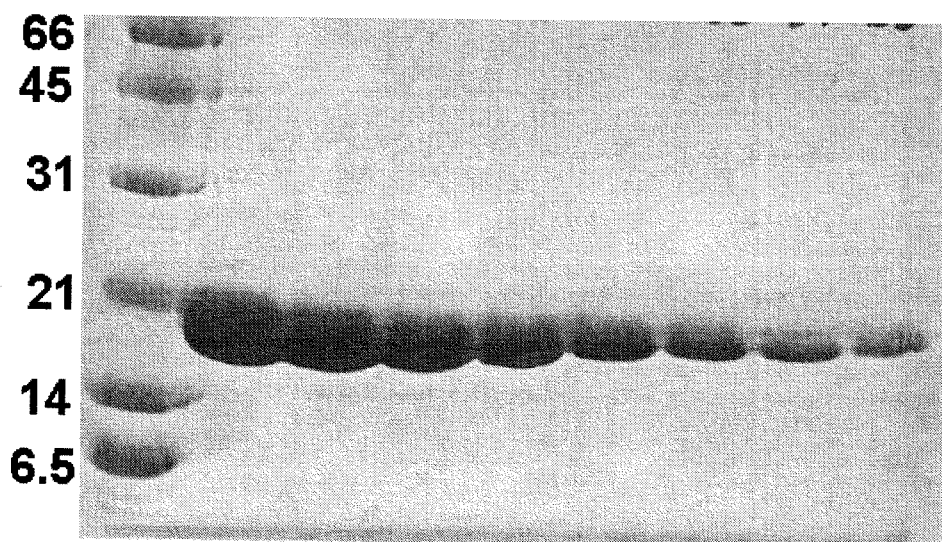


FIG. 2

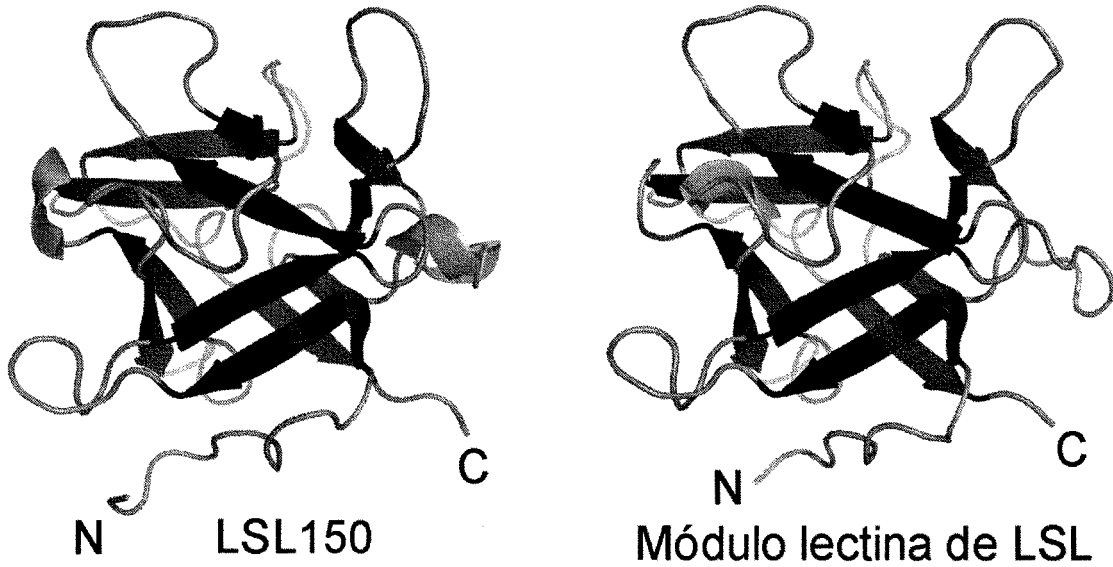


FIG. 3

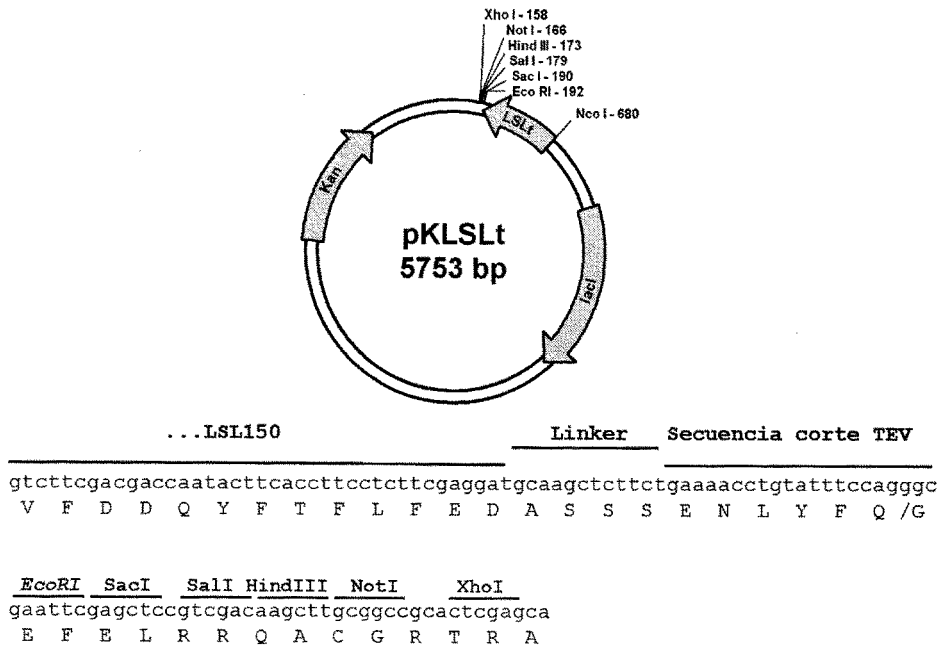
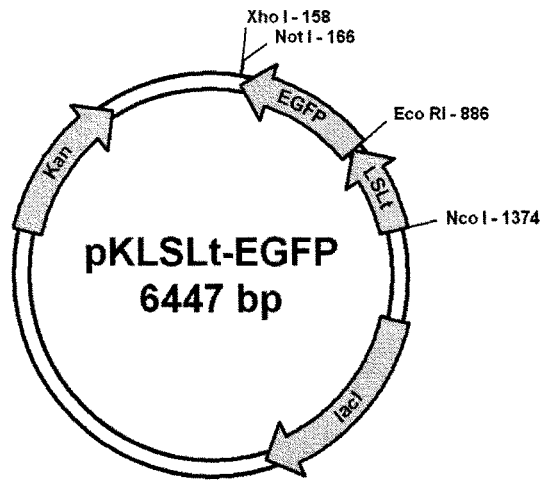


FIG. 4



```

          EcoRI              NotI  XhoI
tgaaaacctgtatttccagggcgaattcatggtgagcaag...aagtaagcgggccgcactcgag
  E  N  L  Y  F  Q / G  E  F  M  V  S  K.....K  -
Secuencia corte TEV             ...EGFP...
    
```

FIG. 5

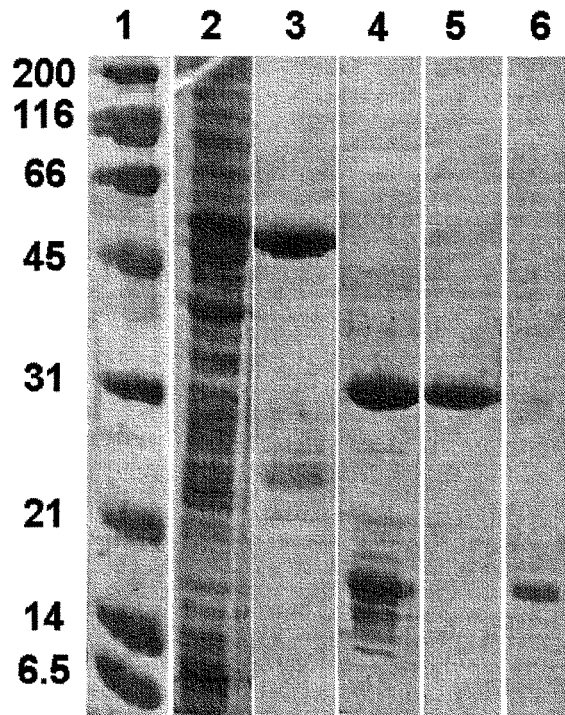
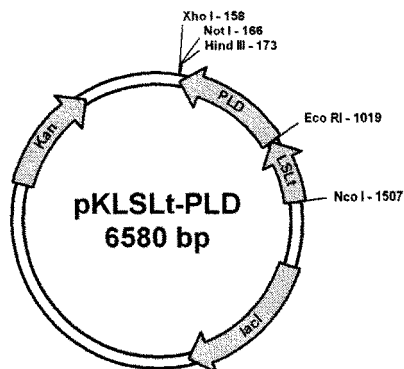


FIG. 6



EcoRI  
 tgaaaacctgtattccaggcgaattgctgcgccactt...attccgtggtgaaagcttcgggccgcactcgagca  
E N L Y F Q / G E F A A P L ... I P W -  
 Secuencia corte TEV ...PLD...

FIG. 7

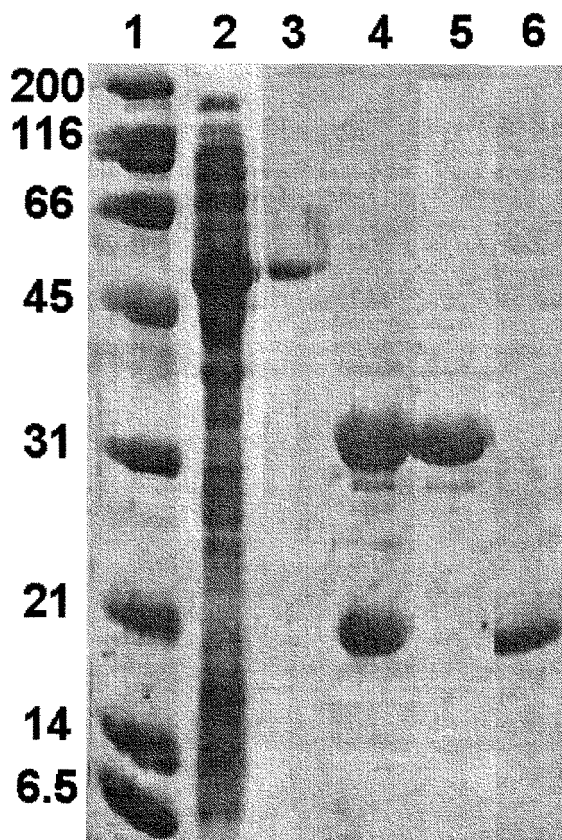
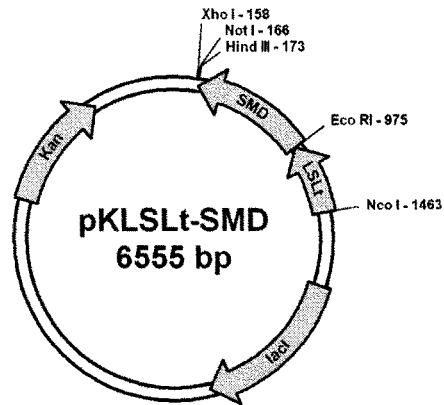


FIG. 8

ES 2 327 309 B1



EcoRI tgaaaacotgtatttccagggcgaattcgcttctacagc...aacccgtggtgaaagcttgccggccgcactcgagca  
E N L Y F Q / G E F A S T A . . . N P W -  
Secuencia corte TEV ...SMD...  
HindIII NotI XhoI

FIG. 9

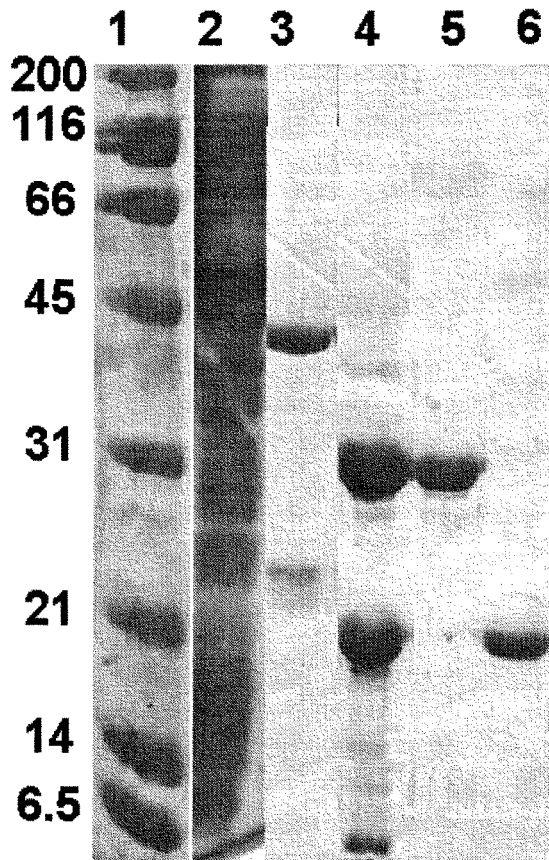
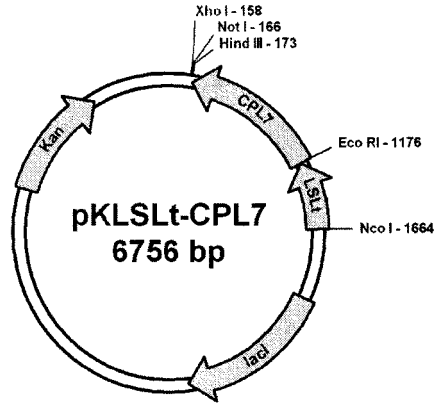


FIG. 10



```

          EcoRI                                HindIII NotI      XhoI
tgaaaaacctgtatttccagggcggaattatgggtaagaaa...ttactttcataaaagcttgccggccgactcgagca
E  N  L  Y  F  Q  /G  E  F  M  V  K  K...L  L  S  -
Secuencia corte TEV                                ...CPL7...
    
```

FIG. 11

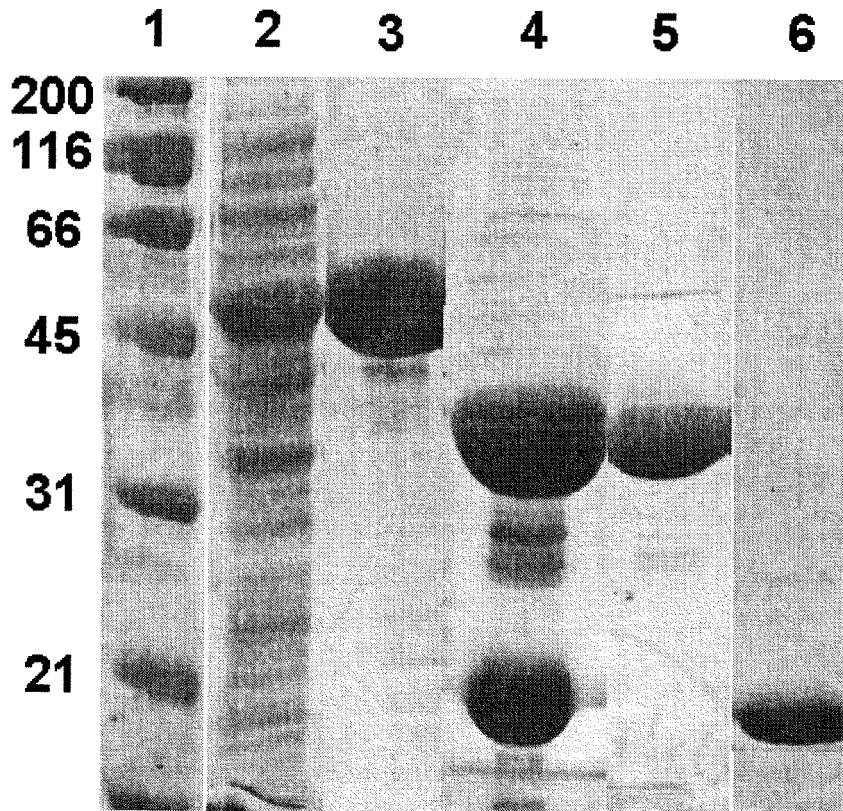
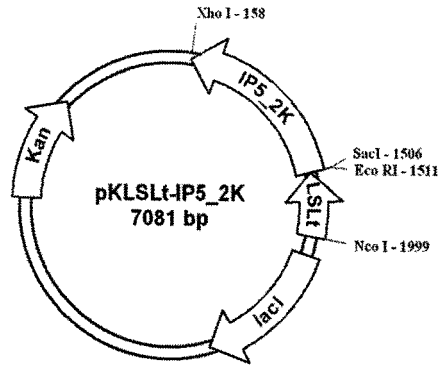


FIG. 12



$\overline{\text{EcoRI}}$   $\overline{\text{XhoI}}$   
 tgaaaacctgtatttccagggcgaattcgagctc**atg**gagatgatt...agctaactcgag  
 E N L Y F Q / G E F E L M E M I ..... S -  
 Secuencia corte TEV IP5\_2K

FIG 13

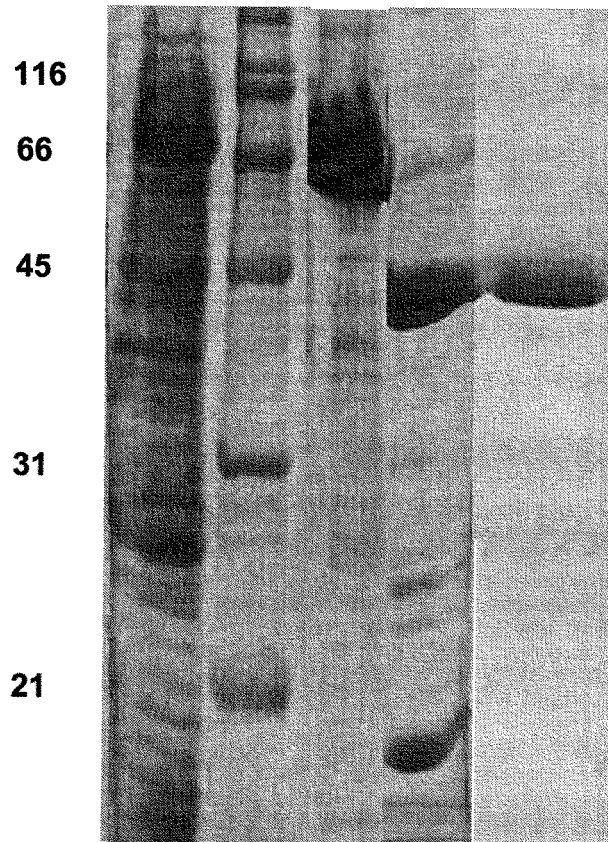


FIG 14



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

① ES 2 327 309

② Nº de solicitud: 200800909

③ Fecha de presentación de la solicitud: 02.04.2008

④ Fecha de prioridad:

## INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤ Int. Cl.: C12N 15/62 (2006.01)  
C07K 14/42 (2006.01)

### DOCUMENTOS RELEVANTES

Categoría	⑥ Documentos citados	Reivindicaciones afectadas
Y	MANCHEÑO, J.M., TATENO, H, GOLDSTEIN, I.J., MARTÍNEZ-RIPOLL, M. y HERMOSO, J.A. Structural Analysis of the <i>Laetiporus sulphureus</i> Hemolytic Pore-forming lectin in Complex with Sugars. J. Biological Chemistry. Abril 2005, Vol. 280, N° 17, páginas 17251-17259. doi:10.1074/jbc.M413933200.	1-8
Y	WO 9966053 A1 (DICTAGENE, S.A.) 23.12.1999, resumen; páginas 1-5.	1-8
A	TIELKER, D., ROSENAU, F., BARTELS, K-M., ROSENBAUM, T. y JAEGER, K-ED. Lectin-based affinity tag for one-step protein purification. Bio-Techniques. Septiembre 2006, Vol. 41, N° 3, páginas 327-332. doi:10.2144/000112236.	1-8
A	TAYLOR, M.E. y DRICKAMER, K. Carbohydrate-recognition domains as tools for rapid purification of recombinant eukaryotic proteins. Biochem. J. Marzo 1991, Vol. 274, páginas 575-580. ISSN 0264-6021 (Impr.).	1-8
A	TATENO, H. y GOLDSTEIN, I. Molecular Cloning, Expression, and Characterization of Novel Hemolytic Lectins from the Mushroom <i>Laetiporus sulphureus</i> , Which Show Homology to Bacterial Toxins. J. Biological Chemistry. Octubre 2003, Vol. 278, N° 42, páginas 40455-40463. doi:10.1074/jbc.M306836200.	1-8

#### 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  
23.06.2009

Examinador  
A. Martín-Falquina Garre

Página  
1/5



Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C12N, C07K

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

EPODOC, WPI, INVENES, EBI, MEDLINE, BIOSIS, EMBASE, NPL, XPESP

Fecha de Realización de la Opinión Escrita: 23.06.2009

**Declaración**

<b>Novedad (Art. 6.1 LP 11/1986)</b>	Reivindicaciones 1-8	<b>SÍ</b>
	Reivindicaciones	<b>NO</b>
<b>Actividad inventiva (Art. 8.1 LP 11/1986)</b>	Reivindicaciones	<b>SÍ</b>
	Reivindicaciones 1-8	<b>NO</b>

Se considera que la solicitud cumple con el requisito de **aplicación industrial**. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

**Base de la Opinión:**

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como ha sido publicada.

**1. Documentos considerados:**

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	Mancheño, J.M., Tateno, H, Goldstein, I.J., Martínez-Ripoll, M. y Hermoso, J.A. Structural Analysis of the <i>Laetiporus sulphureus</i> Hemolytic Pore-forming lectin in Complex with Sugars. J. Biological Chemistry. Vol. 280, Nº 17, páginas 17251-17259. doi:10.1074/jbc.M413933200	29.04.2005
D02	WO9966053 A1 (DICTAGENE, S.A.)	23.12.1999
D03	Tielker, D., Rosenau, F., Bartels, K-M., Rosenbaum, T. y Jaeger, K-ED. Lectin-based affinity tag for one-step protein purification. Bio-Techniques. Vol. 41, Nº 3, páginas 327-332. doi: 10.2144/000112236	09-2006

**2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración**

La invención se refiere a una proteína de fusión basada en la secuencia de aminoácidos del dominio lectina LSL150 del hongo *Laetiporus sulphureus*, su procedimiento de obtención y sus aplicaciones en protocolos de producción/purificación de proteínas recombinantes de interés.

## Reivindicación 1:

El documento D01 divulga la estructura globular beta-trébol del módulo lectina LSL del hongo *Laetiporus sulphureus* (residuos 3-150) y la secuencia de aminoácidos 13 a 148 de dicha lectina (figura 2C), la cual coincide con los aminoácidos correspondientes de la secuencia identificada como SEQ ID NO2 de la reivindicación 1. Sin embargo, esta publicación no divulga que esa secuencia se encuentre formando parte de una proteína de fusión.

El documento D02, divulga una proteína de fusión que lleva como etiqueta de afinidad una lectina del hongo *Dictyostelium discoideum*, lo que permite su utilización para la purificación de proteínas recombinantes de interés por técnicas cromatográficas utilizando galactosa como eluyente.

En vista de lo divulgado tanto en D01 como en D02, la reivindicación 1 cumple el requisito de novedad según el Art. 6.1 LP 11/1986.

Sin embargo, la reivindicación 1 no cumple el requisito de actividad inventiva según el Art. 8.1 LP11/1986, por las siguientes razones:

El documento D02, considerado el estado de la técnica más próximo en relación con la reivindicación 1, divulga, como ya se ha mencionado, una proteína de fusión que lleva una lectina del hongo *Dictyostelium discoideum* como etiqueta de afinidad, lo que permite su utilización para la purificación de proteínas recombinantes de interés.

La diferencia entre el documento D02 y la reivindicación 1 radica en que en esta última la proteína de fusión se basa en una lectina diferente. Por lo tanto, el problema técnico objetivo podría formularse como la provisión de una etiqueta de afinidad alternativa para protocolos de purificación de proteínas recombinantes.

Dado que es conocido que la capacidad de unión específica a carbohidratos de las lectinas las hace aptas como etiquetas de afinidad para la purificación de proteínas recombinantes (ver D02 o D03), se considera que la utilización de la lectina LSL150 del hongo *Laetiporus sulphureus* en lugar de una lectina del hongo *Dictyostelium discoideum*, es simplemente una posibilidad que seleccionaría un experto en la materia a la vista de D01 y D02 para resolver el problema planteado sin el ejercicio de actividad inventiva.

Además, como los ejemplos aportados en la solicitud no permiten distinguir ningún efecto sorprendente que pueda ser atribuido al uso de la LSL150 del hongo *Laetiporus sulphureus* en lugar de cualquier otra lectina, la reivindicación 1 se considera obvia para un experto en la materia.

Hoja adicional

Reivindicaciones dependientes 2-8:

Las reivindicaciones 2-8 en cuanto que son dependientes de la reivindicación 1 cumplen con el requisito de novedad (Art. 6.1 LP 11/1986.).

Sin embargo, no cumplen el requisito de actividad inventiva (Art. 8.1 LP11/1986), puesto que, tanto el procedimiento de obtención de la proteína de fusión (reivindicación 2), como la construcción genética que la codifica (reivindicaciones 3 y 4), el vector de expresión para dicha construcción (reivindicaciones 5 y 6) o la célula huésped que comprende la construcción genética o el vector (reivindicaciones 7 y 8), ya están descritos en D02 o se enmarcan en las técnicas habitualmente utilizadas en el campo de la ingeniería genética y, por tanto, son obvias para un experto en la materia.