

(12) SOLICITUD INTERNACIONAL PUBLICADA EN VIRTUD DEL TRATADO DE COOPERACIÓN EN MATERIA DE PATENTES (PCT)

(19) Organización Mundial de la Propiedad Intelectual  
Oficina internacional



(10) Número de Publicación Internacional  
**WO 2015/185767 A1**

(43) Fecha de publicación internacional  
10 de diciembre de 2015 (10.12.2015) **WIPO | PCT**

(51) Clasificación Internacional de Patentes:

C07K 14/415 (2006.01) C12P 21/02 (2006.01)  
C07K 16/16 (2006.01) G01N 33/68 (2006.01)

(21) Número de la solicitud internacional:

PCT/ES2014/070460

(22) Fecha de presentación internacional:

4 de junio de 2014 (04.06.2014)

(25) Idioma de presentación:

español

(26) Idioma de publicación:

español

(71) Solicitantes: **ABBCN, S.L.** [ES/ES]; Campus UAB, Edificio Eureka, despacho P1M2.1, E-08193 Bellaterra (ES). **IMMUNOSTEP, S.L.** [ES/ES]; Institute of Cancer Research, Avda. Universidad de Coimbra, s/n, E-37008 Salamanca (ES). **FUNDACIÓ INSTITUT DE RECERCA BIOMÈDICA (IRB BARCELONA)** [ES/ES]; Parc Científic de Barcelona, C. Baldiri Reixac, 10-12, E-08028 Barcelona (ES). **INSTITUTO DE BIOLOGÍA MOLECULAR DE BARCELONA - CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS** [ES/ES]; C/ Baldiri i Reixac, 10-12, E-08028 Barcelona (ES).

(72) Inventores: **ALDEA MALO, Martí**; C/ Emili 28, E-08380 Malgrat De Mar (ES). **OROZCO LÓPEZ, Modesto**; C. Tarragona, 8, E-08769 Castellví De Rosanes (ES). **COSTA LEJA, Cristina**; C/ Onze de setembre, 69, 4º 2ª, E-25199 Lleida (ES).

(74) Mandatario: **ZEA CHECA, Bernabé**; Pl. Catalunya, 1, E-08002 Barcelona (ES).

(81) Estados designados (*a menos que se indique otra cosa, para toda clase de protección nacional admisible*): AE,

AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Estados designados (*a menos que se indique otra cosa, para toda clase de protección regional admisible*):

ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), euroasiática (AM, AZ, BY, KG, KZ, RU, TJ, TM), europea (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declaraciones según la Regla 4.17:

— sobre el derecho del solicitante para solicitar y que le sea concedida una patente (Regla 4.17(ii))

— sobre el derecho del solicitante a reivindicar la prioridad de la solicitud anterior (Regla 4.17(iii))

Publicada:

— con informe de búsqueda internacional (Art. 21(3))

— con la parte de lista de secuencias de la descripción (Regla 5.2(a))

(54) Title: PEPTIDE TAGS FOR MARKING PROTEINS BY FUSION, AND ANTIBODIES FOR THE DETECTION THEREOF

(54) Título : ETIQUETAS PEPTÍDICAS PARA EL MARCAJE DE PROTEÍNAS POR FUSIÓN, Y ANTICUERPOS PARA SU DETECCIÓN

(57) Abstract: The invention relates to peptide tags to be fused to proteins on the N-terminal or C-terminal end thereof, which are defined by a series of structural and functional properties that define the harmlessness thereof in terms of the modification of the biological function or structure of the protein or peptide to which they fuse, and where the marker protein is selected from the group consisting of a fragment of Phl P 2 of *Phleum pratense*, a fragment of Hev b 6.02 of *Hevea brasiliensis*, and a fragment of Amb t 5 of *Ambrosia trifida*, or a combination thereof. The invention also relates to methods for producing and detecting said recombinant fusion proteins, and to specific antibodies which bind to said marker tags.

(57) Resumen: Se proporcionan etiquetas peptídicas (tags), para su fusión a proteínas en su extremo N-terminal o C-terminal, las cuales están definidas por una serie de propiedades estructurales y funcionales que definen su inocuidad en cuanto a la modificación de la estructura o la función biológica de la proteína o péptido al cual se fusionan, y en donde la proteína de marcaje se selecciona del grupo consistente en un fragmento de Phl P 2 de *Phleum pratense*, un fragmento de Hev b 6.02 de *Hevea brasiliensis* y un fragmento de Amb t 5 de *Ambrosia trifida*, o una combinación de las mismas. También se proporcionan métodos para la producción y detección de estas proteínas recombinantes de fusión, así como anticuerpos específicos que se unen a estas etiquetas de marcaje.



WO 2015/185767 A1

Etiquetas peptídicas para el marcaje de proteínas por fusión, y anticuerpos para su detección

CAMPO DE LA TÉCNICA

5

La presente invención está relacionada con etiquetas peptídicas (*tags*) y su fusión a proteínas recombinantes para su análisis, detección, separación y purificación. También está relacionada con anticuerpos monoclonales que se unen específicamente a estas etiquetas peptídicas de marcaje. La invención

10 tiene aplicaciones potenciales en diversos campos como el de la ingeniería de proteínas, la biología celular, y la proteómica.

ESTADO DE LA TÉCNICA ANTERIOR

15

La ingeniería genética permite actualmente la manipulación y la expresión de prácticamente cualquier secuencia de DNA (y la producción de la proteína para la que codifica) en un número cada vez mayor de células huésped. La tecnología del DNA recombinante permite fusionar dos o más proteínas de tal manera que las dos son traducidas en una única secuencia polipeptídica.

20

Esta es la idea fundamental en la que se basa el etiquetado de proteínas. Una proteína de interés puede ser marcada (fusionada) con una segunda proteína antigénica de marcaje para la que existen anticuerpos específicos (u otro tipo de ligando de gran afinidad). La proteína de interés puede así ser fácilmente localizada, separada o purificada gracias a la interacción de la

25 proteína antigénica fusionada y su anticuerpo o ligando.

30

El etiquetado de proteínas se utiliza en la separación y purificación de proteínas de interés terapéutico, y también en investigación básica en el ámbito de la biología celular y molecular. En la actualidad se dispone de un cierto número de etiquetas para el marcaje, entre las cuales se encuentran péptidos de pequeño tamaño como FLAG, cMyc y HA, o proteínas como GFP o GST (Terpe K. "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems" Appl. Microbiol Biotechnol. 2003, vol. 60., pp. 523-533).

35

Las etiquetas peptídicas tienen que cumplir con una de serie requisitos muy estrictos, en especial si el objetivo es utilizarlas en estudios celulares *in situ*.

Idealmente, no deben alterar la estructura ni interferir con la actividad biológica de la proteína de interés, y deben ser estables en una variedad de contextos celulares. Además, tienen que ser accesibles a sus anticuerpos correspondientes para propiciar su fácil localización, no deben tener  
5 tendencia a formar complejos con ellas mismas u otras proteínas, y no deben alterar o interferir con la localización, distribución o velocidad de degradación de la proteína de interés.

Aunque algunas etiquetas son más versátiles que otras, en general no todas  
10 tienen la misma fiabilidad en todo tipo de aplicaciones. Habitualmente su eficiencia se mueve en un rango de aceptable a alta en la separación y purificación de proteínas, pero su aplicación en estudios funcionales *in vivo* sigue ofreciendo a menudo problemas de difícil solución, pues las distintas estrategias de marcaje siempre implican el riesgo de alterar la función de la  
15 proteína nativa por interacciones impredecibles con la etiqueta peptídica. Las etiquetas de gran tamaño tienen mayor probabilidad de causar interferencia estérica. Por otra parte, y debido a su desorden estructural intrínseco, las etiquetas de pequeño tamaño pueden adoptar conformaciones muy diversas, lo cual facilita enormemente su interacción con la proteína diana con  
20 consecuencias negativas impredecibles sobre su función. En primer lugar, interacciones espurias de la etiqueta en fases intermedias de plegamiento pueden alterar la conformación final de la proteína diana y favorecer su acumulación en grandes agregados. En segundo lugar, la proteína de fusión puede ser más sensible a degradación. Finalmente, si tienen lugar  
25 interacciones no deseadas con dominios esenciales de la proteína diana, la etiqueta puede modificar la estructura, función, interacciones, o localización. No deja de ser un tanto sorprendente que ca. 20% de las más de 400 proteínas analizadas en un estudio integrativo reciente muestran patrones de localización distintos al comparar el marcaje con proteínas fluorescentes a  
30 los métodos de inmunofluorescencia convencionales (Stadler, C. et al. "Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells" Nat. Methods 2013, vol. 10, pp. 315-323).

35 Las limitaciones mencionadas para los pares etiqueta-anticuerpo de marcaje suponen con frecuencia un obstáculo para la investigación. Los estudios experimentales que se centran en determinar con precisión la localización,

transporte, compartimentalización e interacción de diferentes proteínas en la célula se ven frenados con frecuencia por la falta de etiquetas inocuas en su función y de anticuerpos que sean de gran afinidad de tal manera que permitan su utilización en experimentos de inmunofluorescencia *in situ*.

5 Además, al recomendarse un único par etiqueta-anticuerpo para determinadas aproximaciones experimentales, los investigadores no pueden probar diferentes etiquetas para comprobar si su presencia, ausencia o intercambio repercute en la aparición o desaparición de funciones de la proteína de interés, o en la ganancia o pérdida de determinadas funciones  
10 biológicas. Así, hay una creciente necesidad en este campo de la técnica por descubrir etiquetas nuevas con mayor inocuidad funcional y aplicabilidad en diferentes aproximaciones experimentales, y cuyos anticuerpos exhiban tanto una mayor sensibilidad como una menor reactividad cruzada.

## 15 EXPLICACIÓN DE LA INVENCIÓN

Los inventores han hallado que el marcaje de proteínas de interés con determinadas etiquetas peptídicas que cumplen una serie de requisitos  
20 estructurales y funcionales permite la detección de dicha proteína de interés incluso bajo las condiciones experimentales más exigentes, sin comprometer su función biológica original. La etiqueta peptídica de la invención se selecciona de un grupo que consiste en un fragmento de Phl P 2 de *Phleum pratense* (PDB 1WHP), un fragmento de Hev b 6.02 de *Hevea brasiliensis* (PDB 1WKX), y un fragmento de Amb t 5 de *Ambrosia trifida* (PDB 3BBG), o  
25 una combinación de las mismas.

Sorprendentemente, se ha hallado que cualquiera de estas tres etiquetas antigénicas de marcaje es altamente inocua en cuanto a la alteración de las propiedades intrínsecas de las proteínas que marcan. Estas etiquetas,  
30 cuando se fusionan en la misma cadena polipeptídica con una gran variedad de proteínas, consiguen que no se den en ellas alteraciones en cuanto a su difusión, estabilidad, localización celular, velocidad de degradación, actividad biológica y otros parámetros, al contrario que otras etiquetas ampliamente  
35 utilizadas en el campo de la técnica como son las etiquetas HA o FLAG, las cuales sí alteran a menudo las funciones de las proteínas que marcan.

Por consiguiente, las etiquetas peptídicas de la invención permiten realizar

- estudios celulares sobre expresión, localización y/o tráfico intracelular, interacción proteína-proteína o degradación de una gran variedad de proteínas con una probabilidad mucho menor de alterar su comportamiento nativo, es decir, su comportamiento cuando no están marcadas. Este punto
- 5 permitirá un gran avance en la investigación de numerosas proteínas ya que por primera vez se pueden estudiar a nivel celular (y/o purificar en condiciones nativas) sin que su comportamiento se vea modificado por la presencia de la etiqueta.
- 10 Así, un primer aspecto de la invención es una proteína recombinante de fusión que comprende una proteína o péptido que está fusionada en su extremo N-terminal o C-terminal a una proteína de marcaje, donde la proteína de marcaje está definida por las siguientes propiedades:
- 1.1) comprende un número de residuos aminoacídicos inferior a 100;
  - 15 1.2) es inmunogénica;
  - 1.3) tiene un índice de globularidad superior a 0.5, definido el índice de globularidad como la fracción de residuos con accesibilidad al solvente y calculado mediante el programa Visual Molecular Dynamics (VMD);
  - 1.4) está libre de actividad enzimática;
  - 20 1.5) está libre de sitios de unión para cofactores o efectores;
  - 1.6) está libre de secuencias de degradación;
  - 1.7) está libre de secuencias de localización;
  - 1.8) es de origen vegetal;
  - 1.9) está libre de secuencias de homopolimerización o heteropolimerización;
  - 25 1.10) cuando está fusionada en el extremo N-terminal a la Proteína Verde Fluorescente y se expresa la proteína de fusión en células HEK293T, la proteína recombinante de fusión resultante mantiene los niveles de expresión, dichos niveles medidos mediante inmunodetección; mantiene la estabilidad estructural definida según niveles por inmunodetección durante
  - 30 12 horas de tratamiento con 100 µg/ml de cicloheximida; mantiene la difusión intracelular *in vivo* medida mediante pérdida de fluorescencia durante fotoblanqueado, la fluorescencia medida mediante microscopía confocal de fluorescencia; y mantiene la solubilidad de la GFP en extractos celulares medida mediante centrifugación de extractos a 15000 g durante 15
  - 35 minutos, *in vitro* e *in vivo*;
  - 1.11) cuando está fusionada en el extremo C-terminal de la proteína Cdc28 o de la proteína Ydj1 de *Saccharomyces cerevisiae*, la proteína recombinante

de fusión mantiene la morfología definida por microscopía de campo claro; y mantiene la viabilidad celular medida mediante plaqueo a 42°C;

5 1.12) cuando está fusionada a la Proteína Verde Fluorescente y se expresa en *Escherichia coli*, mantiene los niveles de expresión, dichos niveles medidos mediante inmunodetección; mantiene la solubilidad de la GFP en extractos celulares medida mediante centrifugación de extractos a 15000 g durante 15 minutos; y mantiene la eficiencia en la purificación por cromatografía de afinidad medida mediante inmunodetección;

10 1.13) mantiene los niveles de expresión, esta expresión medida mediante inmunofluorescencia y mantiene la localización celular determinada mediante microscopía confocal de fluorescencia de proteínas que se encuentran en diferentes compartimentos celulares, dichas proteínas seleccionadas del grupo que consiste en: beta-actina del citoesqueleto, la proteína STX6 del aparato de Golgi, HO1 del retículo endoplasmático, la HDAC2 nuclear y  
15 GRB2 de la membrana plasmática en células NIH3T3 de ratón, así como FMRP de la sinapsis neuronal en neuronas del hipocampo; y

donde la proteína de marcaje se selecciona del grupo consistente en un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica  
20 de como mínimo 70% de identidad con la SEQ ID NO:1, un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:2 y un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:3, o una combinación de las mismas.

25

Un segundo aspecto de la invención es una secuencia nucleotídica que codifica por la proteína recombinante de fusión según el primer aspecto de la invención.

30 Un tercer aspecto de la invención es un *cassette* de expresión por recombinación que comprende la secuencia nucleotídica según el segundo aspecto de la invención, operativamente unida a una secuencia de control de la expresión.

35 Un cuarto aspecto de la invención es un vector de expresión que comprende una secuencia nucleotídica según el segundo aspecto de la invención o un *cassette* de expresión según el tercer aspecto de la invención.

Un quinto aspecto de la invención es una célula huésped transfectada o transducida con la secuencia nucleotídica según el segundo aspecto de la invención.

5

Un sexto aspecto de la invención es un método para obtener una proteína recombinante de fusión según el primer aspecto de la invención, que comprende:

- 1) Fusionar al menos una secuencia nucleotídica que codifica para la proteína o péptido con una secuencia nucleotídica que codifica para la etiqueta de marcaje;
- 2) Clonar la secuencia resultante del paso 1 en una célula huésped;
- 3) Expresar la secuencia clonada en el paso 2 en una célula.

10

Un séptimo aspecto de la invención es un método para purificar o aislar una proteína recombinante de fusión según el primer aspecto de la invención, que comprende llevar a cabo los pasos según el sexto aspecto de la invención y adicionalmente un paso en donde la proteína de fusión se purifica o aísla.

15  
20

Un octavo aspecto de la invención es un método para detectar una proteína recombinante de fusión según el primer aspecto de la invención que comprende los pasos que se llevan a cabo en el sexto aspecto de la invención y adicionalmente un paso en donde se añaden medios para la detección de la proteína de marcaje.

25

Asimismo, se ha comprobado que cuando las proteínas recombinantes de fusión que forman parte de la invención se detectan mediante el uso de los anticuerpos específicamente desarrollados contra ellas, permiten realizar estudios celulares *in situ* incluso de las proteínas que tienen niveles de expresión más bajos y que por ende requieren una sensibilidad extremadamente alta para su detección.

30

Un noveno aspecto de la invención es un anticuerpo monoclonal que se une específicamente a un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:1, y que se produce mediante el hibridoma depositado el 27.03.2014 en

35

la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con el número de depósito DSM ACC3234. El híbrido de la invención, fue depositado de acuerdo con el tratado de Budapest el 27.03.2014, en la institución "Leibniz-Institut DSMZ -  
5 Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH", en Inhoffenstrasse 7 B, D-38124 Braunschweig, por los depositantes AbBCN S.L. (Campus UAB 08193 Bellaterra) y Immunostep S.L. (Campus Unamuno, 37007, Salamanca). El híbrido se identificó por el depositante mediante la referencia ET5-1, y recibió el número de acceso DSM ACC3234 y fue  
10 declarado viable.

Un décimo aspecto de la invención es un anticuerpo monoclonal que se une específicamente a un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID  
15 NO:2, y que se produce mediante el híbrido depositado el 14.05.2014 en la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con el número de depósito DSM ACC3242. El híbrido de la invención, fue depositado de acuerdo con el tratado de Budapest el 14.05.2014, en la institución "Leibniz-Institut DSMZ -  
20 Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH", en Inhoffenstrasse 7 B, D-38124 Braunschweig, por los depositantes AbBCN S.L. (Campus UAB 08193 Bellaterra) y Immunostep S.L. (Campus Unamuno, 37007, Salamanca). El híbrido se identificó por el depositante mediante la referencia ET6-1, y recibió el número de acceso DSM ACC3242 y fue  
25 declarado viable.

Un undécimo aspecto de la invención es un anticuerpo monoclonal que se une específicamente a un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID  
30 NO:3, y que se produce mediante el híbrido depositado el 27.03.2014 en la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con el número de depósito DSM ACC3236. El híbrido de la invención, fue depositado de acuerdo con el tratado de Budapest el 27.03.2014, en la institución "Leibniz-Institut DSMZ -  
35 Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH", en



Inhoffenstrasse 7 B, D-38124 Braunschweig, por los depositantes AbBCN S.L. (Campus UAB 08193 Bellaterra) y Immunostep S.L. (Campus Unamuno, 37007, Salamanca). El hibridoma se identificó por el depositante mediante la referencia ET10-1, y recibió el número de acceso DSM ACC3236 y fue  
5 declarado viable.

La combinación de las proteínas de marcaje de la invención junto con los anticuerpos monoclonales desarrollados específicamente, representan una poderosa herramienta en aplicaciones de inmunodetección, separación,  
10 aislamiento y purificación de proteínas que requieren una alta sensibilidad en la formación del complejo antígeno-anticuerpo.

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

15 FIG.1. Tests de expresión y agregación de las etiquetas (*tags*) seleccionadas en el cribado bioinformático. a) Las etiquetas de marcaje ET1 a ET12 fueron fusionadas a GFP para analizar los efectos sobre su expresión y agregación en células HEK293T. Los códigos de entrada en la base de datos PDB se indican como referencia. La barra indica 20  $\mu\text{m}$ . b) Los niveles de expresión  
20 de las fusiones etiqueta-GFP se muestran como valores medios (30 células) e intervalos de confianza para la media ( $\alpha=0.05$ ). El eje de las ordenadas indica expresión en unidades relativas a la proteína GFP nativa. c) Como una medida del grado de agregación, se obtuvo el coeficiente de variación (%) de los valores intracelulares de pixel para las fusiones etiqueta-GFP analizadas,  
25 y se indican los valores medios (30 células) y sus intervalos de confianza ( $\alpha=0.05$ ). En el eje de las coordenadas se indica la variabilidad intracelular como coeficiente de variación(%).

FIG.2. Tests de solubilidad y estabilidad para proteínas de marcaje de la  
30 invención. Representación estructural de ET5 (Phi P2, a), ET6 (Hev b 6.02, b) y ET10 (Hev t 5, c). d) Las fusiones de las etiquetas a GFP fueron expresadas en células HEK293Ty los extractos totales fueron analizados mediante inmunodetección con un anticuerpo  $\alpha\text{GFP}$ . e) Los niveles de las fusiones etiqueta-GFP se analizaron 12 horas después de la adición de  
35 cicloheximida mediante inmunodetección y comparados con el control sin tratar. Los valores medios (N=3) y los intervalos de confianza para la media ( $\alpha=0.05$ ) se indican como barras (eje de la izquierda). Se obtuvo por

centrifugación una fracción soluble de las celular HEK293T que expresaban la fusión etiqueta-GFP indicada, y los niveles obtenidos mediante inmunobloting se compararon en relación con los del extracto celular total. Los valores medios (N=3) y los intervalo de confianza para la media ( $\alpha=0.05$ ) se indican como círculos abiertos (eje de la derecha). f) Las células HEK293T que expresaban las fusiones etiqueta-GFP indicadas se analizaron mediante microscopía confocal a lo largo del tiempo bajo condiciones de fotoblanqueado (FLIP). Se muestra la fluorescencia relativa celular a diferentes tiempos (en segundos) durante el fotoblanqueado.

10

FIG.3. Pruebas de inocuidad funcional de las etiquetas de marcaje de la invención y de las del estado de la técnica sobre Cdc28 e Ydj1, dos proteínas esenciales de la levadura de gemación *S. cerevisiae*. a) Imágenes de campo claro de células de la levadura que expresan niveles endógenos de Cdc28 fusionada con las etiquetas ET5 (Phi P 2), ET6 (Hev b 6.02), 3HA y FLAG. Se muestra una cepa control (sin etiqueta) como referencia. La barra es de 10  $\mu\text{m}$ . b) Se muestran en la gráfica los volúmenes individuales críticos de gemación de las células que se encuentran en el apartado b (N>200). Se muestran también los valores medios (líneas verticales gruesas) y los intervalos de confianza ( $\alpha=0.05$ , líneas verticales finas). c) Diluciones en serie de células de levadura que expresaban niveles endógenos de Ydj1 fusionada a las etiquetas de marcaje ET5 (Phi P 2), ET6 (Hev b 6.02), 3HA y FLAG se hicieron crecer en placas, y se incubaron a las temperaturas indicadas. Se muestra una cepa de control (sin etiqueta de marcaje) como referencia. d) Se muestran en la gráfica los volúmenes individuales críticos de gemación de las células que se encuentran en el apartado e (N>250). Se muestran también los valores medios (líneas verticales gruesas) y los intervalos de confianza ( $\alpha=0.05$ , líneas verticales finas).

30

FIG. 4. Inocuidad de las etiquetas de marcaje en la localización de proteínas que se encuentran en diferentes compartimentos celulares: citoesqueleto ( $\beta$ actina), Golgi (STX6), retículo endoplasmático (HO1), núcleo (HDAC2), membrana plasmática (GRB2), y sinapsis neuronal (FMRP). a) Las etiquetas ET5 (Phi P 2) y ET6 (Hev b 6.02) fusionadas a las proteínas indicadas fueron expresadas en células NIH3T3 y analizadas mediante inmunofluorescencia (señal más clara en la figura) con los correspondientes anticuerpos  $\alpha$ ET5 y  $\alpha$ ET6. La barra es de 10  $\mu\text{m}$ . b) La fusión ET5-FMRP fue expresada en las

35

neuronas de hipocampo de ratón y analizada mediante inmunofluorescencia (señal más clara en la figura) con anticuerpos  $\alpha$ ET5. Una imagen de GFP co-transfectada y tratada para mejorar los límites somáticos y neuríticos se muestra también para referencia. Barra de 25  $\mu$ m. Se muestra también una  
5 amplificación 5x de la región indicada en el panel central. Barra de 5  $\mu$ m. En el panel derecho se muestra una proyección 3D de esta región.

FIG. 5. Estabilidad e integridad de las etiquetas de la invención y otras etiquetas peptídicas comúnmente utilizadas cuando están fusionadas a  
10 proteínas esenciales de la levadura de gemación *S. cerevisiae*. Extractos totales de células de levadura que expresan niveles endógenos de Cdc28 (a) o de Ydj1 (b) fusionadas a ET5 (Phl P 2), ET6 (Hev b 6.02), 3HA o 6FLAG se analizaron mediante inmunodetección con anticuerpos  $\alpha$ Cdc28 o  $\alpha$ Ydj1, respectivamente. Se muestra también una cepa de control como referencia.

15

FIG. 6. Análisis de integridad de las etiquetas de la invención después de su purificación por afinidad a partir de células de *E. coli*. a) Las etiquetas fusionadas a 6His se purificaron mediante cromatografía de afinidad a metales divalentes, y se analizaron mediante SDS-PAGE. Debido a su  
20 pequeño tamaño, las etiquetas ET6 y ET10 fueron expresadas y purificadas como dímeros. Una proteína contaminante común está marcada con un asterisco. b) Las correspondientes fusiones a GFP-6His fueron expresadas, purificadas y analizadas como en el apartado a. La fusión GFP-6His sin etiquetar se muestra como referencia.

25

FIG. 7. Análisis de sensibilidad de anticuerpos que unen a las etiquetas de la invención ET5 (Phl P 2) y ET6 (Hev b 6.02). a) Las etiquetas ET5 y ET6 fusionadas a GFP fueron expresadas en células HEK293T y analizadas mediante inmunofluorescencia (señal más clara en la imagen del panel inferior) con los anticuerpos  $\alpha$ ET5 y  $\alpha$ ET6 correspondientes. Se muestra también la fluorescencia de la GFP (señal más clara en la imagen del panel superior). La barra es de 20  $\mu$ m. b) Cuantificación de señales de inmunofluorescencia producidas por los anticuerpos  $\alpha$ ET5  $\alpha$ ET6 seleccionados relativa a la que se obtiene con un anticuerpo policlonal  
35  $\alpha$ GFP.

Los valores medios (>30 células) y los intervalos de confianza para la media ( $\alpha=0.05$ ) se muestran como barras (eje de la izquierda). Los porcentajes de

células GFP-positivas que fueron también positivas por inmunofluorescencia se muestran también como círculos abiertos (eje de la derecha).

FIG. 8. Señal no específica de inmunofluorescencia de los anticuerpos monoclonales de la invención, contra las etiquetas ET5 (Phl P 2 ) y ET6 (Hev b 6.02). a) Se llevó a cabo un análisis de inmunofluorescencia de las células HEK293T no transfectadas con el anticuerpo  $\alpha$ FLAG a una concentración 50 veces superior a la usual para evaluar su unión no específica a los diferentes componentes celulares (panel superior izquierdo). Las células incubadas únicamente con el anticuerpo secundario fueron utilizadas como control (panel superior derecho). Se muestran también las imágenes de campo brillante de los campos correspondientes (paneles inferiores). La barra es de 50  $\mu$ m. b) Cuantificación de la señal no específica de inmunofluorescencia producida por los anticuerpos  $\alpha$ ET5,  $\alpha$ ET6,  $\alpha$ FLAG,  $\alpha$ HA, y  $\alpha$ GFP a una concentración 50 veces superior a la usual. Los valores se relativizaron al valor obtenido con el anticuerpo secundario. También se indican los valores medios (>30 células) e intervalos de confianza para la media ( $\alpha=0.05$ ).

## 20 DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

Un aspecto de la presente invención es una proteína recombinante de fusión que comprende una proteína o péptido que está fusionada en su extremo N-terminal o C-terminal a una proteína de marcaje, donde la proteína de marcaje está definida por las siguientes propiedades:

- 1.1) comprende un número de residuos aminoacídicos inferior a 100;
  - 1.2) es inmunogénica; 1.3) tiene un índice de globularidad superior a 0.5, definido el índice de globularidad como la fracción de residuos con accesibilidad al solvente y calculado mediante el programa VMD;
  - 1.4) está libre de actividad enzimática; 1.5) está libre de sitios de unión para cofactores o efectores; 1.6) está libre de secuencias de degradación;
  - 1.7) está libre de secuencias de localización; 1.8) es de origen vegetal;
  - 1.9) está libre de secuencias de homopolimerización o heteropolimerización;
- ; y donde la proteína de marcaje se selecciona del grupo consistente en un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:1, un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como

mínimo 70% de identidad con la SEQ ID NO:2 y un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:3 , o una combinación de las mismas.

- 5 Como se ha descrito anteriormente,  
un primer aspecto de la invención es una proteína recombinante de fusión  
que comprende una proteína o péptido que está fusionada en su extremo N-  
terminal o C-terminal a una proteína de marcaje, donde la proteína de  
marcaje está definida por las siguientes propiedades:
- 10 1.1) comprende un número de residuos aminoacídicos inferior a 100; 1.2) es  
inmunogénica; 1.3) tiene un índice de globularidad superior a 0.5, definido el  
índice de globularidad como la fracción de residuos con accesibilidad al  
solvente y calculado mediante el programa VMD; 1.4) está libre de actividad  
enzimática; 1.5) está libre de sitios de unión para cofactores o efectores;
- 15 1.6) está libre de secuencias de degradación; 1.7) está libre de secuencias  
de localización; 1.8) es de origen vegetal; 1.9) está libre de secuencias de  
homopolimerización o heteropolimerización; 1.10) cuando está fusionada en  
el extremo N-terminal a la Proteína Verde Fluorescente y se expresa la  
proteína de fusión en células HEK293T, la proteína recombinante de fusión  
20 resultante mantiene los niveles de expresión, dichos niveles medidos  
mediante inmunodetección; mantiene la estabilidad estructural definida  
según niveles por inmunodetección durante 12 horas de tratamiento con 100  
µg/ml de cicloheximida; mantiene la difusión intracelular *in vivo* medida  
mediante pérdida de fluorescencia durante fotoblanqueado, la fluorescencia  
25 medida mediante microscopía confocal de fluorescencia; y mantiene la  
solubilidad de la GFP en extractos celulares medida mediante centrifugación  
de extractos a 15000 G durante 15 minutos, *in vitro* e *in vivo*; 1.11) cuando  
está fusionada en el extremo C-terminal de la proteína Cdc28 o de la  
proteína Ydj1 de *Saccharomyces cerevisiae* la proteína recombinante de  
30 fusión mantiene la morfología definida por microscopía de campo claro; y  
mantiene la viabilidad celular medida mediante plaqueo a 42°C; 1.12)  
cuando está fusionada a la Proteína Verde Fluorescente y se expresa en  
*Escherichia coli*, mantiene los niveles de expresión, dichos niveles medidos  
mediante inmunodetección; mantiene la solubilidad de la GFP en extractos  
35 celulares medida mediante centrifugación de extractos a 15000 g durante 15  
minutos; y mantiene la eficiencia en la purificación por cromatografía de  
afinidad medida mediante inmunodetección; 1.13) mantiene los niveles de

expresión, esta expresión medida mediante inmunofluorescencia y mantiene la localización celular determinada mediante microscopía confocal de fluorescencia de proteínas que se encuentran en diferentes compartimentos celulares, dichas proteínas seleccionadas del grupo que consiste en: beta-actina del citoesqueleto, la proteína STX6 del aparato de Golgi, HO1 del retículo endoplasmático, la HDAC2 nuclear y GRB2 de la membrana plasmática en células NIH3T3 de ratón, así como FMRP de la sinapsis neuronal en neuronas del hipocampo; y

5 donde la proteína de marcaje se selecciona del grupo consistente en un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:1, un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:2 y un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 70% de

10 identidad con la SEQ ID NO:3 , o una combinación de las mismas.

15

En una realización particular del primer aspecto de la invención, la proteína de marcaje es un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID

20 NO:1

En una realización particular del primer aspecto de la invención, la proteína de marcaje es un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 80% de identidad con la SEQ ID

25 NO:1

En una realización particular del primer aspecto de la invención, la proteína de marcaje es un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 90% de identidad con la SEQ ID

30 NO:1. En otra realización particular de este primer aspecto, la proteína de marcaje es un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica que consiste en la SEQ ID NO:1, o lo que es lo mismo con un 100 % de identidad con la SEQ ID NO: 1.

35 En otra realización particular del primer aspecto de la invención, la proteína de marcaje es un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID

NO:2

5 En otra realización particular del primer aspecto de la invención, la proteína de marcaje es un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 80% de identidad con la SEQ ID NO:2

10 En otra realización particular del primer aspecto de la invención, la proteína de marcaje es un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 90% de identidad con la SEQ ID NO:2. En otra realización particular de este primer aspecto, la proteína de marcaje es un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica que consiste en la SEQ ID NO: 2, o lo que es lo mismo con un 100 % de identidad con la SEQ ID NO: 2.

15

En otra realización particular del primer aspecto, la proteína de marcaje es un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:3.

20 En otra realización particular del primer aspecto, la proteína de marcaje es un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 80% de identidad con la SEQ ID NO:3.

25 En otra realización particular del primer aspecto, la proteína de marcaje es un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 90% de identidad con la SEQ ID NO:3. En otra realización particular de este primer aspecto, la proteína de marcaje es un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica que consiste en la SEQ ID NO: 3, o lo que es lo mismo con un 100 % de identidad con la SEQ ID NO: 3.

30

Todas estas secuencias de las proteínas de marcaje fusionadas con la proteína o péptido de interés están codificadas, de acuerdo con el segundo aspecto de la invención, por una secuencia nucleotídica.

35

Como se ha descrito más arriba, un tercer aspecto de la invención es un cassette de expresión por recombinación que comprende la secuencia

nucleotídica según el segundo aspecto de la invención, operativamente unida a una secuencia de control de la expresión.

5 En una realización particular del tercer aspecto de la invención, el cassette de expresión es tal que la secuencia de control de la expresión comprende una secuencia diana para reguladores transcripcionales, una secuencia de unión al ribosoma y una secuencia terminadora de la transcripción.

10 Como se ha descrito más arriba, el séptimo aspecto de la invención es un método para purificar o aislar una proteína recombinante de fusión según el primer aspecto, que comprende llevar a cabo los pasos según el sexto aspecto de la invención y adicionalmente un paso en donde la proteína de fusión se purifica o aísla.

15 En una realización particular del séptimo aspecto de la invención, el paso adicional de purificación o aislamiento de la proteína de fusión comprende un anticuerpo o fragmento del mismo que se une a la proteína de marcaje.

20 Como también ha sido descrito más arriba, un octavo aspecto de la invención es un método para detectar una proteína recombinante de fusión según el primer aspecto de la invención que comprende llevar a cabo los pasos según el sexto aspecto de la invención y adicionalmente un paso en donde se añaden medios para la detección de la proteína de marcaje.

25 En una realización particular del octavo aspecto de la invención, los medios para la detección de la proteína de marcaje comprenden un anticuerpo o un fragmento del mismo.

30 En otra realización particular del octavo aspecto de la invención, los medios para la detección de la proteína de marcaje comprenden un anticuerpo monoclonal.

35 En una realización particular de los aspectos séptimo y octavo de la invención, los medios se seleccionan de los anticuerpos que se obtienen mediante los hibridomas depositados en la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con los números de depósito DSM ACC3236 (depositado el día 27.03.2014), DSM



ACC3234 (depositado el día 27.03.2014) y DSM ACC3242 (depositado el día 14.05.2014).

5 A lo largo de la descripción y las reivindicaciones la palabra "comprende" y sus variantes no pretenden excluir otras características técnicas, aditivos, componentes o pasos. Además, la palabra "comprende" incluye el caso "consiste en". Para los expertos en la materia, otros objetos, ventajas y características de la invención se desprenderán en parte de la descripción y en parte de la práctica de la invención. Los siguientes ejemplos se  
10 proporcionan a modo de ilustración, y no se pretende que sean limitativos de la presente invención. Además, la presente invención cubre todas las posibles combinaciones de realizaciones particulares y preferidas aquí indicadas.

15 Definiciones.

Todos los términos utilizados en la presente solicitud deben ser entendidos en su significado comúnmente utilizado en el campo de la técnica. Sin embargo, con el objetivo de mejorar la claridad, se incluyen las siguientes  
20 definiciones. Las abreviaturas de las proteínas citadas en la presente solicitud también se listan a continuación, juntamente con su entrada en la base de datos Uniprot.

25 El término "proteína recombinante de fusión" como se utiliza aquí significa una única cadena polipeptídica compuesta de: 1) una proteína de interés (proteína de la cual se quiere estudiar su expresión celular, distribución, compartimentalización, interacción, etc. o proteína que se quiere aislar, purificar o separar), y 2) una proteína (etiqueta) de marcaje.

30 Los términos "proteína de marcaje" o "etiqueta de marcaje" (términos que son intercambiables en la presente memoria y que en alguna ocasión se denominan también "tags" por su nomenclatura en inglés) como se utilizan aquí significan una proteína para la cual existen anticuerpos u otros ligandos de alta afinidad, que se expresa fusionada (en la misma cadena  
35 polipeptídica) a una proteína de interés, de tal manera que la proteína de interés puede ser separada, purificada o localizada en estudios celulares gracias a la interacción de la proteína de marcaje con su ligando de alta

afinidad.

El término "inmunogénico" como se utiliza aquí, significa una sustancia (en el caso de la presente invención, una proteína) que es capaz de inducir una  
5 respuesta inmune.

El término "índice de globularidad" da una idea de la compacidad de la estructura de una proteína, y está relacionada con su estabilidad e integridad estructural. Es sabido que numerosas proteínas (o algunos de sus dominios)  
10 son intrínsecamente desordenadas, es decir, que no adquieren una estructura tridimensional fija. Cuanto más desordenada es una proteína, menor es su globularidad, y se define como la fracción de residuos con accesibilidad al solvente. En una proteína desordenada, el promedio de la accesibilidad al solvente de sus residuos aumenta. Existen numerosos  
15 programas que calculan la accesibilidad al solvente de los residuos de una proteína. En la presente invención se ha utilizado el programa VMD.

El término "secuencia de degradación" como se utiliza aquí significa una secuencia de aminoácidos que, cuando está presente en una cadena  
20 polipeptídica, dirige a ésta última hacia una ruta de degradación o proteólisis parcial. Por ejemplo, una secuencia de degradación comúnmente conocida y denominada como PEST se compone de segmentos ricos en prolina, glutámico, serina y treonina.

El término "secuencia de localización" como se utiliza aquí significa una secuencia de aminoácidos que, cuando está presente en una cadena polipeptídica, dirige a ésta última a una localización concreta dentro de los compartimentos celulares (como puede ser el aparato de Golgi, el retículo endoplasmático, el núcleo de la célula, etc.). Por ejemplo, una secuencia  
30 típica de localización nuclear es la secuencia KRPAATKKAGQAKKKK (SEQ ID NO:4) de la nucleoplasmina, que se compone de dos grupos de residuos cargados positivamente, y que es reconocida por la  $\alpha$ -importina; una secuencia típica del extremo C-terminal que provoca retención de la proteína en el retículo endoplasmático es KDEL (SEQ ID NO:5).

35

El término "secuencia de homopolimerización o heteropolimerización" como se utiliza aquí significa una secuencia que, cuando está presente en una

cadena polipeptídica, propicia que ésta forme complejos con otras cadenas polipeptídicas idénticas a ella (homo) o diferentes en cuanto a secuencia o estructura (hetero).

- 5 El término "efector" como se utiliza aquí es una proteína, una molécula orgánica de pequeño tamaño (como puede ser una hormona, un metabolito, etc.) o un ion, que selectivamente se une a una proteína para regular su función. El efector normalmente tiene un sitio de unión específico en la proteína que regula.

10

El término "libre de" como se utiliza aquí, significa "carente de".

El término "libre de actividad enzimática" significa aquí que la proteína no cataliza ninguna reacción bioquímica.

15

El término "estabilidad estructural" significa aquí que la proteína no pierde la conformación tridimensional responsable de su actividad biológica, no es degradada, metabolizada, o eliminada mediante diferentes sistemas celulares de procesamiento de proteínas.

20

El término "transducción" como se utiliza aquí, significa un proceso mediante el cual un DNA foráneo (heterólogo) es introducido en una célula huésped mediante el uso de un vector vírico. El objetivo último de este proceso es que el DNA foráneo se integre en el genoma de la célula huésped de tal manera que el RNA y/o la proteína por la cual codifica se expresen de manera estable en esta célula. En el caso de la presente invención, la transducción se realiza con el DNA que codifica para las proteínas recombinantes de fusión.

25

- 30 El término "transfección" como se utiliza aquí, significa un proceso mediante el cual un DNA foráneo (heterólogo) es introducido en una célula huésped mediante cualquier método que no sea viral, tal como electroporación o mediante tratamiento químico. En el caso de la presente invención, la transfección se realiza con el DNA que codifica para las proteínas recombinantes de fusión.
- 35

Tal como se utiliza en esta descripción, la expresión "operativamente unida"

significa que el producto o productos de interés es (son) expresado(s) en el marco de lectura correcto bajo el control de las secuencias de control o reguladoras de expresión.

- 5 En la presente invención debe entenderse por “porcentaje de identidad” el porcentaje de aminoácidos o de nucleótidos iguales entre dos o más secuencias independientes que se comparan o alinean de tal modo que se obtenga el mejor valor de coincidencias entre posiciones. El porcentaje de identidad entre secuencias peptídicas y entre secuencias nucleotídicas se  
10 calcula, preferentemente, mediante el algoritmo BLAST por “*Basic Local Alignment Search Tool*”, accesible desde el portal de la NCBI (<http://www.ncbi.nlm.nih.gov/BLAST>) y descrito en Altschul S. "Basic local alignment search tool", J. Mol. Biol. 1990, vol. 215, pp. 403–410. Así, cuando en la presente invención se indica que dos secuencias tienen un mínimo de  
15 70% de identidad, significa que una vez alineadas, las dos secuencias tienen un mínimo de 7 de cada 10 aminoácidos idénticos.

En el sentido de la presente invención, se entiende por “anticuerpo o fragmento del mismo”, cualquier anticuerpo policlonal o monoclonal, o un  
20 fragmento derivado del mismo que aún retiene la capacidad para unirse a su antígeno. Ejemplos de fragmentos son: un fragmento (Fab)<sub>2</sub> derivado de la digestión de un anticuerpo con pepsina, un fragmento Fab derivado de la digestión de un anticuerpo con papaína, un fragmento scFv derivado de fusionar los dos dominios variables de las cadenas pesada y ligera,  
25 nanocuerpos (*nanobodies*) derivados de anticuerpos de camélido o de escualos, etc.

La Proteína Verde Fluorescente (Green Fluorescent Protein, o “GFP”), es una proteína de 238 residuos aislada de la medusa *Aequorea victoria*, que  
30 tiene una marcada fluorescencia de color verde cuando se expone a luz ultravioleta. Esta proteína se utiliza habitualmente en estudios de biología molecular y celular como marcador. Al ser tan fácilmente detectable se utiliza para controlar el nivel de expresión génica, para el control del desarrollo celular, la división en procariontes, etc. Se encuentra en la base de datos  
35 Uniprot con la entrada (código) P42212, versión 118 de la entrada de 14 de Mato de 2014 y versión 1 de la secuencia de 1 de Noviembre de 1995.

Cdc28 de *Saccharomyces cerevisiae* (o “cell division control protein 28” también llamada cyclin dependent kinase 1), es una proteína esencial para el control del ciclo celular de la levadura. Se encuentra en la base de datos Uniprot con la entrada (código) P00546, versión 160 de la entrada de 14 de  
5 Mato de 2014 y versión 1 de la secuencia de 21 de Julio de 1986.

Ydj1 de *Saccharomyces cerevisiae* (o “yeast DNAJ protein 1, también llamada “mitochondrial protein import protein MA S5”) es una proteína de 409 aminoácidos involucrada en la importación de proteínas a la mitocondria de  
10 la levadura. Se encuentra en la base de datos Uniprot con la entrada (código) P25491, versión 146 de la entrada de 14 de Mayo de 2014 y versión 1 de la secuencia de 1 de Mayo de 1992.

Beta-actina de *Mus musculus* es una proteína de 375 residuos que está  
15 involucrada en diversos procesos de motilidad celular y que forma parte integral del citoesqueleto. Se encuentra en la base de datos Uniprot con la entrada (código) P60710, versión 110 de la entrada de 14 de Mayo de 2014 y versión 1 de la secuencia de 1 de Abril de 1988.

20 STX6 de *Mus musculus* (también llamada Syntaxin-6) es una proteína de 255 residuos involucrada en el tráfico vesicular intracelular. Se encuentra en la base de datos Uniprot con la entrada (código) Q9JJK1, versión 101 de la entrada de 14 de Mayo de 2014 y versión 1 de la secuencia 1 de Octubre de 2000.

25 HO1 de *Mus musculus* (Heme oxigenase 1, también llamada proteína P32) es una proteína de 289 residuos que se localiza en el retículo endoplasmático, y que está involucrada en la rotura del anillo del grupo hemo. Se encuentra en la base de datos Uniprot con la entrada (código)  
30 P14901, versión 119 de la entrada de 14 de Mayo de 2014 y versión 1 de la secuencia de 1 de Abril de 1990.

HDAC2 de *Mus musculus* (también llamada HD2, Histone deacetylase 2) es una proteína de 488 residuos que se localiza en el núcleo celular, y que está  
35 involucrada en la deacetilación de residuos de lisina en el extremo N-terminal de las histonas (H2A, H2B, H3 y H4), y por consiguiente está involucrada en el empaquetamiento del DNA y en la regulación epigenética de la expresión

de numerosos genes. Se encuentra en la base de datos Uniprot con la entrada (código) P70288, versión 142 de la entrada de 14 de Mayo de 2014 y versión 1 de la secuencia de 1 de Febrero de 1997.

- 5 GRB2 de *Mus musculus* (*growth factor receptor-bound protein2*, también llamado “*Adapter protein GRB2*”) es una proteína de 217 residuos que se puede localizar en el citoplasma, en endosomas o en el aparato de Golgi, y que desarrolla un papel en la conexión entre los factores de crecimiento de superficie y la vía de señalización de Ras. Se encuentra en la base de datos
- 10 Uniprot con la entrada (código) Q60631, versión 155 de la entrada de 14 de Mayo de 2014 y versión 1 de la secuencia de 1 de Noviembre de 1996.

- FMRP de *Mus musculus* (*Fragile X mental retardation protein*) es una proteína de 589 residuos que se encuentra en neuronas del hipocampo. Se
- 15 encuentra en la base de datos Uniprot con la entrada (código) Q547R0, versión 51 de la entrada de 16 de Abril de 2014 y versión 1 de la secuencia de 24 de Mayo de 2005.

- El término “NIH3T3” hace referencia a la línea celular también denominada
- 20 3T3. Es una línea celular estándar de fibroblastos ampliamente utilizada en investigación celular. Esta línea fue establecida en 1962 por dos científicos en la Universidad de Nueva York, George Todaro y Howard Green.

- “Phl P 2” es el alérgeno del polen Phl P 2 derivado de *Phleum pratense*, una
- 25 hierba perenne que se encuentra en la mayor parte Europa. Esta proteína se encuentra en la base de datos Uniprot con la entrada (código) P43214 (última modificación 16 de Octubre del 2013, versión 80 de la entrada y versión 1 de la secuencia de 1 de Noviembre de 1995). Su estructura tridimensional ha sido resuelta por cristalografía de rayos X y se encuentra
- 30 accesible en el Protein Data Bank (PDB) con el código 1WHP. La secuencia aminoacídica del fragmento del alérgeno Phl P 2 utilizado en la presente invención es (SEQ ID NO 1):

- VPKV TFTVEKGSNE KHLAVLVKYE GDTMAEVELR EHGSDDEWVAM  
TKGEGGVWTF DSEEPLQGPF NFRFLTEKGM KNVFDDVPE
- 35 KYTIGATYAP EE

“Hev b 6.02” es un alérgeno derivado de *Hevea brasiliensis* (árbol del

caucho), la especie más importante de árboles que producen latex en Sud América. Esta proteína se encuentra en la base de datos Uniprot con la entrada (código) Q6JYQ7 (última modificación 19 de Febrero del 2014, versión 51 de la entrada y versión 1 de la secuencia de 5 de Julio de 2004).  
5 Su estructura tridimensional ha sido resuelta por cristalografía de rayos X y se encuentra accesible en el Protein Data Bank (PDB) con el código 1WKX. La secuencia aminoacídica del fragmento del alérgeno Hev b 6.02 utilizado en la presente invención es (SEQ ID NO 2):

EQCGRQAGGKLCPDNLCCSQWGWCGSTDEYCSPDHNCQSNCKD

10

“Amb t 5” es el alérgeno del polen Amb t 5 derivado de *Ambrosia trifida* (una especie del género Ambrosia, plantas herbáceas muy comunes en América). Esta proteína se encuentra en la base de datos Uniprot con la entrada (código) P10414 (última modificación 19 Febrero del 2014, versión 82 de la  
15 entrada y versión 2 de la secuencia de 1 de Agosto de 1992). Su estructura tridimensional ha sido resuelta por Resonancia Magnética Nuclear y está accesible en el Protein Data Bank con el código 3BBG. La secuencia aminoacídica del fragmento Amb t 5 utilizado en la presente invención es (SEQ ID NO 3):

20

DDGLCYEGTNCGKVGKYCCSPIGKYCVCYDSKAICNK NCT

## EJEMPLOS

### Material y Métodos

25

#### Cribado bioinformático.

El objetivo inicial del screening bioinformático fue obtener una lista de dominios de proteínas aptos para preparar proteínas de fusión. Debido a la  
30 dificultad de predecir antigenicidad basada en aproximaciones puramente teóricas, se empezó con ca. 2000 familias de proteínas reportadas en la base de datos VarDB (Hayes, C. N. et al. “varDB: a pathogen-specific sequence database of protein families involved in antigenic variation” Bioinformatics 2008, vol. 24, pp. 2564–2565) y epítomos reportados en la base IEDB (Vita,  
35 R. et al. “The immune epitope database 2.0” Nucl. Acids Res. 2010, vol. 38, D854-D862). Las estructuras de proteínas que contenían los 63.424 epítomos listados en IEDB se obtuvieron utilizando BLASTP contra el PDB (se

utilizaron los parámetros estándar). Finalmente, la lista se extendió para incluir posibles homólogos usando los correspondientes dominios del PFAM (Punta, M. et.al. "The Pfam protein families database". Nucl. Acids Res. 2012, vol. 40, D290-D301). La lista final incluyó 3.500 proteínas antigénicas que cubrían 1.383 dominios PFAM.

La lista inicial se chequeó para comprobar las propiedades siguientes:

- Antigenicidad utilizando BCIPRED (Saha, S. & Raghava, G. P. S. "BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties" en Nicosia, G., Cutello, V., Bentley, P.J. & Timis, J. (eds.) pp. 197-204, Springer 2004, Heidelberg). El método utiliza una combinación de estimaciones basadas en secuencia que incluyen: hidrofobicidad, flexibilidad, accesibilidad, propensión a formar giros beta, propensión antigénica y polaridad. Los dominios considerados antigénicos por al menos dos de los criterios citados, fueron aceptados.
- La accesibilidad al solvente, la distancia del epítipo a los extremos N y C-terminal, la accesibilidad del epítipo al solvente y el índice de globularidad fueron calculados con VMD (Humphrey, et.al. "VMD - Visual molecular dynamics" J. Molec. Graphics 1996, vol. 14, pp. 33-38.)
- La caracterización funcional fue obtenida de las anotaciones en PFAM (*vide supra*) y Gene Ontology (GO, The Gene Ontology Consortium. "Gene ontology annotations and resources", Nucl. Acids Res. 2013, vol. 41, pp. D530-D535. Los términos GO descartados afectaban a las siguientes categorías: Solubilidad y relación con la membrana, actividad, unión, estado polimérico y promiscuidad en la unión a proteínas.

El cribado dio como resultado una lista de 226 dominios que fue manualmente curada como se describe en el apartado de Resultados, para descartar actividades específicas o propiedades que no estaban anotadas en las bases de datos comentadas más arriba.

Cultivo de células de mamífero, vectores de expresión y análisis celular



Se mantuvo a las células NIH3T3 and HEK293T en un medio DMEM (Dulbecco's modified Eagle's médium) que contenía glutamina suplementado con antibióticos y 10% de FCS. Los DNAs sintéticos que codificaban las etiquetas peptídicas fueron sub-clonados en un plásmido derivado del pEGFP-N1 (Clontech) donde algunas de las secuencias polylinker habían sido sustituidas por el promotor T7, secuencias que unen el ribosoma para la expresión génica en *E.coli* y en células de mamífero, y un péptido 6His para la purificación por afinidad. Las células fueron transfectadas con Lipofectamine 2000 siguiendo instrucciones del fabricante (Invitrogen), y analizadas 24 horas después de la transfección. Los niveles de expresión y el índice de agregación de las fusiones GFP en las células HEK293T fueron analizadas mediante microscopia de epifluorescencia con la ayuda del programa de análisis de imagen ImageJ (Wayne Rasband, NIH). La pérdida de fluorescencia por fotoblanqueado (Fluorescence loss in photobleaching, FLIP) se realizó con un microscopio confocal Zeiss LSM780. Una pequeña región circular del citoplasma ( $3.6 \mu\text{m}^2$ ) fue repetitivamente fotoblanqueada a potencia de láser máxima durante 3 segundos cada vez y, entre los períodos de blanqueo, la célula fue examinada por imagen con luz de baja intensidad para medir la pérdida de fluorescencia. Para el análisis cuantitativo, la intensidad de fondo (*background*) fue restada, y las intensidades fuera del área de fotoblanqueo fueron medidas a lo largo del tiempo y normalizadas a las de una célula transfectada no blanqueada.

#### Cepas de levadura, condiciones de crecimiento y medidas celulares

La cepa parental de levadura CML 128 y los métodos utilizados para sustitución génica por recombinación han sido descritos en Gallego, C., et al. "The Cln3 cyclin is downregulated by translational repression and degradation during the G1 arrest caused by nitrogen deprivation in budding yeast". *EMBO J.* 1997, vol. 16, pp. 7196–7206. Las células de levadura se hicieron crecer bajo condiciones exponenciales durante 7-8 generaciones en medio SC (Sambrook, J. & Russell, D. W. "Molecular cloning: A laboratory manual". 2001 3a edición, pp. 323-325, CSHL Press, Cold Spring Harbor, NY) con 2% de glucosa a 30°C, a no ser que se indique lo contrario. El volumen celular en el momento de la gemación fue determinado a partir de imágenes de campo brillante con la ayuda de un BudJ (Ferrezuelo, F. et al. "The critical size is set at a single-cell level by growth rate to attain

homeostasis and adaptation". Nat. Comm. 2012, vol. 3, pp. 1012), un plugin ImageJ que se puede obtener en [www.ibmb.csic.es/home/maldea](http://www.ibmb.csic.es/home/maldea).

#### Purificación de las etiquetas peptídicas

5

La expresión de la etiqueta ET5 (Phl P 2) monomérica marcada con el hexapéptido 6His, de la etiqueta ET6 (Hev b 6.02) dimérica, o de las fusiones correspondientes a GFP fue inducida en *E. coli* BL21 (DE3) con 0.5 mM IPTG y crecidas a 25°C durante 18h a 220 rpm. Las proteínas expresadas fueron purificadas por afinidad en Ni-NTA (Qiagen) siguiendo instrucciones del fabricante.

10

#### Producción de los anticuerpos monoclonales

15

Ratones hembra BALB/c fueron inmunizados intraperitonealmente con 75 µg de cada una de las etiquetas peptídicas de la invención, emulsificadas en adyuvante (Stimune Adjuvant; Prionics). 30 y 60 días más tarde, los ratones fueron inyectados con 75 µg de cada uno de los antígenos en adyuvante. 4 días antes del procedimiento de fusión, cada ratón recibió una inyección intravenosa del mismo antígeno en PBS 20 mM fosfato. Los esplenocitos fueron fusionados con la línea celular de mieloma Sp2/0-Ag14. Los clones de hibridoma fueron seleccionados mediante ELISA indirecto convencional, y las células provenientes de pocillos positivos fueron clonadas mediante dilución limitante. Los anticuerpos monoclonales fueron purificados a partir de los sobrenadantes mediante cromatografía de afinidad en proteína G (HiTrap Protein G Sepharose High Performance; GE Healthcare).

20

25

#### Análisis por inmunodetección en membrana

30

El análisis por inmunodetección en membrana fue realizado con αGFP (clones de ratón 7.1 y 13.1, Roche), αCdc28 (anticuerpos policlonales de conejo, cedidos por C. Mann), y αYdj1 (clones de ratón 1G10.H8, Abnova). Las solubilidades de las fusiones de las etiquetas de la invención junto con GFP fue evaluada por centrifugación a 20.800 g durante 15 minutos de los extractos celulares preparados en tampón de lisis (50 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM DTT e inhibidores de proteasas y fosfatasa).

35

### Inmunofluorescencia

Las células HEK293T o NIH3T3 fueron rápidamente lavadas con PBS y fijadas en 4% formaldehído y 4% de sacarosa durante 30 minutos en temperatura ambiente. Las células fijadas fueron permeabilizadas con 0.1% triton en PBS durante 5 minutos, y bloqueadas con 1% de BSA en PBS. El bloqueo de las neuronas del hipocampo se llevó a cabo en 5% de serum de cabra en PBS. Los anticuerpos primarios contra HA (clon de ratas 3F10 de Roche), FLAG (clon de ratón M2, Sigma), ET5 (clon de ratón R19/8-11/18) y ET6 (clon de ratón R19/4-11/15) fueron utilizados con anticuerpos secundarios etiquetados con Alexa568 (Molecular Probes) en una solución bloqueante.  $\alpha$ GFP (clones de ratón 7.1 y 13.1, Roche) fue utilizada como referencia para el análisis de señal y de fondo.

### 15 Resultados

#### Screening bioinformático inicial para la selección de candidatos

Con el fin de identificar los polipéptidos con uso potencial como proteínas de marcaje con un mínimo riesgo de interferencia física o funcional con las proteínas de interés, se realizó una búsqueda de dominios de proteínas con estructura tridimensional conocida y con marcada capacidad inmunogénica, utilizando las bases de datos Protein Data Bank (PDB) y Protein Families Database (Pfam). Debido a la baja fiabilidad de la predicción sobre las propiedades antigénicas, la búsqueda utilizó como punto de partida las regiones inmunogénicas confirmadas experimentalmente obtenidas a partir del The Immune Epitope Database (IEDB) así como del Antigenic Variation Database (VarDB). Se aplicaron filtros funcionales para prevenir la selección de dominios insolubles, hidrolíticos, poliméricos o promiscuos, y se seleccionaron aquéllos con un índice de globularidad alto, pequeño tamaño y gran capacidad como inmunógenos basado en una combinación de factores físico-químicos. Adicionalmente, se consideraron también los parámetros de área de superficie accesible y flexibilidad en los extremos N o C-terminal para facilitar la construcción de la proteína de fusión resultante. Estas estructuras fueron filtradas en una ronda posterior para descartar todos aquellos dominios que estuviesen anotados como dominios con capacidad enzimática o con capacidad para unir cofactores o efectores. También se

rechazaron dominios con capacidad de formar homopolímeros o que tuviesen secuencias reguladoras o de degradación o localización celular. Finalmente, se descartaron los dominios derivados de un organismo vertebrado. Los 12 dominios resultantes con potencial para ser desarrollados como etiquetas antigénicas se listan en la Tabla 1.

10 Tabla 1 | Proteínas antigénicas candidatas.

Epitag	PDB id	Descripción	Organismo	aa	Globularidad
1	1BK8	Antimicrobial protein 1	Horse chestnut ( <i>Aesculus hippocastanum</i> )	50	0.654
2	1BW3	Barwin, basic seed protein	Barley ( <i>Hordeum vulgare</i> )	125	0.413
3	1ICX	Protein LLR18A	Yellow lupin ( <i>Lupinus luteus</i> )	155	0.450
4	1PJW	Envelope protein	Encephalitis virus	111	0.389
5	1WHP	Allergen Phl p2	Timothy grass ( <i>Phleum pratense</i> )	96	0.562
6	1WKX	Hevein isoform 2	Rubber tree ( <i>Hevea brasiliensis</i> )	43	0.598
7	1X6R	Fimbrial protein	<i>Pseudomonas aeruginosa</i>	123	0.507
8	2JMH	Mite allergen Blo t5	Mite ( <i>Blomia tropicalis</i> )	119	0.371
9	2JTY	Type-1 fimbrial protein, A chain	<i>Escherichia coli</i>	184	0.452
10	3BBG	Pollen allergen 5	Great ragweed ( <i>Ambrosia trifida</i> )	40	0.587
11	3FT9	Allergen Phl p3	Timothy grass ( <i>Phleum pratense</i> )	100	0.568
12	3K78	Major pollen allergen Bet v1-D/H	Birch ( <i>Betula verrucosa</i> )	160	0.424

#### Tests de viabilidad bioquímica de las etiquetas de marcaje *in vivo*

Las proteínas plegadas incorrectamente se acumulan normalmente como agregados que pueden ser fácilmente observados en el microscopio como cuerpos de inclusión. Las alteraciones bioquímicas que caracterizan a los agregados normalmente incluyen una reducida solubilidad (en los extractos celulares), una densidad más alta y menor movilidad de difusión (medido en un microscopio de fluorescencia) y una menor accesibilidad a los anticuerpos (medido mediante métodos de inmunofluorescencia). Así, para evaluar si las etiquetas que forman parte de la invención contienen determinantes estructurales claros que les permiten plegarse rápida y eficientemente, éstas se expresaron en altas concentraciones como proteínas de fusión junto con

la GFP en células de riñón embrionarias (HEK293T), y se analizaron las propensiones a formar agregados, así como sus patrones de localización en experimentos *in vivo* de microscopía de fluorescencia (FIG. 1a). La mayor parte de las fusiones con las etiquetas de la invención adquirieron niveles de expresión similares al del experimento control con GFP (FIG. 1b). Sólo una etiqueta (ET8) no era detectable, y no fue analizada seguidamente. Con respecto a su patrón intracelular (FIG. 1c), 3 de las etiquetas analizadas en profundidad (ET1, ET2, y ET12) mostraron una marcada acumulación en las células, acoplada a unos niveles muy bajos de difusión en el citoplasma. Otras tres etiquetas (ET7, ET9 y ET11) mostraron agregados en la mayoría de las células, aunque una importante fracción de la proteína de fusión produjo un patrón difuso en la célula. ET3 y ET4 mostraron un patrón difuso en la mayoría de las células, pero mostraron claramente agregados en algunas de ellas. Finalmente, los 3 dominios restantes (ET5, ET6 y ET10) produjeron un patrón difuso en todas las células que era indistinguible de aquél producido en el experimento con GFP de control, lo que sugirió que estos 3 dominios no perturbarían las propiedades físico-químicas y bioquímicas de la GFP *in vivo*. Así estos 3 dominios fueron seleccionados para pruebas posteriores.

A pesar de su pequeño tamaño, se infirió que la estructura de las tres etiquetas (ET5 que se corresponde con la proteína Phl P 2, ET6 que se corresponde con la proteína Hev b 6.02 y ET10 que se corresponde con la proteína Amb t 5) está basada en un empaquetamiento compacto basado en hojas beta que, en el caso de ET6 y ET10 está estabilizado por la presencia de puentes disulfuro. Las fusiones de estas tres etiquetas a la GFP se expresaron en las células HEK293T y se detectaron básicamente como una única banda con el peso molecular aparente esperado para cada caso en experimentos de western blot (FIG. 2d), y su estabilidad en presencia de cicloheximida se reveló como prácticamente idéntica a la de la GFP de control (FIG. 2e). Estos datos indican que estas tres etiquetas peptídicas no afectan significativamente *per se* a la integridad y a las tasas de degradación de la proteína de fusión.

La presencia de cualquiera de estas tres etiquetas no afectó a la solubilidad de la GFP en extractos celulares (FIG. 2e), ni tampoco modificó su tasa de difusión intracelular *in vivo* como muestran los resultados de análisis de

pérdida de fluorescencia (FIG.2f), confirmándose así la ausencia de asociaciones estables de las etiquetas de la invención con estructuras del citoesqueleto u otras estructuras celulares de gran tamaño. Finalmente, ya que ni la fluorescencia ni los niveles de proteína no se vieron afectados por la presencia de ninguna de las etiquetas peptídicas de la invención, se concluyó que estas tres etiquetas no afectan ni a la maduración ni al plegamiento de la GFP como modelo de proteína de interés.

#### 10 Inocuidad funcional de las etiquetas de marcaje de la invención en proteínas esenciales modelo

La fusión de etiquetas de marcaje en los extremos N o C-terminal de las proteínas de interés puede producir efectos inesperados en sus funciones, localización intracelular o asociación a complejos macromoleculares, particularmente si las etiquetas no tienen una estructura claramente definida y son intrínsecamente desordenadas. Las tres etiquetas que forman parte de la invención contienen claros determinantes estructurales y de ahí que deberían tener menos propensión a producir efectos inesperados en las proteínas de fusión bajo estudio cuando se comparan con otras etiquetas de marcaje como por ejemplo FLAG o HA, dos de las más ampliamente utilizadas. Para probar este último punto, se decidió realizar una serie de análisis comparativos de los efectos producidos por el etiquetaje en dos proteínas esenciales de la levadura *S. cerevisiae*, una quinasa dependiente de ciclina y una chaperona de dominio J.

Las quinasas dependientes de ciclina (Cdk) establecen múltiples interacciones para llevar a cabo una serie de procesos que desencadenan el ciclo celular, y la modificación de estas interacciones tiene un impacto profundo en su consecución. La levadura depende de una única Cdk, la Cdc28 para desencadenar todas las fases del ciclo celular de una manera coordinada con la maquinaria de crecimiento celular. El aislamiento de mutantes con un tamaño celular alterado ha sido crucial en el descubrimiento de procesos moleculares claves que gobiernan el ciclo celular y, a la vez, el tamaño celular es un indicador particularmente válido de los procesos moleculares en donde la Cdc28 juega un papel central.

Así, se decidió etiquetar el locus CDC28 endógeno con fusiones en el extremo C-terminal con las etiquetas ET5 (Phl P 2, 96 residuos aminoacídicos) y ET6 (Hev b 6.02, 43 residuos aminoacídicos), así como con repeticiones tándem de FLAG (6x, 62 residuos aminoacídicos) y HA (3x, 32 residuos aminoacídicos), para obtener *tags* (etiquetas) similares en cuanto al tamaño. Ninguna de estas etiquetas causó efectos significativos en los niveles de Cdc28 ni en su integridad (FIG.5a). Sin embargo, como se deduce de una inspección visual de las células de levadura (FIG.3b), la fusión Cdc28-FLAG causó un cambio muy importante en el tamaño celular y en la forma, indicando importantes alteraciones en las vías de señalización mediadas por Cdc28 en fases G1 y G2 del ciclo celular. A destacar, los mismos efectos se vieron cuando sólo 3 copias de la etiqueta FLAG se utilizaron para marcar Cdc28 en el extremo C-terminal. Este último resultado apunta a que algún determinante intrínseco de secuencia en la etiqueta FLAG es el responsable de los efectos deletéreos sobre la función biológica de Cdc28. En cambio, mientras que ET5 tiene un tamaño similar a la etiqueta 6xFLAG, éste no produjo ningún cambio en la morfología de las células (FIG.3b) y únicamente produjo un mínimo efecto en el tamaño medio crítico en la transición G1/S cuando se comparó con la etiqueta 6xFLAG (FIG. 3c). La etiqueta 3xHA no produjo ningún defecto morfológico a gran escala, pero indujo un aumento significativo del tamaño crítico en la transición G1/S (FIG.3c), demostrando una alteración clara de los mecanismos que controlan la entrada del ciclo celular en función del crecimiento celular. Por contra, la etiqueta ET6 (Hev b 6.02) (de tamaño muy parecido a 3xHA) no causó cambios significativos ni en la media, ni en los valores de distribución del tamaño crítico.

Las chaperonas moleculares utilizan la energía derivada de la hidrólisis de los enlaces fosfato del ATP para transmitir cambios conformacionales a sus proteínas cliente, facilitando así su plegamiento correcto y promoviendo el ensamblaje o desensamblaje de complejos multiproteína. Una de las clases más ubicuas de chaperonas es la familia de la Heat Shock 70 kDa (Hsp70s), que invariablemente requiere de la participación de las co-chaperonas de dominio J, las cuales estimulan su actividad ATPasa y cooperan en la unión de las proteínas cliente. Ydj1 es la chaperona de dominio J más abundante en la levadura y, a parte de sus actividades en la homeostasis del proteoma, juega un papel crucial en la liberación de la ciclina Cln3 del retículo

endoplasmático para desencadenar la entrada en el ciclo celular.

Adicionalmente a su dominio J localizado en el extremo N-terminal, Ydj1 contiene un dominio de dimerización en su extremo C-terminal que es esencial para su función. Por consiguiente, se decidió obtener fusiones en el extremo C-terminal del locus YDJ1 endógeno (FIG. 5b) y probar posibles efectos deletéreos de las etiquetas peptídicas ET5, ET6, 6xFLAG y 3xHA en la función de Ydj1. Ninguna de las etiquetas peptídicas causaron efectos detectables en los niveles de Ydj1 o en su integridad. De manera similar a Cdc28, la fusión de Ydj1 con la etiqueta 6xFLAG produjo los efectos más severos, los cuales pudieron ser observados incluso bajo condiciones normales de crecimiento a 30 grados Celsius (FIG. 3e). Además, tanto la etiqueta 6xFLAG como la etiqueta 3xHA produjeron una disminución clara del crecimiento a altas temperaturas, donde la función de la co-chaperona Ydj1 se convierte en esencial. En claro contraste, la fusión de Ydj1 con ET5 o ET6 produjo efectos muy moderados en el crecimiento a 37 grados Celsius (FIG. 3e). Ydj1 es también muy importante para coordinar el crecimiento y la entrada en el ciclo celular asociado al tamaño. En línea con sus efectos en la velocidad de crecimiento, tanto la etiqueta 6xFLAG como la etiqueta 3xHA causaron un aumento marcado en el tamaño crítico en la transición G1/S (FIG. 3f). Muy al contrario, ni ET5 ni ET6 produjeron un efecto obvio en el tamaño crítico, lo cual sugiere que estas etiquetas de marcaje son funcionalmente inocuas en cuanto al rol desempeñado por Ydj1 en los procesos de entrada del ciclo celular.

#### Las etiquetas de marcaje como herramientas inocuas para el análisis de proteínas *in situ*.

Los prospectos que habitualmente acompañan a los anticuerpos monoclonales disponibles en el mercado a menudo incluyen información sobre su uso para la detección de proteínas de interés en técnicas de inmunofluorescencia *in situ*. Sin embargo, muy raramente se proporciona una prueba fehaciente de su especificidad en células donde la expresión de la proteína de interés ha sido disminuida. Esta es una de las razones por las que etiquetas peptídicas como la FLAG o la HA son ampliamente utilizadas para confirmar la localización intracelular de la proteína bajo unas condiciones muy determinadas. Teniendo presente lo anterior, los inventores han ideado y implementado la producción eficiente de anticuerpos



monoclonales para la detección *in situ* de las etiquetas de marcaje de la invención, mediante técnicas de inmunofluorescencia.

5 Bien en forma monomérica (ET5, Phi P 2) o bien en forma dimérica (ET6 que corresponde a Hev b 6.02 y ET10 que corresponde a Amb t 5), las etiquetas y sus correspondientes fusiones GFP fueron expresadas y purificadas en células de *E. coli* (FIG. 6), para posteriormente inmunizar ratones y realizar una criba inicial mediante métodos ELISA indirectos. Los anticuerpos monoclonales que dieron positivos fueron probados a concentraciones bajas en células HIK293T mediante inmunofluorescencia para comprobar su sensibilidad y eficiencia (FIG.7) y, por encima de todo, una baja señal de fondo a altas concentraciones en células control (FIG.8). Estos anticuerpos son los que se generan mediante los hibridomas depositados en la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con los números de depósito DSM ACC3236, DSM ACC3234 y DSM ACC3242. Los anticuerpos monoclonales seleccionados fueron posteriormente utilizados para analizar una serie de fusiones con las etiquetas ET5 y ET6 a proteínas paradigmáticas que localizan en diferentes compartimentos celulares en fibroblastos 3T3 y neuronas del hipocampo de ratón (*Mus musculus*); en la FIG. 4 se muestran imágenes representativas. Independientemente de la etiqueta utilizada, todas las proteínas de fusión mostraron la localización intracelular esperada y, lo que es más importante, no se observaron agregados o fracciones de las proteínas en compartimentos anómalos que se pudiesen interpretar como una consecuencia de la interacción de las etiquetas peptídicas con las proteínas marcadas. En resumen, estos resultados demuestran la inocuidad funcional de las etiquetas de la invención en cuanto a la distribución de un variado grupo de proteínas que localizan en diversos compartimentos de la célula como son el núcleo, el aparato de Golgi, el citoesqueleto, el retículo endoplasmático, la membrana plasmática o la sinapsis neuronal.

#### REFERENCIAS CITADAS EN LA SOLICITUD

35 Terpe K. "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems" Appl. Microbiol Biotechnol. 2003, vol. 60., pp. 523-533

- Stadler, C. et al. "Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells" *Nat. Methods* 2013, vol. 10, pp. 315-323
- 5 Altschul S. "Basic local alignment search tool", *J. Mol. Biol.* 1990, vol. 215, pp. 403–410
- Hayes, C. N. et al. "varDB: a pathogen-specific sequence database of protein families involved in antigenic variation" *Bioinformatics* 2008, vol. 24, pp. 2564–2565
- Vita, R. et al. "The immune epitope database 2.0" *Nucl. Acids Res.* 2010, vol. 38, D854-D862
- 10 Punta, M. et al. "The Pfam protein families database". *Nucl. Acids Res.* 2012, vol. 40, D290-D301
- Humphrey, et al. "VMD - Visual molecular dynamics" *J. Molec. Graphics* 1996, vol. 14, pp. 33-38
- 15 GO, The Gene Ontology Consortium. "Gene ontology annotations and resources", *Nucl. Acids Res.* 2013, vol. 41, pp. D530-D535
- Gallego, C., et al. "The Cln3 cyclin is downregulated by translational repression and degradation during the G1 arrest caused by nitrogen deprivation in budding yeast". *EMBO J.* 1997, vol. 16, pp. 7196–7206
- 20 Sambrook, J. & Russell, D. W. "Molecular cloning: A laboratory manual". 2001 3a edición, CSHL Press, Cold Spring Harbor, NY
- Ferrezuelo, F. *et al.* "The critical size is set at a single-cell level by growth rate to attain homeostasis and adaptation". *Nat. Comm.* 2012, vol. 3, pp. 1012
- 25 Saha, S. & Raghava, G. P. S. "BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties" en Nicosia, G., Cutello, V., Bentley, P.J. & Timis, J. (eds.) pp. 197-204, Springer 2004, Heidelberg

## REIVINDICACIONES

1. Proteína recombinante de fusión que comprende una proteína o péptido que está fusionado en su extremo N-terminal o C-terminal a una proteína de marcate, donde la proteína de marcate está definida por las siguientes propiedades:
- 1.1) comprende un número de residuos aminoacídicos inferior a 100;
  - 1.2) es inmunogénica;
  - 1.3) tiene un índice de globularidad superior a 0.5, definido el índice de globularidad como la fracción de residuos con accesibilidad al solvente y calculado mediante el programa VMD;
  - 1.4) está libre de actividad enzimática;
  - 1.5) está libre de sitios de unión para cofactores o efectores;
  - 1.6) está libre de secuencias de degradación;
  - 1.7) está libre de secuencias de localización;
  - 1.8) es de origen vegetal;
  - 1.9) está libre de secuencias de homopolimerización o heteropolimerización;
  - 1.10) cuando está fusionada en el extremo N-terminal a la Proteína Verde Fluorescente y se expresa la proteína de fusión en células HEK293T, la proteína recombinante de fusión resultante mantiene los niveles de expresión, dichos niveles medidos mediante inmunodetección; mantiene la estabilidad estructural definida según niveles por inmunodetección durante 12 horas de tratamiento con 100 µg/ml de cicloheximida; mantiene la difusión intracelular *in vivo* medida mediante pérdida de fluorescencia durante fotoblanqueado, la fluorescencia medida mediante microscopía confocal de fluorescencia; y mantiene la solubilidad de la GFP en extractos celulares medida mediante centrifugación de extractos a 15000 g durante 15 minutos, *in vitro* e *in vivo*;
  - 1.11) cuando está fusionada en el extremo C-terminal de la proteína Cdc28 o de la proteína Ydj1 de *Saccharomyces cerevisiae* la proteína recombinante de fusión mantiene la morfología definida por microscopía de campo claro; y mantiene la viabilidad celular medida mediante plaqueo a 42°C;
  - 1.12) cuando está fusionada a la Proteína Verde Fluorescente y se expresa en *Escherichia coli*, mantiene los niveles de expresión, dichos niveles medidos mediante inmunodetección; mantiene la solubilidad de la GFP en extractos celulares medida mediante centrifugación de extractos a 15000 g

durante 15 minutos; y mantiene la eficiencia en la purificación por cromatografía de afinidad medida mediante inmunodetección;

- 1.13) mantiene los niveles de expresión, esta expresión medida mediante inmunofluorescencia y mantiene la localización celular determinada mediante microscopía confocal de fluorescencia de proteínas que se encuentran en diferentes compartimentos celulares, dichas proteínas seleccionadas del grupo que consiste en: beta-actina del citoesqueleto, la proteína STX6 del aparato de Golgi, HO1 del retículo endoplasmático, la HDAC2 nuclear y GRB2 de la membrana plasmática en células NIH3T3 de ratón, así como FMRP de la sinapsis neuronal en neuronas del hipocampo; y

donde la proteína de marcaje se selecciona del grupo consistente en un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:1, un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:2 y un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:3, o una combinación de las mismas.

2. Proteína recombinante de fusión según la reivindicación 1, en donde la proteína de marcaje es un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:1
3. Proteína recombinante de fusión según la reivindicación 1, en donde la proteína de marcaje es un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:2
4. Proteína recombinante de fusión según la reivindicación 1, en donde la proteína de marcaje es un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:3
5. Secuencia nucleotídica que codifica por la proteína recombinante de fusión según una cualquiera de las reivindicaciones 1 a 4.

6. *Cassette* de expresión por recombinación que comprende la secuencia nucleotídica según la reivindicación 5 operativamente unida a una secuencia de control de la expresión.
- 5 7. *Cassette* de expresión según la reivindicación 6, en donde la secuencia de control de la expresión comprende una secuencia diana para reguladores transcripcionales, una secuencia de unión al ribosoma y una secuencia terminadora de la transcripción.
- 10 8. Vector de expresión que comprende la secuencia nucleotídica según la reivindicación 5 o un *cassette* de expresión según la reivindicación 6.
9. Célula huésped transfectada o transducida con la secuencia nucleotídica según la reivindicación 5.
- 15 10. Método para obtener una proteína recombinante de fusión según cualquiera de las reivindicaciones 1-4, que comprende:
- 1) Fusionar al menos una secuencia nucleotídica que codifica para la proteína o péptido con una secuencia nucleotídica que codifica para la proteína de marcaje;
- 20 2) Clonar la secuencia resultante del paso 1 en una célula huésped;
- 3) Expresar la secuencia clonada en el paso 2 en una célula.
11. Método para purificar o aislar una proteína recombinante de fusión según cualquiera de las reivindicaciones 1-4, que comprende llevar a cabo los pasos según la reivindicación 10 y adicionalmente un paso en donde la proteína de fusión se purifica o aísla.
- 25 12. Método según la reivindicación 11 en donde el paso adicional de purificación o aislamiento de la proteína de fusión comprende un anticuerpo o fragmento del mismo que se une a la proteína de marcaje.
- 30 13. Método para detectar una proteína recombinante de fusión según cualquiera de las reivindicaciones 1-4, que comprende llevar a cabo los pasos según la reivindicación 10 y adicionalmente un paso en donde se añaden medios para la detección de la proteína de marcaje.
- 35

14. Método según la reivindicación 13 en donde los medios para la detección de la proteína de marcaje comprenden un anticuerpo o un fragmento del mismo.

5 15. Método según la reivindicación 14 en donde los medios para la detección de la proteína de marcaje comprenden un anticuerpo monoclonal.

16. Método según las reivindicaciones 12-15 en donde los medios se seleccionan de los anticuerpos que se obtienen mediante los hibridomas depositados en la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con los números de depósito DSM ACC3236, DSM ACC3234 y DSM ACC3242.

17. Anticuerpo monoclonal que se une específicamente a un fragmento de Phl P 2 de *Phleum pratense* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:1, y que se produce mediante el hibridoma depositado en la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con el número de depósito DSM ACC3234.

18. Anticuerpo monoclonal que se une específicamente a un fragmento de Hev b 6.02 de *Hevea brasiliensis* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:2, y que se produce mediante el hibridoma depositado en la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con el número de depósito DSM ACC3242.

19. Anticuerpo monoclonal que se une específicamente a un fragmento de Amb t 5 de *Ambrosia trifida* con una secuencia aminoacídica de como mínimo 70% de identidad con la SEQ ID NO:3, y que se produce mediante el hibridoma depositado en la institución "Leibniz-Institut DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" con el número de depósito DSM ACC3236.

FIG. 1a

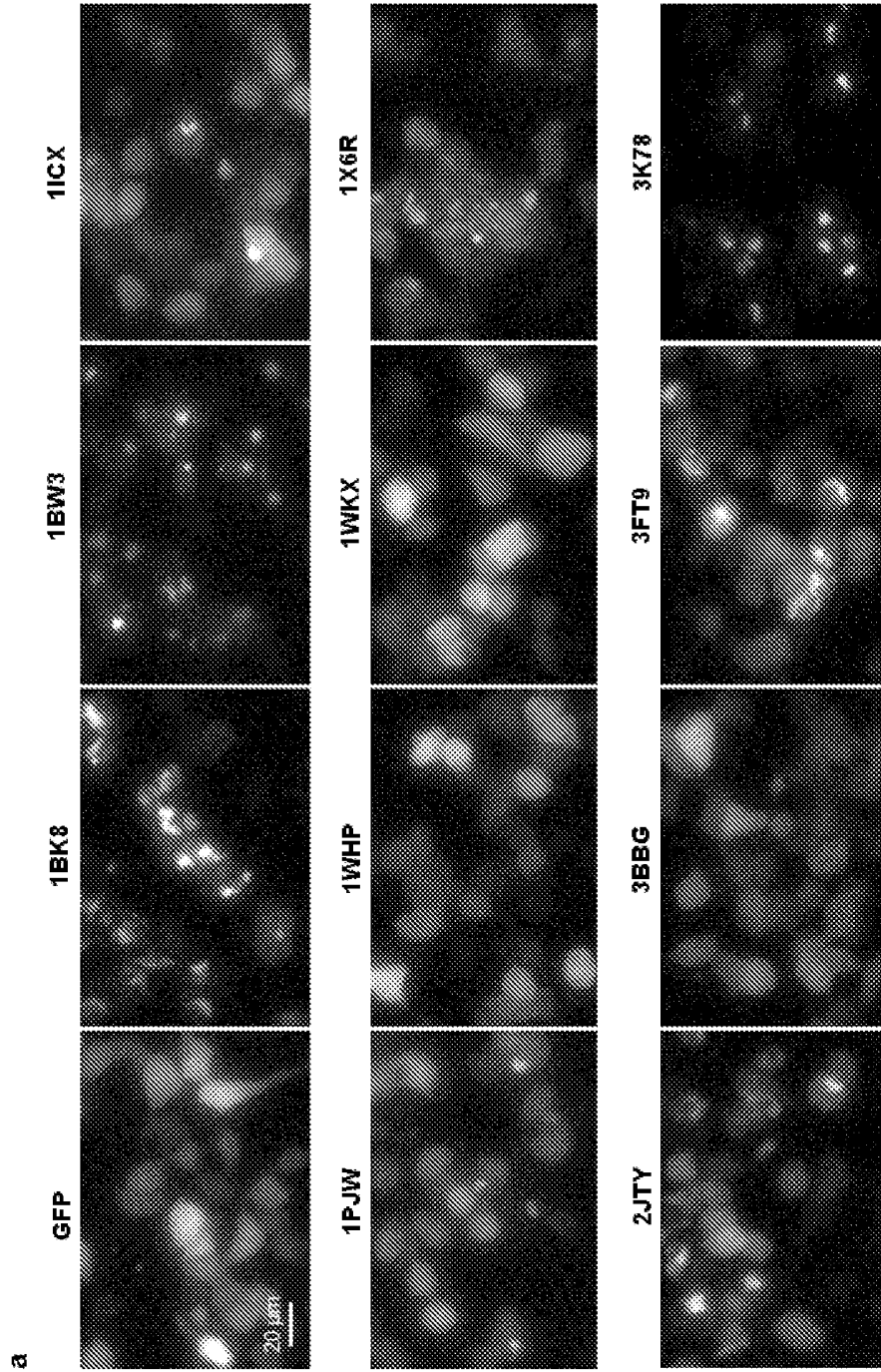


FIG. 1b

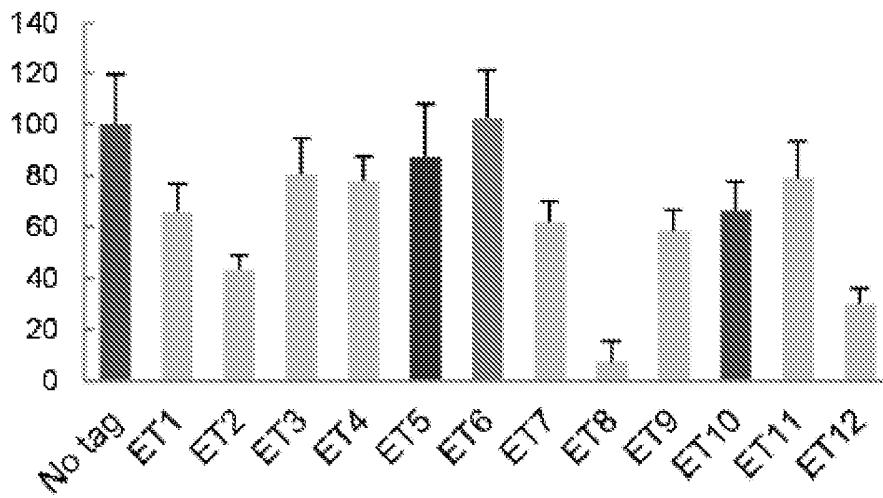


FIG. 1c

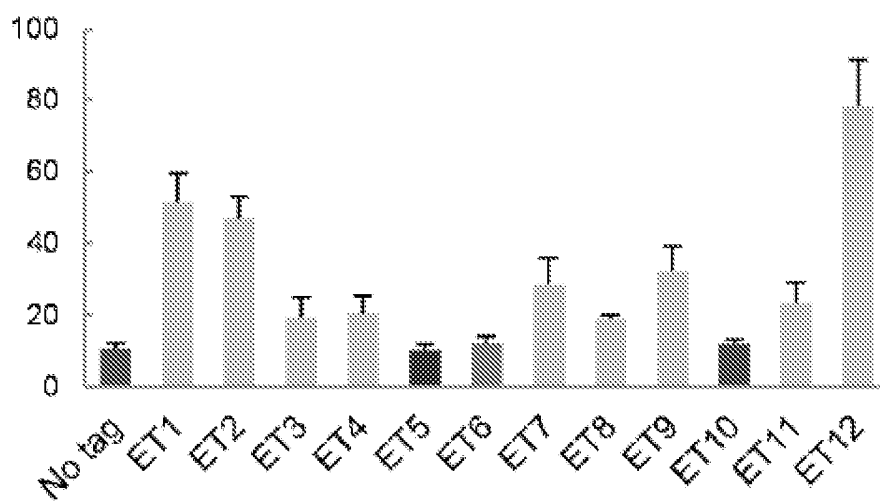
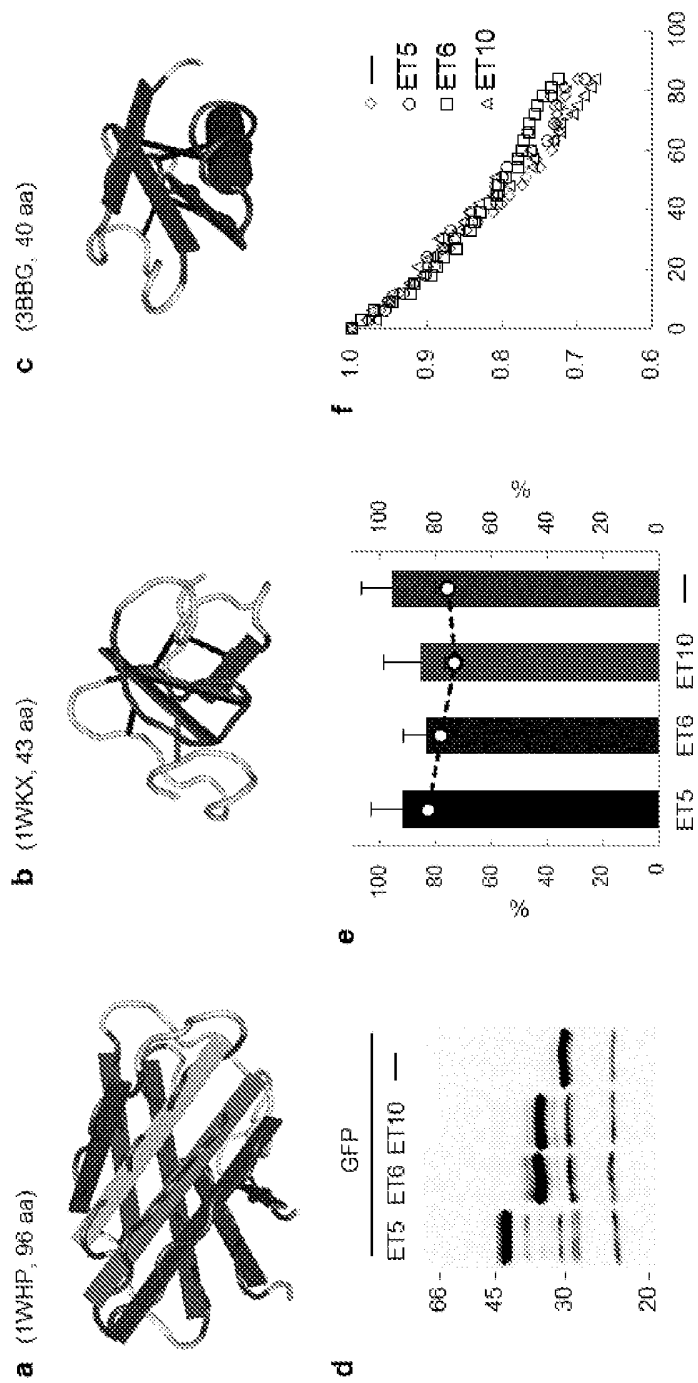




FIG. 2



4/10

FIG. 3a

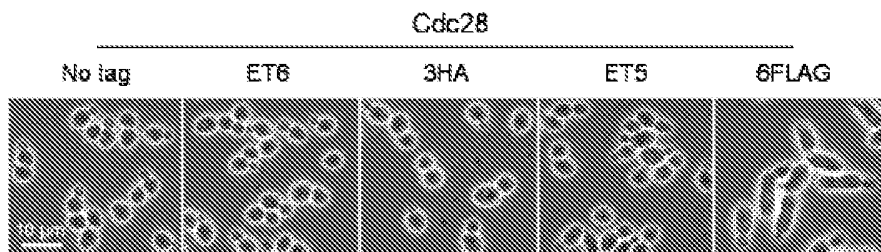


FIG. 3b

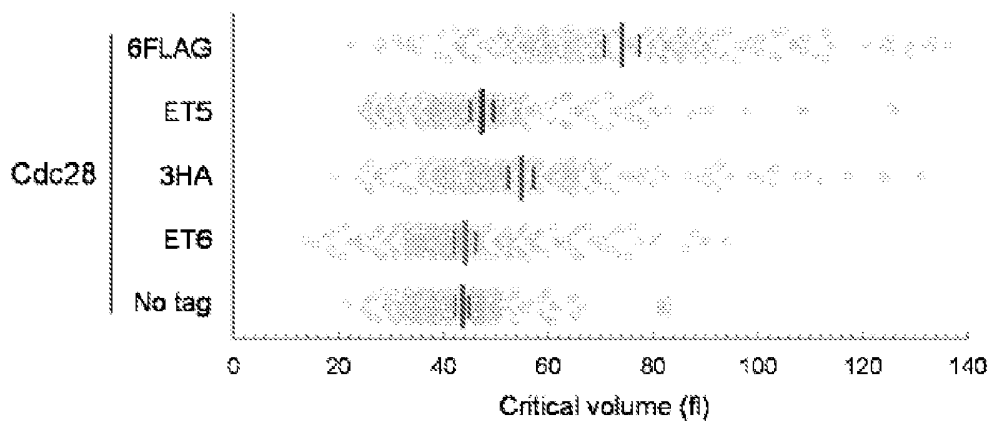


FIG.3c

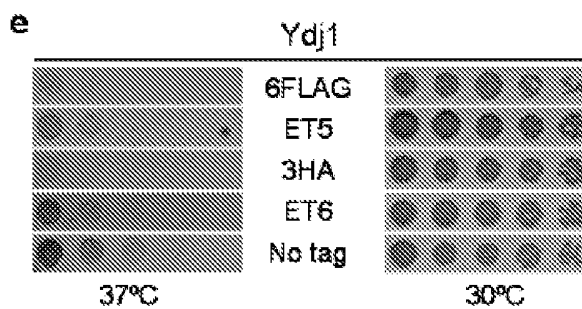


FIG.3d

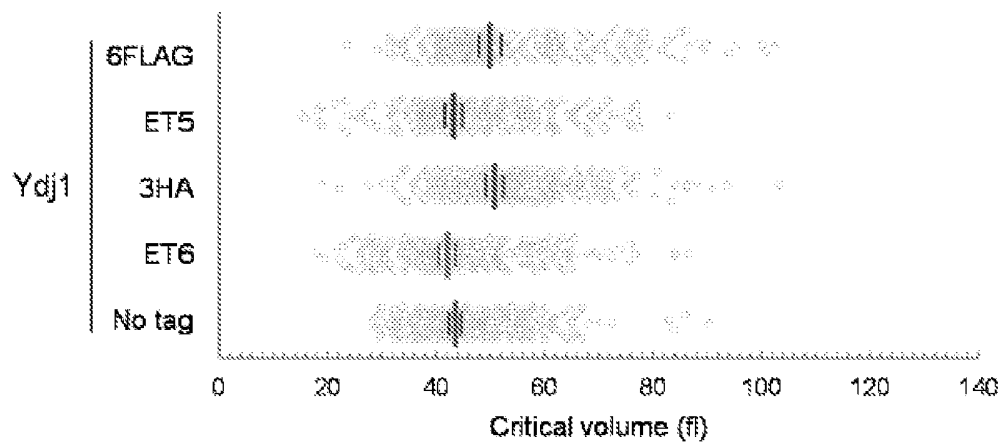


FIG. 4

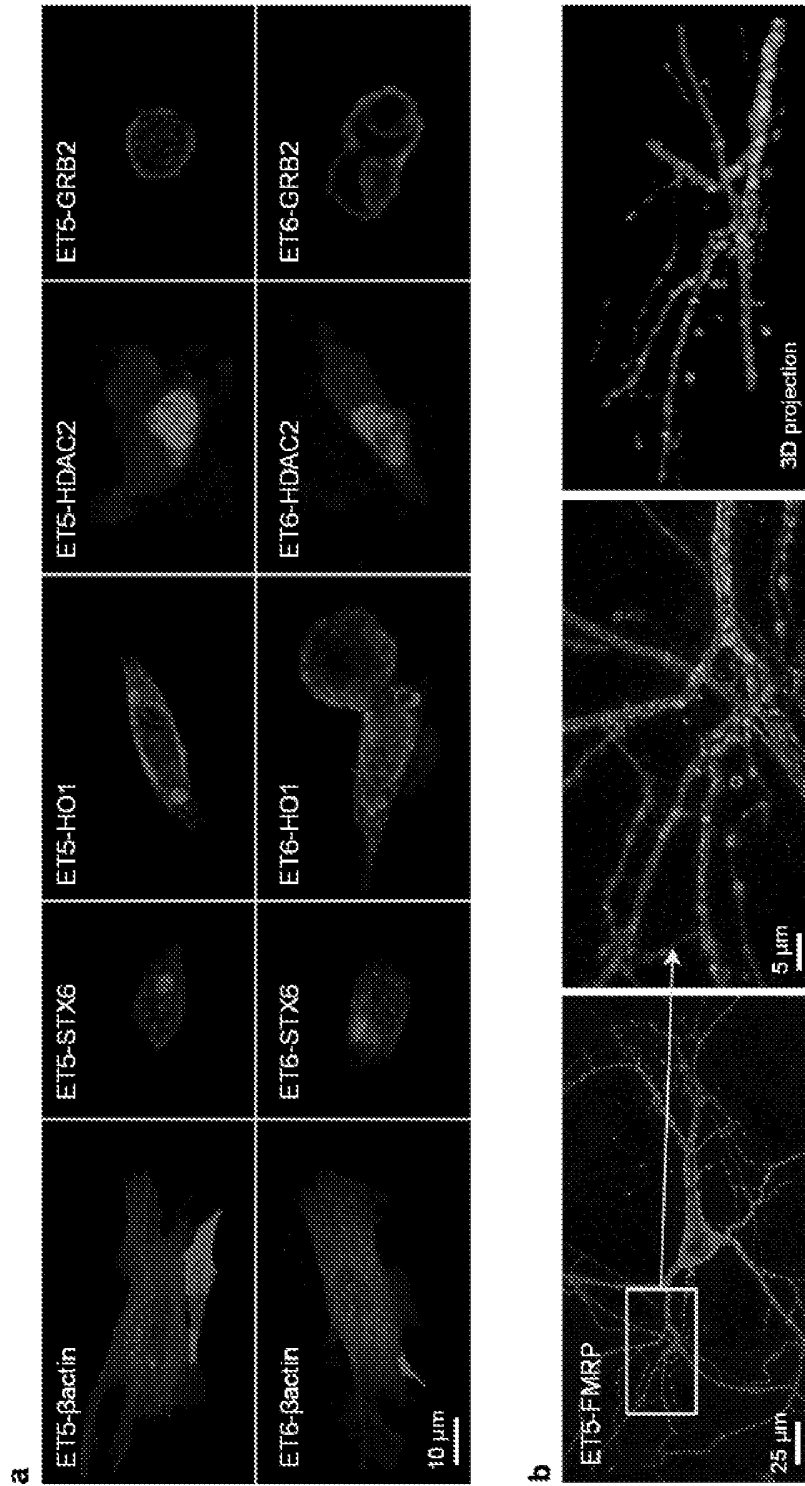
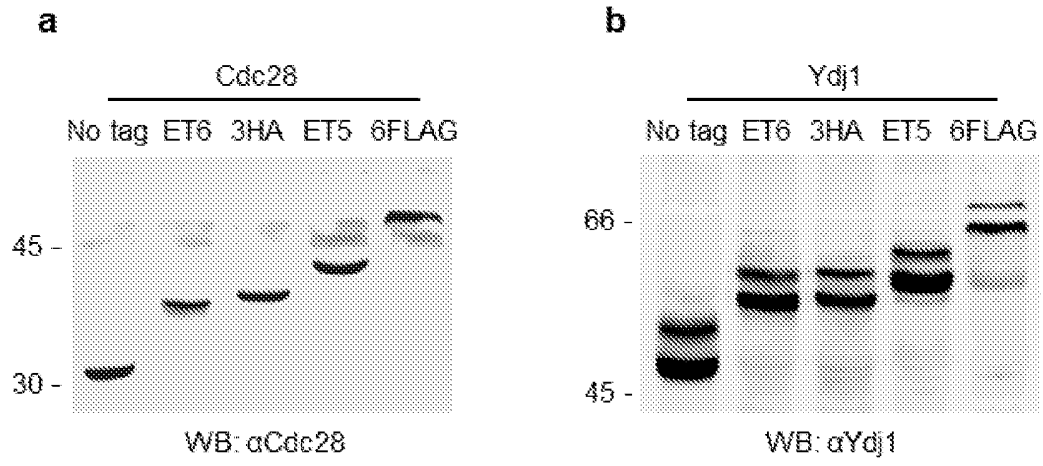
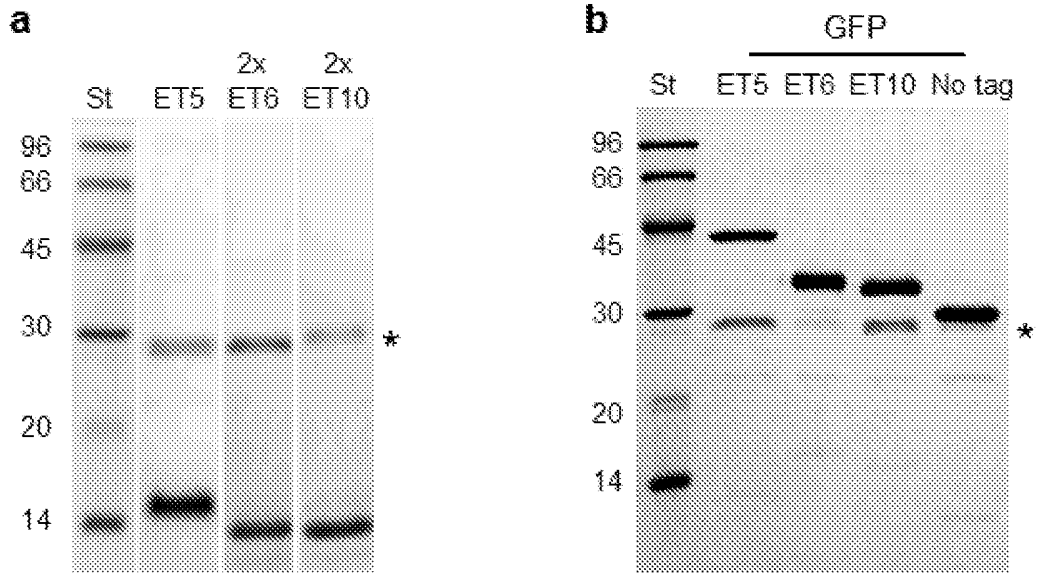


FIG. 5



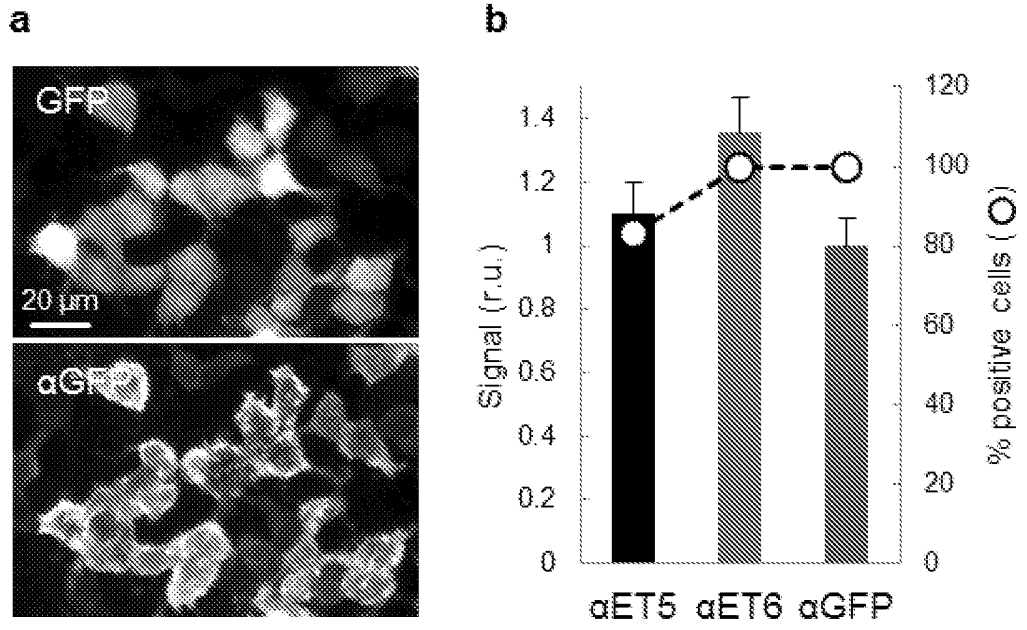
8/10

FIG. 6



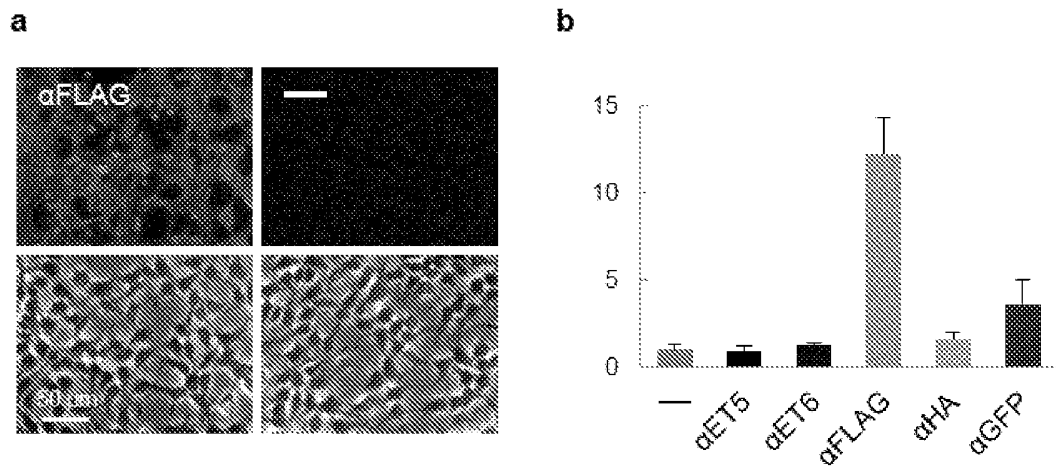
9/10

FIG. 7



10/10

FIG. 8





# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/ES2014/070460

## A. CLASSIFICATION OF SUBJECT MATTER

**See extra sheet**

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K, C12P, G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPODOC, INVENES, REGISTRY, HCAPLUS, BIOSIS, GOOGLE

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DOLECEK, C. et al. "Molecular characterization of <i>Phl p</i> II, a major Timothy grass ( <i>Phleum pratense</i> ) pollen allergen. FEBS. December 1993. Vol. 335, N° 3, pages 299-304, introduction and Fig. 2.	1, 2, 5-14
X	DE MARINO, S. et al. "An immunoglobulin-like fold in a major plant allergen: the solution structure of Phl p 2 from timothy grass pollen". STRUCTURE. 15.08.1999. Vol. 7, N° 8, pages 943-952, the whole document.	1, 2, 5-14
X	MARTH, K. et al. "Human monoclonal antibody-based quantification of group 2 grass pollen allergens". J. ALLERGY & CLIN. IMMUNOL. March 2004. Vol. 113, N° 3, pages 470-474, the whole document.	1, 2, 5-17

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance.</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure use, exhibition, or other means.</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>
--	--

Date of the actual completion of the international search  
27/03/2015

Date of mailing of the international search report  
**(27/03/2015)**

Name and mailing address of the ISA/

OFICINA ESPAÑOLA DE PATENTES Y MARCAS  
Paseo de la Castellana, 75 - 28071 Madrid (España)  
Facsimile No.: 91 349 53 04

Authorized officer  
M. Novoa Sanjurjo

Telephone No. 91 3498466

**CLASSIFICATION OF SUBJECT MATTER**

*C07K14/415* (2006.01)

*C07K16/16* (2006.01)

*C12P21/02* (2006.01)

*G01N33/68* (2006.01)

**INTERNATIONAL SEARCH REPORT**

International application No.

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

The International Searching Authority has found multiple inventions in this international application, as follows:

Invention 1 - claim 1 (in part); claim 2; claims 5-16 (in part); claim 17. Tagging protein derived from *Phleum pratense* and having SEQ ID No. 1, monoclonal antibody, purification method, etc.

Invention 2 - claim 1 (in part); claim 3; claims 5-16 (in part); claim 18. Tagging protein derived from *Hevea brasiliensis* and having SEQ ID No. 2, monoclonal antibody, purification method, etc.

Invention 3 - claim 1 (in part); claim 4; claims 5-16 (in part); claim 19. Tagging protein derived from *Ambrosia trifida* and having SEQ ID No. 3, monoclonal antibody, purification method, etc.

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
**Claim 1 (in part); claim 2; claims 5-16 (in part), claim 17.**

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/ES2014/070460

C (continuation).			DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of documents, with indication, where appropriate, of the relevant passages	Relevant to claim No.			
X	KAHLERT; H. et al. " <i>Phleum pratense</i> group 2-specific monoclonal antibodies – tools for identification and quantification in recombinant allergen preparations and allergen extracts". ALLERGY. November 2012. Vol. 67, suppl. 96, page 543, abstract 1450.	1, 2, 5-17			

# INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional nº

PCT/ES2014/070460

## A. CLASIFICACIÓN DEL OBJETO DE LA SOLICITUD

### Ver Hoja Adicional

De acuerdo con la Clasificación Internacional de Patentes (CIP) o según la clasificación nacional y CIP.

## B. SECTORES COMPRENDIDOS POR LA BÚSQUEDA

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

C07K, C12P, G01N

Otra documentación consultada, además de la documentación mínima, en la medida en que tales documentos formen parte de los sectores comprendidos por la búsqueda

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

EPODOC, INVENES, REGISTRY, HCAPLUS, BIOSIS, GOOGLE

## C. DOCUMENTOS CONSIDERADOS RELEVANTES

Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones nº
X	DOLECEK, C. et al. "Molecular characterization of <i>Phl p</i> II, a major Timothy grass ( <i>Phleum pratense</i> ) pollen allergen. FEBS. Diciembre 1993. Vol. 335, Nº 3, páginas 299-304, introducción y Fig. 2.	1, 2, 5-14
X	DE MARINO, S. et al. "An immunoglobulin-like fold in a major plant allergen: the solution structure of Phl p 2 from timothy grass pollen". STRUCTURE. 15.08.1999. Vol. 7, Nº 8, páginas 943-952, todo el documento.	1, 2, 5-14
X	MARTH, K. et al. "Human monoclonal antibody-based quantification of group 2 grass pollen allergens". J. ALLERGY & CLIN. IMMUNOL. Marzo 2004. Vol. 113, Nº 3, páginas 470-474, todo el documento.	1, 2, 5-17

En la continuación del recuadro C se relacionan otros documentos  Los documentos de familias de patentes se indican en el anexo

* Categorías especiales de documentos citados:	"T" documento ulterior publicado con posterioridad a la fecha de presentación internacional o de prioridad que no pertenece al estado de la técnica pertinente pero que se cita por permitir la comprensión del principio o teoría que constituye la base de la invención.
"A" documento que define el estado general de la técnica no considerado como particularmente relevante.	
"E" solicitud de patente o patente anterior pero publicada en la fecha de presentación internacional o en fecha posterior.	
"L" documento que puede plantear dudas sobre una reivindicación de prioridad o que se cita para determinar la fecha de publicación de otra cita o por una razón especial (como la indicada).	"X" documento particularmente relevante; la invención reivindicada no puede considerarse nueva o que implique una actividad inventiva por referencia al documento aisladamente considerado.
"O" documento que se refiere a una divulgación oral, a una utilización, a una exposición o a cualquier otro medio.	"Y" documento particularmente relevante; la invención reivindicada no puede considerarse que implique una actividad inventiva cuando el documento se asocia a otro u otros documentos de la misma naturaleza, cuya combinación resulta evidente para un experto en la materia.
"P" documento publicado antes de la fecha de presentación internacional pero con posterioridad a la fecha de prioridad reivindicada.	"&" documento que forma parte de la misma familia de patentes.

Fecha en que se ha concluido efectivamente la búsqueda internacional.  
27/03/2015

Fecha de expedición del informe de búsqueda internacional.  
**27 de marzo de 2015 (27/03/2015)**

Nombre y dirección postal de la Administración encargada de la búsqueda internacional  
OFICINA ESPAÑOLA DE PATENTES Y MARCAS  
Paseo de la Castellana, 75 - 28071 Madrid (España)  
Nº de fax: 91 349 53 04

Funcionario autorizado  
M. Novoa Sanjurjo  
Nº de teléfono 91 3498466

# INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional n°

PCT/ES2014/070460

## Recuadro II Observaciones cuando se estime que algunas reivindicaciones no pueden ser objeto de búsqueda (continuación del punto 2 de la primera hoja)

Este informe de búsqueda internacional no se ha realizado en relación a ciertas reivindicaciones según el artículo 17.2.a) por los siguientes motivos:

1.  Las reivindicaciones n°s:  
se refieren a un objeto con respecto al cual esta Administración no está obligada a proceder a la búsqueda, a saber:
  
2.  Las reivindicaciones n°s:  
se refieren a elementos de la solicitud internacional que no cumplen con los requisitos establecidos, de tal modo que no pueda efectuarse una búsqueda provechosa, concretamente:
  
3.  Las reivindicaciones n°s:  
son reivindicaciones dependientes y no están redactadas de conformidad con los párrafos segundo y tercero de la regla 6.4(a).

## Recuadro III Observaciones cuando falta unidad de invención (continuación del punto 3 de la primera hoja)

La Administración encargada de la Búsqueda Internacional ha detectado varias invenciones en la presente solicitud internacional, a saber:

**Invención 1.- (1) parcialmente, 2, (5-16) parcialmente, 17. Proteína de marcaje derivada de *Phleum pratense* de SEQ ID n° 1, anticuerpo monoclonal, procedimiento de purificación, etc.**

**Invención 2.- (1) parcialmente, 3, (5-16) parcialmente, 18. Proteína de marcaje derivada de *Hevea brasiliensis* de SEQ ID n° 2, anticuerpo monoclonal, procedimiento de purificación, etc.**

**Invención 3.- (1) parcialmente, 4, (5-16) parcialmente, 19. Proteína de marcaje derivada de *Ambrosia trifida* de SEQ ID n° 3, anticuerpo monoclonal, procedimiento de purificación, etc.**

1.  Dado que todas las tasas adicionales requeridas han sido satisfechas por el solicitante dentro del plazo, el presente informe de búsqueda de tipo internacional comprende todas las reivindicaciones que pueden ser objeto de búsqueda.
2.  Dado que todas las reivindicaciones que pueden ser objeto de búsqueda podrían serlo sin realizar un esfuerzo que justifique tasas adicionales, esta Administración no requirió el pago de tasas adicionales.
3.  Dado que tan sólo una parte de las tasas adicionales requeridas ha sido satisfecha dentro del plazo por el solicitante, el presente informe de búsqueda de tipo internacional comprende solamente aquellas reivindicaciones respecto de las cuales han sido satisfechas las tasas, concretamente las reivindicaciones n°s:
4.  Ninguna de las tasas adicionales requeridas ha sido satisfecha por el solicitante dentro de plazo. En consecuencia, el presente informe de búsqueda de tipo internacional se limita a la invención mencionada en primer término en las reivindicaciones, cubierta por las reivindicaciones n°s: **(1) parcialmente, 2, (5-16) parcialmente, 17.**

### Indicación en cuanto a la protesta

- Se acompañó a las tasas adicionales la protesta del solicitante y, en su caso, el pago de una tasa de protesta.
- Se acompañó a las tasas adicionales la protesta del solicitante, pero la tasa de protesta aplicable no se pagó en el plazo establecido para ello.
- El pago de las tasas adicionales no ha sido acompañado de ninguna protesta.

**CLASIFICACIONES DE INVENCION**

*C07K14/415* (2006.01)

*C07K16/16* (2006.01)

*C12P21/02* (2006.01)

*G01N33/68* (2006.01)

# INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional nº

PCT/ES2014/070460

C (Continuación).		DOCUMENTOS CONSIDERADOS RELEVANTES
Categoría *	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones nº
X	KAHLERT; H. et al. " <i>Phleum pratense</i> group 2-specific monoclonal antibodies – tools for identification and quantification in recombinant allergen preparations and allergen extracts". ALLERGY. Noviembre 2012. Vol. 67, supl. 96, página 543, resumen 1450.	1, 2, 5-17