

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 401 359**

51 Int. Cl.:

C07K 16/18 (2006.01)

C07K 16/46 (2006.01)

C07K 19/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.11.2006 E 08019358 (4)**

97 Fecha y número de publicación de la concesión europea: **26.12.2012 EP 2055718**

54 Título: **Direccionamiento y seguimiento de antígenos en células vivas**

30 Prioridad:

11.11.2005 EP 05024739

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

18.04.2013

73 Titular/es:

**LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
(100.0%)
GESCHWISTER-SCHOLL-PLATZ 1
80359 MÜNCHEN, DE**

72 Inventor/es:

**ROTHBAUER, ULRICH y
LEONHARDT, HEINRICH**

74 Agente/Representante:

ILLESCAS TABOADA, Manuel

ES 2 401 359 T3

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

DESCRIPCIÓN

Direccionamiento y seguimiento de antígenos en células vivas

5 La presente invención se refiere a la realización como se caracteriza específicamente en las reivindicaciones, la presente invención se refiere a una proteína de fusión que comprende una primera secuencia (poli) peptídica que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* (incluyendo cualquier camello o dromedario) y una segunda secuencia (poli) peptídica, que es una proteína detectable, preferentemente derivable de una proteína detectable, por ejemplo fluorescente, cromófora o fosforescente, en la que dicha (a) primera secuencia (poli) peptídica está codificada por la secuencia de ácido nucleico de SEQ ID NOs: 11 o 31; y (b) segunda secuencia (poli) peptídica, si deriva de una proteína fluorescente o cromófora, puede ser (i) la proteína verde fluorescente derivable de *Aequorea victoria* codificada por la secuencia de ácido nucleico de SEQ ID NOs: 7, o un mutante o fragmento fluorescente de la misma; (ii) la proteína roja fluorescente derivable de *Discosoma* (DsRed) codificada por la secuencia de ácido nucleico de SEQ ID NO: 9, o un mutante fluorescente o fragmento del mismo; o (iii) un homólogo funcional de (i) o (ii) con al menos 80% de identidad de secuencia; en la que dicha primera secuencia (poli) peptídica se localiza en dirección N terminal de dicha segunda secuencia (poli) peptídica, separándose opcionalmente dichas secuencias por un enlazador de al menos un resto de aminoácido. La presente memoria descriptiva también describe un método para purificar una estructura antigénica de interés, que comprende a) poner en contacto una muestra que contiene dicha estructura antigénica con I. una proteína de fusión dirigida a dicho estructura antigénica, en la que dicha proteína de fusión comprende una primera secuencia (poli) peptídica que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* y una segunda secuencia (poli) peptídica, que es una proteína detectable, preferentemente derivable de una proteína fluorescente o cromófora, en la que dicha (1.) primera secuencia (poli) peptídica está compuesta de marco 1, CDR1, marco 2, CDR2, marco 3, y CDR3, codificadas por la secuencia de ácido nucleico de SEQ ID NO: 2 o codificadas por una secuencia de ácido nucleico con al menos 70% de identidad de secuencia o un fragmento de la misma; y (2.) segunda secuencia (poli) peptídica es una proteína detectable, preferentemente (i) la proteína verde fluorescente derivable de *Aequorea victoria* codificada por la secuencia de ácido nucleico de SEQ ID NO: 7, o un mutante o fragmento fluorescente de la misma; (ii) la proteína roja fluorescente derivable de *Discosoma* (DsRed) codificada por la secuencia de ácido nucleico de SEQ ID NO: 9, o un mutante fluorescente o fragmento del mismo; o (iii) un homólogo funcional de (i) o (ii) con al menos 80% de identidad de secuencia; en la que dicha primera secuencia (poli) peptídica está localizada en dirección N terminal o C terminal de dicha segunda secuencia (poli) peptídica, separándose opcionalmente dichas secuencias por un enlazador de al menos un resto de aminoácido, o II. un (poli) péptido que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae*, compuesta de marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, codificado por la secuencia de ácido nucleico de SEQ ID NOs: 2 o codificado por una secuencia de ácido nucleico con al menos 70% de identidad de secuencia o un fragmento de la misma; en la que la proteína de fusión está unida a un soporte sólido; b) opcionalmente lavar el soporte sólido de la etapa a) para retirar constituyentes unidos de forma no específica; c) eluir la estructura antigénica.

40 Todas o cualquier combinación de etapas (incluyendo solamente etapas únicas) llevadas a cabo en el método de la presente invención y citadas a lo largo de la presente memoria descriptiva pueden llevarse a cabo en cualquier combinación de *in vivo*, *ex vivo* o *in vitro*.

45 Los anticuerpos son herramientas valiosas para identificar y visualizar estructuras celulares. Desafortunadamente, la aplicación de anticuerpos de origen natural para la detección de antígenos intracelulares requiere permeabilización (y con frecuencia fijación) de células. Además, la detección basada en anticuerpo de los antígenos dentro de células intactas se evita esencialmente por el hecho de que están, por naturaleza, diseñadas para actuar en un ambiente oxidante (extracelular); el ambiente reductor en el citoplasma conduce a una formación de enlaces disulfuro deficiente, dando como resultado un ensamblaje ineficaz de partes de reconocimiento de epítopos de la cadena ligera y pesada variable^{1,2}. Solamente en algunos casos se han usado anticuerpos intracelulares (ICAb) para afectar a la función proteica *in vivo* pero aún se sabe poco acerca de sus propiedades en células vivas³⁻⁷.

55 En un intento de evitar los problemas asociados con la aplicación de anticuerpos en el citoplasma de células intactas, se ha estudiado en el pasado la expresión proteica fusionando proteínas de interés con proteínas fluorescentes, habitualmente GFP ("marcaje con GFP"). El marcaje con GFP se ha convertido en un método extremadamente popular para estudiar el tráfico intracelular de proteínas y, en combinación con técnicas de fotoblanqueo de fluorescencia, ha proporcionado información única sobre la dinámica proteica en células vivas. Sin embargo, solamente puede medirse la dinámica de proteínas químicas, mientras que las proteínas auténticas, su modificación postraducciona así como componentes no proteicos de la célula no puede evaluarse por los métodos disponibles. Para superar estas limitaciones, sería deseable generar agentes de unión proteicos detectables que eviten los problemas y limitaciones de anticuerpos de origen natural y establecer su aplicación en la célula viva preferentemente evitando la interferencia con procesos celulares.

60 Por lo tanto, el problema técnico que subyace a la presente invención fue proporcionar nuevos métodos y compuestos que permitieran la detección intracelular de antígenos en células intactas. La solución a este problema técnico se consigue proporcionando las realizaciones caracterizadas en las reivindicaciones.

Se describe un método para detectar la presencia, cantidad o localización subcelular de una estructura antigénica de interés en una célula, que comprende las etapas de: (a) (i) expresar una proteína de fusión dirigida a la estructura antigénica de interés en dicha célula o (ii) introducir una proteína de fusión dirigida a la estructura antigénica de interés y acoplada a un (poli) péptido capaz de transducirse a dicha célula; comprendiendo dicha proteína de fusión una primera secuencia (poli) peptídica que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* y una segunda secuencia (poli) peptídica derivable de una proteína fluorescente o cromófora, en la que dicha (1.) primera secuencia (poli) peptídica está compuesta de marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, de un anticuerpo de *Camelidae* y, preferentemente, codificada por la secuencia de ácido nucleico de SEQ ID NO: 2 o codificada por una secuencia de ácido nucleico con al menos 70% de identidad o un fragmento de la misma y un segundo polipéptido, que es un (poli) péptido detectable (tal como un marcador) y preferentemente derivable de una proteína fluorescente o cromófora, en el que dicha (1.) primera secuencia (poli) peptídica está compuesta de marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, codificada por la secuencia de ácido nucleico de un anticuerpo de *Camelidae*, preferentemente de SEQ ID NO: 7, o un mutante o fragmento fluorescente de la misma; b. la proteína roja fluorescente derivable de *Discosoma* (DsRed) codificada por la secuencia de ácido nucleico de SEQ ID NO: 9, o un mutante o fragmento fluorescente de la misma; o c. un homólogo funcional de a. o b. con al menos 80% de identidad de secuencia; (b) revelar la presencia, cantidad o localización subcelular de dicha estructura antigénica de interés, si la hubiera, en dicha célula por medio de dicha proteína detectable; estando dicha primera secuencia (poli) peptídica en dirección N terminal de dicha segunda secuencia (poli) peptídica, separándose opcionalmente dichas secuencias por un enlazador de al menos un resto de aminoácido.

La expresión “localización subcelular” o “distribución” se refiere a la presencia de un compuesto, por ejemplo una proteína o un antígeno, dentro de la célula, en la célula o relacionada/conectada a la célula.

La expresión “estructura antigénica” o “antígeno” se refiere a cualquier tipo de compuesto capaz de inducir una respuesta inmune. Un antígeno puede ser un compuesto molecular pequeño o una macromolécula tal como una proteína o una molécula de ácido nucleico o fragmentos de las mismas así como combinaciones de moléculas grandes y moléculas pequeñas acopladas por ejemplo por enlace químico. En particular, dichos fragmentos pueden comprender o consistir en, por ejemplo, modificaciones post-traduccionales tales como fosforilación, metilación o glicosilación. En el contexto de la presente invención las expresiones se usan para describir dianas celulares de las proteínas de fusión usadas para llevar a cabo el método de la invención. La parte de anticuerpo, que deriva de anticuerpos de *Camelidae* puede dirigirse a cualquier estructura, contra la que pueden inducirse anticuerpos.

De acuerdo con la presente invención dicha proteína de fusión puede prepararse por un método para generar una proteína detectable capaz de unirse a un antígeno de interés, que comprende las etapas de: (a) obtener de una célula productora de anticuerpos de *Camelidae* o un grupo de tales células, una primera molécula de ácido nucleico o un grupo de tales moléculas de ácido nucleico, que codifica la región variable de una inmunoglobulina o producir de forma recombinante o (semi) sintética dicha primera molécula de ácido nucleico o grupo de primeras moléculas de ácido nucleico; (b) opcionalmente seleccionar de dicho grupo una molécula de ácido nucleico particular que codifica la región variable de una inmunoglobulina específica; (c) fusionar la región codificante de la primera molécula de ácido nucleico, que codifica la región variable de una inmunoglobulina en fase con la región codificante de una segunda molécula de ácido nucleico, que codifica un (poli) péptido marcador detectable, en la que la región codificante de la primera molécula de ácido nucleico está localizada 5' de la región codificante de la segunda molécula de ácido nucleico y en la que las regiones codificantes se separan opcionalmente por una región codificante que codifica un enlazador de al menos un resto de aminoácido; y (d) expresar la molécula de ácido nucleico fusionada que codifica la proteína de fusión en una célula o extracto sin células.

A lo largo de la presente invención la expresión “dirigido a” se usa para describir la propiedad de una proteína de fusión o (poli) péptido como se ha mencionado en la presente memoria descriptiva para unirse específicamente a una estructura antigénica. Esta unión específica excluye o esencialmente excluye la unión de antígenos cercanamente relacionados estructuralmente habitualmente conocida en la técnica como reactividad cruzada. La unión específica puede medirse, por ejemplo, en ensayos de ELISA u otros ensayos inmunológicos en los que la estructura antigénica de interés se une al pocillo de una placa de ELISA y se unen antígenos estructuralmente relacionados a diferentes pocillos de esta placa de ELISA. Si la proteína de fusión se une solamente a la estructura antigénica (antígeno) de interés, no habrá señal más allá del nivel de fondo o esencialmente más allá del nivel de fondo (es decir, como máximo 25% por encima del nivel de fondo) cuando se evalúe la unión con respecto a los antígenos relacionados.

La expresión “(poli) péptido capaz de transducirse en dicha célula” se refiere a la capacidad de ciertas proteínas para transducirse, es decir translocarse, del exterior de una célula al interior. El mecanismo aún es en gran medida desconocido. Los (poli) péptidos capaces de transducirse son principalmente básicos. La proteína de fusión se translocará a la célula junto con la molécula de transducción.

El término “(poli) péptido” se refiere alternativamente a péptidos o a polipéptidos. Los péptidos convencionalmente son aminoácidos ligados covalentemente de hasta 30 restos, mientras que los polipéptidos (también denominados en el presente documento “proteínas”) comprenden 31 y más restos de aminoácidos.

Las expresiones “proteína detectable” y “(poli) péptido detectable” y “(poli) péptido marcador detectable” se refieren a (poli) péptidos que son detectables por medios visuales o dan lugar a un producto detectable por medios visuales (tal como sucede con las enzimas). Por ejemplo, tras la excitación o mediante escisión de un sustrato cromogénico, dichas proteínas/(poli) péptidos dan como resultado la emisión de energía o color detectable. En este sentido, la proteína debe entenderse como una proteína marcada de forma detectable, correspondiente al segundo (poli) péptido (que es el marcador). En una realización preferida de la presente invención dichas proteínas detectables o (poli) péptidos marcadores detectables son (poli) péptidos fluorescentes, fosforescentes o cromóforos. De hecho, cualquier (poli) péptido que dé como resultado, después de exposición a una energía de excitación, la emisión de energía detectable, se entiende de acuerdo con la presente invención como un marcador detectable. En una realización particular de la presente invención, la energía emitida desde una primera proteína puede transferirse a una segunda proteína detectable (diferente) que también es excitable y que puede emitir la energía obtenida de la primera proteína detectable. Un ejemplo de dicha resonancia de energía es el sistema de FRET que puede adaptarse de acuerdo con las enseñanzas de la presente invención. La expresión “energía emitida...” preferentemente se refiere a fluorescencia, fosforescencia o la emisión de luz con una longitud de onda particular que es detectable.

El término “derivable” a lo largo de la presente invención también abarca el término “derivado”.

La expresión “célula productora de anticuerpos de *Camelidae*” se refiere a cualquier célula de *Camelidae* capaz de producir anticuerpos. Dichas células incluyen células del sistema linfático tales como células del bazo. En particular, las células productoras de anticuerpos pueden obtenerse o derivarse de linfocitos B. Preferentemente, la primera molécula de ácido nucleico se obtiene de linfocitos de sangre periférica (PBL). Sin embargo, puesto que la información genética completa está presente en cualquier célula del organismo, al menos a nivel teórico, la primera molécula de ácido nucleico también puede derivar de células no linfáticas. Incluso aunque las enseñanzas de la presente invención se realizan preferentemente usando células que han experimentado recombinación somática para generar exones de región V intactos, debe observarse que también puede generarse un exón de región V intacto por técnicas de ADN recombinante. El término *Camelidae* se refiere a la familia de *Camelidae* que incluye el género *Camelus* que comprende las especies *Camelus bactrianus* (Camello Bactriano) y *Camelus dromedarius* (dromedario), el género *Lama* que comprende las especies *Lama glama* (llama), *Lama guanicoe* (guanaco) y *Lama pacos* (alpaca) y el género *Vicugna* que comprende la especie *Vicugna vicugna* (Vicuña). Los *Camelidae* preferidos incluyen *Camelus bactrianus*, *Camelus dromedarius* y *Lama pacos*.

La expresión “molécula de ácido nucleico” como se usa a lo largo de la memoria descriptiva de la presente invención se refiere a ADN o ARN, incluyendo ADN genómico, ADNc, ARNm, ARNhn.

La expresión “obtener de... una primera molécula de ácido nucleico” se refiere al uso de la molécula de ácido nucleico que codifica la región variable de la inmunoglobulina. Un ejemplo de la obtención de una molécula de ácido nucleico es el aislamiento físico de dicha molécula de ácido nucleico de una célula. Dicha molécula de ácido nucleico puede por ejemplo obtenerse usando clonación convencional y procedimientos de exploración, tales como los de clonación de ADNc usando ARNm como material de partida.

Otro ejemplo de obtención de una molécula de ácido nucleico implica amplificación de ácidos nucleicos. De hecho, en muchos casos, una primera etapa de aislamiento físico de la molécula de ácido nucleico puede preceder a una segunda etapa de amplificación de ácido nucleico. En el caso de que la molécula de ácido nucleico sea una molécula de ARN tal como ARNhn o ARNm, la secuencia de nucleótidos de dicha molécula puede transcribirse de forma inversa a ADN.

El término “amplificación” o “amplificar” significa aumentar el número de copias. El experto en la materia conocerá diversos métodos para amplificar moléculas de ácido nucleico, estos métodos también pueden usarse en los métodos de la presente invención. Los métodos de amplificación incluyen, pero sin limitación, “reacción en cadena de la polimerasa” (PCR), “reacción en cadena de la ligasa” (LCR, documento EPA320308), “reacción de sonda cíclica” (CPR), “amplificación por desplazamiento de cadena” (SDA, Walker *et al.* 1992, Nucleic Acid Res. 7: 1691-1696), “sistemas de amplificación basados en transcripción” (TAS, Kwoh *et al.* 1989, Proc. Nat. Acad. Sci. USA 86: 1173; Gingeras *et al.*, solicitud de PCT WO 10315/88). Preferentemente, la amplificación de ADN se consigue usando reacción en cadena de la polimerasa (PCR) [Methods in Molecular Biology, Vol. 226 (Bartlett J. M. S. y Stirling D., eds.): PCR protocols, 2ª edición; PCR Technology: Principles and Applications for DNA Amplification (Erich H. a., ed.), Nueva York 1992; PCR Protocols: A guide to methods and applications (Innis M. A. *et al.*, eds.), Academic Press, San Diego 1990]. Los métodos de amplificación de ácido nucleico pueden ser particularmente útiles en casos en los que la muestra contiene solamente cantidades pequeñas de ácido nucleico. Si dicho ácido nucleico es ARN, puede realizarse una RT-PCR. Posteriormente, puede realizarse otra etapa de amplificación que implique PCR. Como alternativa, si dicho ácido nucleico contenido en la muestra es ADN, puede realizarse PCR.

La PCR, en general, consiste en muchas repeticiones de un ciclo que consiste en: (a) una etapa de desnaturalización, que funde ambas hebras de una molécula de ADN; (b) una etapa de hibridación, que se dirige a permitir que los cebadores hibriden específicamente con las hebras fundidas de la molécula de ADN; y (c) una etapa de extensión, que elonga los cebadores hibridados usando la información proporcionada por la hebra molde. En

general, puede realizarse PCR por ejemplo en una mezcla de reacción de 50 μ l que contiene 5 μ l de tampón de PCR 10x con $MgCl_2$ 1,5 mM, 200 μ M de cada desoxinucleósido trifosfato, 0,5 μ l de cada cebador (10 μ M), aproximadamente 10 a 100 ng de ADN molde y de 1 a 2,5 unidades de Taq polimerasa. Los cebadores para la amplificación pueden estar marcados o no estar marcados. La amplificación de ADN puede realizarse, por ejemplo, con un termociclador modelo 2400 (Applied Biosystems, Foster City, CA): 2 minutos a 94 °C, seguido de 35 ciclos consistentes en hibridación (30 segundos a 50 °C), extensión (1 minuto a 72 °C), desnaturalización (10 segundos a 94 °C) y una etapa de hibridación final a 55 °C durante 1 minuto así como una etapa de extensión final a 72 °C durante 5 minutos. Sin embargo, el experto en la materia sabe cómo optimizar estas condiciones para la amplificación de moléculas de ácido nucleico específicas o reducir la escala o aumentar el volumen de la mezcla de reacción.

Un método adicional de amplificación de ácido nucleico es la "reacción en cadena de la polimerasa de transcriptasa inversa" (RT-PCR). Este método se usa cuando el ácido nucleico para amplificar consiste en ARN. La expresión "transcriptasa inversa" se refiere a una enzima que cataliza la polimerización de desoxirribonucleósido trifosfatos para formar productos de extensión de cebadores que son complementarios de un molde de ácido ribonucleico. La enzima inicia la síntesis en el extremo 3' del cebador y continúa hacia el extremo 5' del molde hasta que termina la síntesis. Son ejemplos de agentes de polimerización adecuados que convierten la secuencia diana de ARN en una secuencia de ADN copia (ADNc) complementaria, la transcriptasa inversa del virus de mieloblastosis aviar y ADN polimerasa de *Thermus thermophilus*, una ADN polimerasa termoestable con actividad transcriptasa inversa comercializada por Perkin Elmer. Típicamente, el molde bicatenario de ARN/ADNc genómico se desnaturaliza por calor durante la primera etapa de desnaturalización después de la etapa de transcripción inversa inicial dejando la hebra de ADN disponible como un molde de amplificación. Las polimerasas adecuadas para su uso con un molde de ADN incluyen, por ejemplo, ADN polimerasa I de *E. coli* o su fragmento Klenow, ADN polimerasa T4, Tth polimerasa y Taq polimerasa, una ADN polimerasa estable frente al calor aislada de *Thermus aquaticus* y desarrollada y fabricada por Hoffmann-La Roche y disponible en el mercado de Perkin Elmer. Esta última enzima se usa ampliamente en la amplificación y secuenciación de ácidos nucleicos. Las condiciones de reacción para usar Taq polimerasa se conocen en la técnica y se describen, por ejemplo, en: PCR Technology, Erlich, H. A. 1989, Stockton Press, Nueva York; o en: Innis, M. A., D. H. Gelfand, J. J. Sninsky y T. J. White. 1990, PCR Protocols: A guide to methods and applications. Academic Press, Nueva York. La RT de alta temperatura proporciona mayor especificidad de cebadores y mejor eficacia. La Solicitud de Patente de Estados Unidos N° de serie 07/746.121, presentada el 15 de agosto de 1991, describe una "RT-PCR homogénea" en la que los mismos cebadores y polimerasa son suficientes para las etapas tanto de transcripción inversa como de amplificación por PCR, y las condiciones de reacción se optimizan de modo que ambas reacciones se produzcan sin cambio de reactivos. La ADN polimerasa de *Thermus thermophilus*, una ADN polimerasa termoestable que puede actuar como una transcriptasa inversa, puede usarse para todas las etapas de extensión de cebadores, independientemente del molde. Ambos procesos pueden realizarse sin tener que abrir el tubo para cambiar o añadir reactivos; solamente debe ajustarse el perfil de temperatura entre el primer ciclo (molde de ARN) y el resto de los ciclos de amplificación (molde de ADN). La reacción de RT puede realizarse, por ejemplo, en una mezcla de reacción de 20 μ l que contiene: 4 μ l de tampón ANV-RT 5x, 2 μ l de Oligo dT (100 μ g/ml), 2 μ l de dNTP 10 mM, 1 μ l de ARN total, 10 unidades de transcriptasa inversa AMV, y H_2O hasta un volumen final de 20 μ l. La reacción puede, por ejemplo, realizarse usando las siguientes condiciones: la reacción se mantiene a 70 °C durante 15 minutos para permitir la transcripción inversa. La temperatura de reacción se eleva después a 95 °C durante 1 minuto para desnaturalizar la doble cadena ARN-ADNc. A continuación, la temperatura de reacción experimenta dos ciclos de 95 °C durante 15 segundos y 60 °C durante 20 segundos seguido de 38 ciclos de 90 °C durante 15 segundos y 60 °C durante 20 segundos. Finalmente, la temperatura de reacción se mantiene a 60 °C durante 4 minutos para la etapa de extensión final, se enfría a 15 °C, y se mantiene a esa temperatura hasta el procesamiento posterior de la muestra amplificada. Cualquiera de las condiciones de reacción anteriormente mencionadas puede aumentarse de escala de acuerdo con las necesidades del caso particular.

La expresión "un grupo de dichas moléculas de ácido nucleico" se refiere a dos o más moléculas de ácido nucleico que pueden obtenerse de una pluralidad de células, preferentemente de forma simultánea. Estas moléculas pueden ser idénticas en secuencia o pueden diferir. En una realización preferida de la presente invención, las moléculas de ácido nucleico contenidas en dicho grupo de moléculas de ácido nucleico solo difieren en los codones de la llamada "región determinante de complementariedad" (CDR) que forma una región crucial dentro del bolsillo de unión de la inmunoglobulina, que es responsable de determinar su especificidad antigénica.

El término "inmunoglobulina" como se usa en el presente documento se refiere a cualquier molécula que comprenda un pliegue de inmunoglobulina o dominio de inmunoglobulina. Los dominios o pliegues de inmunoglobulina son elementos característicos de las proteínas de la superfamilia de proteínas de inmunoglobulina que incluyen moléculas tales como anticuerpos, receptores de linfocitos T y moléculas del MHC. Hay dos tipos principales de dominio de inmunoglobulina, dominios C con una lámina de tres cadenas y una de cuatro cadenas, y dominios V con una cadena extra en cada lámina.

La expresión "la región variable de una inmunoglobulina" o "el dominio variable de una inmunoglobulina" se refiere al dominio más amino-terminal de una inmunoglobulina, que se forma por recombinación de segmentos génicos V, D y

J durante el desarrollo de linfocitos. La región variable comprende tres regiones de variabilidad de secuencia particular que se denominan regiones hipervariables y se indican con HV1, HV2 y HV3. El resto de los dominios V muestran menos variabilidad de secuencia y las regiones entre las regiones hipervariables, que son relativamente invariantes, se denominan regiones marco conservadas (FR1, FR2, FR3 y FR4). Las regiones hipervariables están formadas por bucles en la estructura del dominio V, que forman juntos el sitio de unión a antígenos de la inmunoglobulina. Como los tres bucles hipervariables constituyen el sitio de unión para el antígeno y determinan la especificidad formando una superficie complementaria al antígeno, se denominan más habitualmente regiones determinantes de complementariedad, o CDR, y se indican con CDR1, CDR2 y CDR3. En el caso de los anticuerpos de cadena pesada de origen natural (HCAb) observados en *Camelidae*⁸, el sitio de unión a antígeno se forma solamente a partir de la región determinante de complementariedad de la cadena pesada, mientras que en anticuerpos convencionales el sitio de unión a antígeno se forma a partir de la región determinante de complementariedad de la cadena tanto ligera como pesada.

Como se usa en el presente documento, la región variable comprende preferentemente los restos de aminoácidos 1 a 117 como se codifican por la secuencia de ácido nucleico de SEQ ID NO: 1.

Sin embargo, la expresión "región variable" también se refiere a mutantes y fragmentos de los mismos que contienen una o más mutaciones que conservan preferentemente la especificidad de unión de la región variable no mutada que puede ser una región o tramo de aminoácidos de origen natural. El término "mutación" como se usa a lo largo de la memoria descriptiva de la presente invención se refiere a delección, adición o sustitución de un resto de aminoácido. "Uno o más" significa 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 pero también hasta 15 o incluso 20 restos de aminoácidos. La mutación puede residir en posición N o C terminal o dentro de la región variable. Preferentemente, la región variable comprende tres CDR intactas. Sin embargo, es concebible que en casos particulares aún sea posible unión de antígenos suficiente con solamente dos o incluso una CDR. En tales casos se prefiere que una de las CDR restantes sea CDR3. Un aspecto central de la presente invención es la provisión de moléculas de fusión capaces de reconocer un antígeno. Para mejorar el reconocimiento de antígenos, puede ser deseable sustituir uno o más restos de aminoácidos. Preferentemente, esta sustitución puede realizarse en CDR1, CDR2 y/o CDR3. Puede elegirse un enfoque aleatorio o un enfoque selectivo para generar estos mutantes de sustitución. En ocasiones puede ser necesario intercambiar restos de aminoácidos adicionales fuera de las CDR para generar anticuerpos con suficiente estabilidad o suficiente afinidad a un antígeno dado. A diferencia de los dominios VH de anticuerpos convencionales, los dominios VHH de anticuerpos de cadena pesada de camello se expresan de forma eficaz como proteínas recombinantes no agregantes solubles debido a sustituciones hidrófilas únicas en el marco 2: V37F/Y, G44E, L45R. Esto muestra la importancia de los restos de aminoácidos fuera de las CDR para la estabilidad y solubilidad de los fragmentos. Una mutagénesis dirigida de uno o más restos de aminoácidos podrían conducir a una mejora de la estabilidad o solubilidad. También pueden introducirse sustituciones de aminoácidos fuera de las CDR para humanizar VHH de camélido (esto se resume en una revisión de Muyldermans S., 2001, *Reviews in Molecular Biology*, 74:277-302). Como se usa en el presente documento, los fragmentos que comprenden el dominio o región variable también se describen en ocasiones como "fragmentos de reconocimiento de antígeno".

La expresión "seleccionar... una molécula de ácido nucleico particular que codifica la región variable de una inmunoglobulina específica" se refiere a una etapa de selección que puede realizarse opcionalmente como parte del método descrito y para identificar moléculas de ácido nucleico con afinidad particularmente fuerte por un antígeno dado. Para este fin, se expresan los diversos dominios variables codificados por la primera molécula de ácido nucleico. Después de la expresión, se ensaya la afinidad de los dominios variables con un antígeno específico. El ensayo puede realizarse por cualquier técnica adecuada conocida por el experto en la materia incluyendo ELISA, RIA, EIA, FACS, MS-MALDI, etc. En un ejemplo, la región variable se expresa en la superficie de un fago o virus. El contacto del fago o virus con un antígeno de interés permitirá aislar moléculas de ácido nucleico que codifiquen regiones variables con afinidad suficientemente alta por el antígeno. Dichos sistemas pueden modificarse por el experto en la materia para identificar solamente regiones variables con afinidad particularmente alta con el antígeno. Esta modificación puede implicar, por ejemplo, una modificación de las condiciones del tampón o la adición de competidores.

La expresión "fusionar la región codificante... en fase" se refiere a la generación de proteínas de fusión. Las proteínas de fusión son proteínas quiméricas que consisten en secuencias derivadas de al menos dos moléculas diferentes. De acuerdo con las enseñanzas de la presente invención, la región codificante que codifica la región variable de una inmunoglobulina se fusiona en fase con la región codificante que codifica uno o más (poli) péptidos detectables, tales como (poli) péptidos fluorescentes. La fusión puede realizarse por cualquier técnica conocida por el experto en la materia, siempre que dé como resultado la fusión en fase de dichas regiones codificantes. Convencionalmente, la generación de una proteína de fusión a partir de dos proteínas o dominios separados se basa en el "corte y empalme de dos caras por extensión solapante" descrito en Horton R., *et al.*, 1989, *Gene* 77:61-68. Los fragmentos que codifican los dominios o proteínas únicos se generan en dos reacciones de PCR primarias separadas. Los cebadores internos para las reacciones de PCR primarias contienen una región complementaria significativa, de aproximadamente 20 pb, que permite la fusión de los dos fragmentos de dominio en la segunda PCR. Como alternativa, las regiones codificantes pueden fusionarse haciendo uso de sitios de restricción que pueden ser de origen natural o introducirse por tecnología de ADN recombinante.

La expresión “la región codificante de la primera molécula de ácido nucleico se localiza 5' de la región codificante de la segunda molécula de ácido nucleico” se refiere a la disposición de regiones codificantes correspondientes en el nivel de ADN o ARNm e indica que los restos codificados por la primera secuencia de ácido nucleico son los restos N terminales de la proteína de fusión, mientras que los restos codificados por la segunda secuencia de ácido nucleico son los restos C terminales de la proteína de fusión.

La expresión “expresar la molécula de ácido nucleico fusionada que codifica la proteína de fusión en una célula o extracto sin células” se refiere a la transcripción y traducción de la proteína de fusión usando elementos de control de la expresión apropiados que actúen en la célula o extracto sin células seleccionado. De esta manera, las propiedades de proteínas de fusión individuales pueden ensayarse en sistemas de expresión celulares o en extractos sin células. Para este fin, la molécula de ácido nucleico que codifica la proteína de fusión puede clonarse en un vector de expresión adecuado, cuya composición, generalmente, depende del sistema de expresión. El sistema de expresión puede ser procariota o eucariota. Un vector de expresión de mamífero típico contiene el elemento promotor, que media en el inicio de la transcripción de ARNm, la secuencia codificante de proteína y señales requeridas para la terminación de la transcripción y poliadenilación del transcrito. Los elementos adicionales podrían incluir potenciadores, secuencias de Kozak y secuencias intermedias flanqueadas por sitios donadores y aceptores para corte y empalme de ARN. Puede conseguirse transcripción altamente eficaz con los promotores tempranos y tardíos de SV40, las repeticiones terminales largas (LTR) de retrovirus, por ejemplo, RSV, HTLV1, VIH1, y el promotor temprano del citomegalovirus (CMV). Sin embargo, también pueden usarse elementos celulares (por ejemplo, el promotor de actina humano). Los vectores de expresión adecuados para su uso en la práctica de la presente invención incluyen, por ejemplo, vectores tales como pSVL y pMSG (Farmacia, Uppsala, Suecia), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) y pBC12MI (ATCC 67109). Las células hospedadoras de mamífero que podrían usarse incluyen células Hela, 293, H9 y Jurkat humanas, células NIH3T3 y C127 de ratón, Cos 1, Cos 7 y CV1, células QC1-3 de codorniz, células L de ratón y células de ovario de hámster chino (CHO). Como alternativa, el (poli) péptido recombinante puede expresarse en líneas celulares estables que contienen la construcción génica integrada en un cromosoma. La co-transfección con un marcador seleccionable tal como dhfr, gpt, neomicina, higromicina permite la identificación y aislamiento de las células transfectadas. El ácido nucleico transfectado también puede ampliarse para expresar grandes cantidades del (poli) péptido codificado. El marcador de DHFR (dihidrofolato reductasa) es útil para desarrollar líneas celulares que porten varios cientos o incluso varios miles de copias del gen de interés. Otro marcador de selección útil es la enzima glutamina sintasa (GS) (Murphy *et al.* 1991, *Biochem J.* 227:277-279; Bebbington *et al.* 1992, *Bio/Technology* 10:169-175). Usando estos marcadores, las células de mamífero se cultivan en medio selectivo y se seleccionan las células con la mayor resistencia. Los vectores de expresión incluirán preferentemente al menos un marcador seleccionable. Tales marcadores incluyen dihidrofolato reductasa, G418 o resistencia a neomicina para cultivo de células eucariotas y genes de resistencia a tetraciclina, kanamicina o ampicilina para cultivar en *E. coli* y otras bacterias. Los ejemplos representativos de hospedadores apropiados incluyen, pero sin limitación, células bacterianas, tales como células de *E. coli*, *Streptomyces* y *Salmonella typhimurium*; células fúngicas, tales como células de levadura; células de insecto tales como células de *Drosophila* S2 y *Spodoptera* Sf9; células animales tales como CHO, COS, 293 y células del melanoma de Bowes; y células vegetales. Se conocen en la técnica medios y condiciones de cultivo apropiados para las células hospedadoras anteriormente descritas.

El polipéptido que se expresa de forma recombinante puede contener restos de aminoácidos adicionales para aumentar la estabilidad o para modificar la dirección de la proteína. Por ejemplo, puede añadirse una región de aminoácidos adicionales, particularmente aminoácidos cargados, al extremo N terminal del polipéptido para mejorar la estabilidad y persistencia en la célula hospedadora, durante la purificación, o durante la manipulación y el almacenamiento posteriores. Además, pueden añadirse restos peptídicos al polipéptido para facilitar la purificación. Tales regiones pueden retirarse antes de la preparación final del polipéptido. La adición de restos peptídicos a polipéptidos para generar secreción o excreción, para mejorar la estabilidad y para facilitar la purificación, entre otros, son técnicas familiares y rutinarias en este campo. Una proteína de fusión preferida comprende una región heteróloga de inmunoglobulina que es útil para estabilizar y purificar proteínas. Por ejemplo, el documento EP-A-0 464 533 describe proteínas de fusión que comprenden diversas partes de región constante de moléculas de inmunoglobulina junto con otra proteína humana o partes de la misma. En muchos casos, la parte Fc en una proteína de fusión es verdaderamente ventajosa para su uso en terapia y diagnóstico y por lo tanto da como resultado, por ejemplo, propiedades farmacocinéticas mejoradas (documento EP-A 0 232 262). Debe entenderse, sin embargo, que la parte Fc provechosamente no es parte de la proteína de fusión que puede obtenerse de acuerdo con la invención. Por otro lado, para algunos usos sería deseable poder suprimir la parte Fc después de que la proteína de fusión se haya expresado, detectado y purificado de la manera ventajosa descrita. Este es el caso cuando la parte Fc resulta ser un impedimento por ejemplo para la actividad de la proteína de fusión. En descubrimiento de fármacos, por ejemplo, se han fusionado proteínas humanas, tales como hIL-5, con partes Fc con el fin de realizar ensayos de exploración de alto rendimiento para identificar antagonistas de hIL-5. Véase, D. Bennett *et al.*, *J. Molecular Recognition* 8:52-58 (1995) y K. Johanson *et al.*, *J. Biol Chem* 270:9459-9471 (1995). La proteína de fusión de la presente invención puede recuperarse y purificarse de cultivos de células recombinantes por métodos bien conocidos incluyendo precipitación con sulfato de amonio o etanol, extracción ácida, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad y/o cromatografía de hidroxipatita. Más preferentemente, se emplea cromatografía líquida de proteína rápida (FPLC) para purificación, sin embargo, también puede usarse cromatografía líquida de alto

rendimiento ("HPLC").

De acuerdo con la presente invención se prefiere que la proteína de fusión se exprese a partir de un sistema de expresión viral o un sistema de expresión que implique un fago. Se han descrito varios de dichos sistemas de expresión en la técnica (véase por ejemplo Hoogenboom H.R. *et al.*, 1998, Antibody phage display and its applications, Immunotechnology 4:1-20; Pluckthun A, 1994, Escherichia coli producing recombinant antibodies, Bioprocess Technol. 19:233-252, Verma R. *et al.*, 1998, Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems, J. Immunol. Methods 216:165-181). Se prefieren particularmente sistemas de expresión que impliquen exposición en superficie de los restos de aminoácidos que forman la región variable, puesto que esto permite realizar selección basándose en la interacción de dicha región variable con un antígeno de interés.

Las expresiones "fluorescente", "cromóforo" o "péptido o proteína fosforescente" se refieren a (poli) péptidos con propiedades fluorescentes, cromóforas o fosforescentes. Puede usarse una diversidad de proteínas fluorescentes y cromoproteínas como segunda secuencia (poli) peptídica. Un grupo de dichas proteínas fluorescentes incluye proteína verde fluorescente aislada de *Aequorea victoria* (GFP), así como varias variantes de GFP, tales como proteína cian fluorescente, proteína azul fluorescente, proteína amarilla fluorescente, etc. (Zimmer, 2002, Chem Rev. 102:759-781; Zhang *et al.*, 2002, Nature Reviews 3:906-918). Típicamente, estas variantes comparten aproximadamente el 80%, o mayor identidad de secuencia con SEQ ID NOs: 6 o 7, respectivamente. Se han desarrollado varios mutantes de desplazamiento de color de GFP y pueden usarse para generar los fragmentos de reconocimiento de antígenos fluorescentes de la invención. Estos mutantes de GFP con desplazamiento de color tienen colores de emisión de azul a amarillo-verde, mayor brillo y fotoestabilidad (Tsien, 1998, Annual Review of Biochemistry 67: 509-544). Uno de dichos mutantes de GFP, denominado proteína amarilla fluorescente potenciada, presenta una emisión máxima a 529 nm. También se han descrito variantes basadas en GFP adicionales que tienen espectros de excitación y emisión modificados (Tsien *et al.*, Solicitud de Patente de Estados Unidos 20020123113A1), intensidad de fluorescencia potenciada y tolerancia térmica (Thastrup *et al.*, Solicitud de Patente de Estados Unidos 20020107362A1; Bjorn *et al.*, Solicitud de Patente de Estados Unidos 20020177189A1) y formación de cromóforos a niveles de oxígeno reducidos (Fisher, Patente de Estados Unidos N° 6.414.119). Más recientemente, se han descrito GFP de los antozoos *Renilla reniformis* y *Renilla kollikeri* (Ward *et al.*, Solicitud de Patente de Estados Unidos 20030013849).

Otro grupo de dichas proteínas fluorescentes incluye las proteínas fluorescentes aisladas de antozoos, incluyendo sin limitación la proteína roja fluorescente aislada de la especie de coral *Discosoma*, DsRed (Matz *et al.*, 1999, Nat. Biotechnol. 17:969-973), por ejemplo, SEQ ID NO: 8 o 9, respectivamente (véase, por ejemplo, número de referencia AF168419 versión AF16849.2). DsRed y las otras proteínas fluorescentes de antozoos comparten solamente aproximadamente el 26-30% de identidad de secuencia de aminoácidos con la GFP natural de *Aequorea victoria*, pero se conservan todos los motivos cruciales, lo que indica la formación de la estructura en barril beta de 11 cadenas característica de GFP. La estructura cristalina de DsRed también se ha resuelto, y muestra la conservación de la estructura en barril beta de 11 cadenas de GFP MMDB Id: 5742.

También se han descrito varios mutantes de la proteína roja fluorescente DsRed de longitud de onda más larga, y de forma similar, pueden emplearse en la generación de los fragmentos de reconocimiento de antígenos fluorescentes de la invención. Por ejemplo, pueden emplearse mutantes de DsRed descritos recientemente con espectros de emisión desplazados más hacia el rojo en la práctica de la invención (Wiehler *et al.*, 2001, FEBS Letters 487: 384-389; Terskikh *et al.*, 2000, Science 290:1585-1588; Baird *et al.*, 2000, Proc. Natl. Acad. Sci. USA 97: 11984-11989).

Se ha descrito recientemente un número crecientemente mayor de otras proteínas fluorescentes de varias formas de vida oceánicas, y el banco de datos de proteínas enumera en la actualidad varias estructuras cristalinas de GFP y mutantes de GFP, así como las estructuras cristalinas de diversos análogos de GFP. Se han descrito proteínas fluorescentes relacionadas con estructuras que se ha inferido que son similares a GFP, de corales, pennatuláceos, ascidias y anémonas de mar y pueden usarse en la generación de los fragmentos de reconocimiento de antígenos fluorescentes de la invención (para revisiones véase, Zimmer, 2002, Chem Rev. 102:759-781; Zhang *et al.*, 2002, Nature Reviews 3:906-918).

También se han presentado proteínas fluorescentes de *Anemonia majano*, *Zoanthus sp.*, *Discosoma striata*, *Discosoma sp.* y *Clavularia sp.* (Matz *et al.*, mencionado anteriormente). Se ha indicado que una proteína fluorescente clonada a partir de la especie de coral pétreo, *Trachyphyllia geoffroyi*, emite luz verde, amarilla y roja y convierte de emisión de luz verde a luz roja tras su exposición a luz ultravioleta (Ando *et al.*, 2002, Proc. Natl. Acad. Sci. USA 99:12651-12656). Las proteínas fluorescentes recientemente descritas de anémonas de mar incluyen proteínas verdes y naranja fluorescentes clonadas de *Anemonia sulcata* (Wiedenmann *et al.*, 2000, Proc. Natl. Acad. Sci. USA 97: 14091-14096), una proteína verde fluorescente potenciada de forma natural clonada a partir de los tentáculos de *Heteractis magnifica* (Hongbin *et al.*, 2003, Biochem. Biophys. Res Commun. 301:879-885), y una cromoproteína púrpura generalmente no fluorescente que presenta fluorescencia roja débil clonada de *Anemonia sulcata* y un mutante de la misma que presenta espectro de emisión de desplazamiento al rojo lejano (595 nm) (Lukyanov *et al.*, 2000, J. Biol. Chem. 275: 25879-25882).

Adicionalmente, se ha descrito otra clase de proteínas relacionadas con GFP que tienen propiedades cromóforas y

fluorescentes. Uno de dichos grupos de proteínas derivadas de coral, la pociloporinas, muestra una serie amplia de características espectrales y fluorescentes (Dove y Hoegh-Guldberg, 1999, Solicitud PCT WO 00146233; Dove *et al.*, 2001, Coral Reefs 19:197-204). Recientemente, se ha descrito la purificación y cristalización de la pociloporina Rtms5 del coral constructor de arrecifes *Montipora efflorescens* (Beddoe *et al.*, 2003, Acta Cryst. D59: 597-599). Rtms5 es de color azul oscuro, pero es débilmente fluorescente. Sin embargo, se ha indicado que Rtms5, así como otras cromoproteínas con homología de secuencia con Rtms5, puede interconvertirse en una proteína fluorescente de rojo lejano mediante sustituciones de un solo aminoácido (Beddoe *et al.*, 2003, mencionado anteriormente; Bulina *et al.*, 2002, BMC Biochem. 3: 7; Lukyanov *et al.*, 2000, mencionado anteriormente).

También se conocen otras diversas cromoproteínas derivadas de corales cercanamente relacionadas con las pociloporinas (véase, por ejemplo, Lukyanov *et al.* 2000, J. Biol Chem 275:25879-82; Gurskaya *et al.*, 2001, FEBS Letters 507:16-20). Puede usarse cualquiera de las proteínas fluorescentes o cromóforas o fragmentos fluorescentes o cromóforos de las mismas de acuerdo con las enseñanzas de la presente invención. Son ejemplos adicionales de proteínas fluorescentes GFP de *Renilla reniformis*, mKO de *Fungia concinna*, verde Azami de *Galaxeidae* o cOFP de *Cerianthus*. Los fragmentos de la proteína fluorescente o cromófora son preferentemente fragmentos funcionales.

Las proteínas fluorescentes, cromóforas o fosforescentes pertenecen al grupo denominado "proteínas detectables" y "(poli) péptidos detectables" y "(poli) péptidos marcadores detectables". En general, estas expresiones se refieren a (poli) péptidos que, tras su excitación, dan como resultado la emisión de energía detectable. Son ejemplos de proteínas detectables proteínas fluorescentes, proteínas cromóforas o "fosforescente", "fluorescente" y "fosforescente" se usan de forma intercambiable. Esto significa por ejemplo que, cuando una realización se refiere a "fluorescente", esta realización comprende "...fosforescente", "...fluorescente" y "...fosforescente".

La expresión "la región codificante de la primera molécula de ácido nucleico se localiza 5' de la región codificante de la segunda molécula de ácido nucleico" se refiere a la disposición de regiones codificantes correspondientes en el nivel de ADN o ARNm e indica que los restos codificados por la primera secuencia de ácido nucleico son los restos N-terminales de la proteína de fusión, mientras que los restos codificados por la segunda secuencia de ácido nucleico son los restos C terminales de la proteína de fusión.

La expresión "fragmento funcional" o "homólogo funcional", como se usa a lo largo de la memoria descriptiva, se refiere a fragmentos de las proteínas de fusión de la presente invención, en las que estos fragmentos conservan la actividad de unión a antígeno completa o al menos una parcial y/o conservan una actividad detectable, por ejemplo fluorescente o cromófora, de la proteína de fusión de la presente invención. La expresión "al menos una actividad de unión a antígeno parcial" significa una actividad que está reducida a no más de preferentemente el 10%, más preferentemente el 1% en comparación con la proteína de fusión de longitud completa. Una medida conveniente para determinar la actividad de unión a antígeno es la Kd para el par de antígeno/anticuerpo.

La expresión "enlazador de al menos un resto de aminoácido" se refiere a restos de aminoácidos que pueden disponerse entre los restos codificados por la primera y segunda secuencias de nucleótidos. Dicho enlazador puede en ocasiones ser útil, por ejemplo, para mejorar el plegamiento separado de los dominios individuales o para modular la estabilidad de la proteína de fusión. Además, dichos restos enlazadores pueden contener señales para transporte, secuencias de reconocimiento de proteasas o señales para modificación secundaria. Los restos de aminoácidos que forman el enlazador pueden estar estructurados o no estructurados. Preferentemente, el enlazador puede ser tan corto como un resto de aminoácido o hasta 2, 3, 4, 5, 10, 20 o 50 restos. En casos particulares, el enlazador puede implicar incluso hasta 100 o 150 restos.

Se describe el concepto de detectar antígenos intracelulares usando compuestos proteicos detectables que son capaces de unirse específicamente a antígenos. Para este fin, se han usado anticuerpos de cadena pesada (HCAb) desprovistos de cadenas ligeras del *Camelidae*⁸, (véase Figura 6). Los HCAb reconocen y se unen a su antígeno mediante regiones hipervariables (CDR, regiones determinantes de complementariedad) localizadas en el dominio variable de cadena pesada (denominado VHH), que presentan el fragmento de unión a antígeno intacto más pequeño (-15 kDa)⁹⁻¹¹. En comparación con otros fragmentos de anticuerpo pequeños derivados de otros mamíferos como Fab, Fv o scFv, los VHH tienen varias ventajas. En primer lugar, solo hay que clonar o sintetizar y expresar un dominio para generar un fragmento de unión a antígeno madurado *in vivo*. En segundo lugar, los VHH específicos pueden seleccionarse fácilmente clonando el VHH en vectores de presentación de fagos seguido de varios ciclos de selección con respecto a agentes de unión a antígeno. En tercer lugar, los VHH son altamente solubles y estables y pueden expresarse eficazmente en sistemas heterólogos⁹. Las afinidades halladas para VHH hasta la fecha estuvieron en el intervalo nanomolar y son comparables a las de scFv¹²⁻¹⁴. Para detección, la parte de unión a epítipo (VHH) se fusiona con un polipéptido fluorescente.

El pequeño tamaño del dominio de unión a antígeno permite la detección de antígenos de otro modo inaccesibles en una célula. Además, permite una mayor resolución de estructuras biológicas complejas. Mediante la visualización directa del antígeno en una superficie celular usando el cromocuerpo, puede tener lugar una detección estequiométrica para cuantificación lineal. Este es el primer método para detectar antígenos en células vivas. El método incluye cualquier estructura antigénica como, por ejemplo, proteínas o ácidos nucleicos pero también modificaciones post-traduccionales y modificaciones no proteicas tales como, por ejemplo, lípidos, azúcares y ácidos

nucleicos. Este método proporciona información acerca de la concentración y localización de antígenos intracelulares y puede por lo tanto aplicarse en investigación básica (detección, cuantificación e inactivación/enmascaramiento de antígenos celulares), diagnóstico, validación de diana (inhibición selectiva de componentes celulares y/o enmascaramiento de sitios de unión específicos), terapia de mamíferos, en particular mamíferos humanos o en el desarrollo de ensayos para exploración de fármacos de alto rendimiento. La lectura del método descrito en el presente documento, por ejemplo niveles de energía diferentes, permite las aplicaciones descritas anteriormente.

La aplicación de diagnóstico del método que se describe en el presente documento se basa en la comparación de la presencia detectada, cantidad o localización subcelular de un antígeno de interés en una célula de un individuo que padece una afección patológica con las de una célula de un individuo sano. Por ejemplo, las diferencias en la cantidad, presencia o localización subcelular de una proteína que tiene o se corresponde con el antígeno de interés, se puede correlacionar con progreso de la enfermedad a lo largo del tiempo. De forma similar, la eficacia de un medicamento puede examinarse siguiendo los cambios de la cantidad, presencia o localización de la estructura antigénica correlacionados con un estado patológico.

Para fines de investigación pueden detectarse alteraciones potenciales del resultado tras el tratamiento de células con ciertos agentes. El mismo principio subyace a la aplicación de la presente invención para validación de diana o exploración de fármacos.

En general, la fluorescencia total por célula depende del nivel de expresión de una proteína de fusión detectable de la invención en lugar de la concentración del antígeno. En ausencia del antígeno, la proteína de fusión/cromocuerpos se distribuirán por toda la célula. En presencia del antígeno, la distribución de cromocuerpo se desplazará hacia el compartimento subcelular respectivo del antígeno, por ejemplo, el citoesqueleto, lámina nuclear, sitios de replicación o cromatina. Dicho desplazamiento no solo es indicativo de la localización subcelular de la estructura antigénica, sino también de su presencia en general. Dependiendo de la intensidad de la señal, puede determinarse la cantidad de la estructura antigénica dentro de la célula. El método de análisis de imágenes depende del antígeno y puede ser un análisis de variación de intensidad media (derivación típica de la intensidad de fluorescencia) o cualquier otro método que identifique la estructura o compartimento de interés por selección basada en umbral o por marcadores fluorescentes específicos y determinando después la relación entre esta estructura/compartimento frente al resto de la célula.

Aplicando la enseñanza de la presente invención, en un ejemplo se desvela un fragmento de anticuerpo de unión a epítipo de 15 kDa inducido en *Lama pacos* frente a la proteína verde fluorescente (GFP). Como se muestra en el presente documento, la parte de unión a epítipo (VHH) puede fusionarse por ejemplo con la versión monomérica de la proteína roja fluorescente (mRFP1). Este "cromocuerpo" anti-GFP puede expresarse de forma estable en células de mamífero en las que reconoce su epítipo en diferentes compartimentos y estructuras celulares. Con microscopía de lapso de tiempo los inventores demuestran que el cromocuerpo localiza eficazmente cambios dinámicos de proteínas diana a lo largo del ciclo celular en células vivas. Como se usa en el presente documento, el término "cromocuerpo" se entiende como un dominio de reconocimiento de antígenos fusionado con una proteína fluorófora o cromogénica.

Los resultados descritos en el presente documento son sorprendentes, en particular porque un enfoque más o menos similar que implicaba proteínas de fusión basado en los mismos dominios funcionales (domino variable y GFP) no consiguió proporcionar ninguna herramienta útil (véase documento WO 03/091415A2 y Zeytun *et al.*, Nature Biotechnology 2003, Vol. 21, 12: 1473-1479) y tuvo que ser retirado por los autores (Zeytun *et al.*, Nature Biotechnology 2004, Vol. 22, 5: 601). En este trabajo, Zeytun *et al.* describieron la inserción de diversos bucles de unión a anticuerpo en los cuatro bucles expuestos en un extremo de la proteína verde fluorescente (GFP). En más detalle, el bucle de unión a anticuerpo derivado de la región de complementariedad 3 de la cadena pesada se insertó en uno de los cuatro bucles de GFP, que conectan las cadenas β de la estructura en lámina β . Las construcciones resultantes se llamaron fluorocuerpos. Se seleccionaron fluorocuerpos específicos que reconocían varios antígenos por presentación de fagos. La competencia por reconocimiento de antígenos de los diferentes fluorocuerpos se ensayó por diversos métodos incluyendo experimentos de desplazamiento de banda e inmunofluorescencia. Este enfoque habría sido una combinación interesante de un dominio de reconocimiento de antígenos y las propiedades fluorescentes de las proteínas verdes fluorescentes. Resulta interesante, sin embargo, que los propios autores retiraron su manuscrito porque los resultados inicialmente presentados no eran reproducibles. En la retractación (Zeytun *et al.*, 2004, Nature Biotechnology Vol. 22, 5: 601), los autores indicaron que las secuenciación de los ácidos nucleicos que codificaban las construcciones descritas revelaron unos llamados "mutantes fuera de fase". En otras palabras, ninguno de los fluorocuerpos inicialmente descritos por Zeytun *et al.* tiene la secuencia de aminoácidos o ácido nucleico codificante supuesta. Además, Zeytun *et al.*, declararon que el fenómeno biológico observado no puede explicarse. En este caso, la combinación del dominio de reconocimiento de antígenos y una proteína fluorescente no funcionó.

En el presente documento se describe una nueva técnica que permite dirigir y localizar antígenos en células vivas. La enseñanza de la presente invención se ilustra por los ejemplos, que implican cromocuerpos generados por inducción de anticuerpos de domino único contra GFP, laminina Dm0 de *Drosophila*, y citoqueratina-8 en *Lama*

pacos, *Camelus dromedarius* respectivo, y fusionando la parte de unión a epítipo de VHH con mRFP1. Como ejemplo se muestra claramente que el cromocuerpo anti-GFP puede expresarse de forma estable en células como un monómero activo y tiene acceso a todos los compartimentos y estructuras subcelulares ensayados. No se detectaron agregados, como se describe para un gran número de intracuerpos expresados de forma intracelular como scFVs²⁶. Esta funcionalidad intracelular del dominio VHH derivado de llama sugiere que su estabilidad intrínseca ($\Delta G = 30 \text{ kJ/mol}$)²⁷ incluso en ausencia del enlace disulfuro C22-C92 conservado es suficientemente alta para plegarse de forma apropiada. La ausencia de un par extra de cisteínas en las CDR en el dominio de unión a epítipo de HCAB derivados de llama podría ser por lo tanto una ventaja natural frente a dominios VHH de dromedarios, en los que es esencial un enlace disulfuro entre bucles de aparición frecuente para el reconocimiento de antígenos^{9, 28}. Otro ejemplo describe un cromocuerpo de GFP, en el que una de las cisteínas conservadas en la región VHH se ha mutado a serina. El ejemplo muestra claramente no solamente una mayor estabilidad del cromocuerpo resultante, sino sorprendentemente también una mejor accesibilidad de la molécula dentro de la célula.

Además la viabilidad de esta técnica se demuestra con cromocuerpos inducidos contra dos proteínas endógenas: citoqueratina-8 y laminina. Este nuevo enfoque que emplea cromocuerpos expande de forma drástica las posibilidades de, por ejemplo, la microscopía de células en vivo puesto que prácticamente cualquier componente o estructura potencialmente antigénico puede marcarse como objetivo y localizarse dentro de células vivas. El hecho de que puedan seleccionarse cromocuerpos específicos de bibliotecas recombinantes y su afinidad pueda mejorarse adicionalmente por ciclos de mutagénesis reducirá en el futuro, posiblemente eliminará, la necesidad de inmunización animal²⁹. No obstante, los VHH recuperados de "bibliotecas inmunes" se benefician de la maduración de afinidad y especificidad de antígeno inherente que se produce durante la inmunización. Las proteínas de fusión/cromocuerpos generados de acuerdo con la presente invención pueden usarse, por ejemplo, para aplicaciones de anticuerpos convencionales ya que la parte fluorescente, la proteína roja fluorescente (mRFP1) en este estudio, puede reemplazarse fácilmente por cualquier otra proteína fluorescente, enzima cromogénica o puede por ejemplo acoplarse con colorantes fluorescentes y puntos cuánticos. Son ejemplos de dichas enzimas cromogénicas peroxidasa de rábano rústico (HRP), fosfatasa alcalina o β -galactosidasa.

La proteína de fusión o la primera secuencia polipeptídica que comprende el dominio variable de anticuerpos de *Camelidae* como se describe en la presente invención también puede acoplarse con fluorocromos no proteicos. Los fluorocromos están disponibles como N-Hidroxisuccinimidiléster (NHS-éster) y pueden acoplarse con proteínas, por ejemplo, los dominios variables (VHH) de la presente invención, mediante grupos amino primarios de dicha proteína. Los fluorocromos habitualmente disponibles muestran máximos de extinción en diferentes regiones de longitud de onda. Son ejemplos colorantes de Cy (por ejemplo Cy3, Cy5), Dy549, Oyster556, colorantes de Alexa o Tamra. El acoplamiento y la aplicación exitosos de la proteína de fusión de la invención se ilustran en los ejemplos.

Como cualquier otra herramienta, los cromocuerpos también necesitarán controles cuidadosos. Como la fusión de proteínas con GFP también la unión de cromocuerpos con sus dianas puede afectar a su actividad y regulación especialmente evitando la interacción con otros componentes celulares. Sin embargo, la facilidad de seleccionar y desarrollar fragmentos de anticuerpo específicos producirá regularmente un conjunto de agentes de unión con diferentes dominios de las proteínas diana. Por lo tanto pueden seleccionarse agentes de unión que pueden estar mejor adaptados para localización o para interferencia funcional dependiendo de sus sitios de unión.

En resumen, se ha demostrado que la proteína de fusión aplicada en el método descrito y en particular el cromocuerpo presentado en el presente documento es adecuado para la dirección y localización de antígenos en todos los compartimentos y estructuras subcelulares ensayados. Incluso los antígenos de partes centrales de la maquinaria de replicación así como antígenos incluidos en profundidad en la cromatina pudieron localizarse a lo largo de la fase S y mitosis lo que demuestra la idoneidad de los cromocuerpos para el estudio de células en vivo.

En una realización preferida del método descrito, la etapa (b) comprende (a) exponer la célula a luz correspondiente a la longitud de onda de excitación de la proteína de fusión; (b) detectar la energía emitida desde la célula y/o detectar la distribución subcelular de la energía emitida; (c) comparar la energía detectada en la etapa (d) con: (i) la energía detectada en una célula de referencia que contiene una cantidad de referencia de dicha estructura antigénica de interés; o expresa una proteína de fusión de referencia para la que no se expresa compañero de unión en la célula; o (ii) un control de datos y (e) concluir, a partir de una energía diferente, el estado de salud de un individuo o evaluar la presencia de la estructura antigénica de interés; en la que una energía mayor detectada en la etapa (c) en comparación con la de la etapa (d) es indicativa de la presencia de dicho antígeno; y/o concluir, a partir de la cantidad y/o localización/distribución subcelular de la energía emitida, la cantidad o localización/distribución subcelular de la estructura antigénica de interés.

La presente memoria descriptiva describe un método en el que la proteína de fusión que comprende una primera secuencia (poli) peptídica que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* y una segunda secuencia (poli) peptídica derivable de una proteína fluorescente o cromófora, en la que dicha (a) primera secuencia (poli) peptídica está compuesta de marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, codificada por la secuencia de ácido nucleico de SEQ ID NO: 2 o codificada por una secuencia de ácido nucleico con al menos 70% de identidad de secuencia o fragmento de la misma: y (b) segunda secuencia (poli) peptídica es (i) la proteína verde fluorescente derivable de *Aequorea victoria* codificada por la secuencia de ácido nucleico de SEQ ID NO: 7, o un

mutante fluorescente o fragmento del mismo; (ii) la proteína roja fluorescente derivable de *Discosoma* (DsRed) codificada por la secuencia de ácido nucleico de SEQ ID NO: 9, o un mutante fluorescente o fragmento del mismo; o (iii) un homólogo funcional de (i) o (ii) con al menos 80% de identidad de secuencia, más preferentemente al menos 85%, aún más preferentemente al menos 90%, aún más preferentemente al menos 95% y más preferentemente con al menos 98% de identidad de secuencia, en la que dicha primera secuencia (poli) peptídica se localiza en dirección N terminal de dicha segunda secuencia (poli) peptídica, separándose opcionalmente dichas secuencias por un enlazador o al menos un resto de aminoácido. El término "fragmento" se refiere a fragmentos funcionales capaces de unirse específicamente al antígeno.

En una realización preferida dicha región variable comprende marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, y está codificada por la secuencia de ácido nucleico de SEQ ID NO: 2 o está codificada por una secuencia de ácido nucleico con al menos 70%, más preferentemente 80%, más preferentemente al menos 85%, aún más preferentemente al menos 90%, aún más preferentemente al menos 95% y más preferentemente al menos 98% de identidad de secuencia o un fragmento de la misma. Preferentemente dicho fragmento es un fragmento funcional, es decir un fragmento capaz de unión de antígeno específica. La "unión específica" puede describirse, por ejemplo, con respecto a reactividad cruzada. Preferentemente, los agentes de unión específicos son proteínas (proteínas de fusión o fragmentos de las mismas) que tienen una constante de disociación o K_D de menos de 10^{-9} M. En una realización preferida de la presente invención, la secuencia de (a) CDR1 consiste en los restos mostrados en SEQ ID NO: 3; (b) CDR2 consiste en los restos mostrados en SEQ ID NO: 4; y (c) CDR3 consiste en los restos mostrados en SEQ ID NO: 5.

En una realización más preferida de la presente invención, dicha primera secuencia (poli) peptídica de dicha proteína de fusión o el (poli) péptido que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* tiene la secuencia de una cualquiera de SEQ ID NOs: 10 o 30 o está codificada por la secuencia de una cualquiera de SEQ ID NOs: 11 o 31.

En una realización más preferida de la presente invención, dicha segunda secuencia (poli) peptídica comprende los restos 1 a 239 de SEQ ID NOs: 6 o 1 a 226 de SEQ ID NOs: 8 o un mutante fluorescente o fragmento del mismo.

En otra realización preferida de la presente invención, dicho mutante de la proteína roja fluorescente es mRFP1 como se codifica por SEQ ID NO: 17 o cualquiera de los mutantes de DsRed fluorescentes codificados por una cualquiera de SEQ ID NOs: 17 y 24 a 27 o por un fragmento de las mismas. mRFP es la proteína roja fluorescente monomérica derivada de DsRed y con frecuencia también se denomina mRFP1 (Campbell *et al.*, (2002) PNAS, 99: 7877-7882, secuencia codificante: número de referencia: AF506027.1 (SEQ ID NO: 17)). Se han generado entre tanto varios otros derivados monoméricos de DsRed, incluyendo mRFPmars (secuencia codificante: número de referencia: AY679163.1 (SEQ ID NO: 24)), publicado por Fischer *et al.*, (2004) FEBS Letters, 577: 227 - 232, mCherry (secuencia codificante: número de referencia: AY678264 (SEQ ID NO: 25)), publicado por Shaner, N.C. *et al.*, (2004) Nat. Biotech., 22: 1567 - 1572), mRaspberry (secuencia codificante: número de referencia: AY86536 (SEQ ID NO: 26)), publicado por Wang, L. *et al.*, (2004) PNAS, 22: 1567 - 1572, mPlum (secuencia codificante: número de referencia: AY86537.1 (SEQ ID NO: 27)), publicado por Wang, L. *et al.*, (2004) PNAS, 22: 1567 - 1572. Cualquiera de estos mutantes fluorescentes derivados de DsRed puede usarse de acuerdo con las enseñanzas de la presente invención. En particular, las proteínas de fusión generadas de acuerdo con las enseñanzas de la presente invención pueden comprender una segunda secuencia (poli) peptídica fluorescente que comprende cualquiera de los mutantes de DsRed anteriormente mencionados o mutantes fluorescentes o fragmentos de los mismos.

En otra realización preferida de la presente invención, la proteína de fusión de la presente invención comprende además una secuencia de dirección seleccionada del grupo que consiste en señal de localización nuclear (NLS), secuencia de importación al retículo endoplásmico, secuencia de importación a la mitocondria. Un ejemplo de una NLS es la secuencia peptídica PKKKRKV (señal de localización nuclear (NLS) del antígeno T grande de SV40, D. Kalderon *et al.*, 1984, Cell 39: 499) que es capaz de dirigir proteínas heterólogas al núcleo. Un ejemplo de la "Secuencia de Importación al Retículo Endoplásmico" es el péptido MMSFVSLLLVILFWATEAEQLTRCVFQ (señal de localización de RE de cadena ligera de inmunoglobulina, Blobel G. y Dobberstein B., 1975 J. Cell. Biol. 67: 835 - 851) que es capaz de dirigir las proteínas heterólogas al lumen del RE. Un ejemplo de una "Secuencia de Importación a Mitocondrias" es el péptido MLSLRQSIRFFRPATRTLCSRYLL (Neupert W. 1997 Annu. Rev. Biochem. 66: 863-917) que es capaz de dirigir proteínas heterólogas a la mitocondria.

También se describe una proteína de fusión que tiene la secuencia de una cualquiera de SEQ ID NOs: 18, 32, 20 o 22 o que está codificada por una molécula de ácido nucleico que comprende la secuencia de una cualquiera de SEQ ID NOs: 19, 33, 21 o 23.

En una realización preferida, la célula es una célula obtenida de un individuo.

En otra realización preferida dicha estructura antigénica de interés se detecta dentro de una célula. En una realización más preferida dicha célula es una célula viva. En otra realización preferida, dicha célula es una célula dentro de un organismo eucariota vivo, preferentemente dentro de un mamífero.

En una realización preferida dicha estructura antigénica se selecciona de proteína, modificación de proteína, cofactor, compuesto molecular pequeño, ADN y ARN.

5 En una realización preferida el (poli) péptido capaz de transducir se selecciona del grupo de (poli) péptidos básicos que comprenden péptido TAT, poli-arginina y poli-lisina.

En una realización más preferida, la proteína es una proteína citoplásmica, nuclear o nucleolar.

10 El término "fragmento" se refiere a mutantes de delección de la región variable como se han definido anteriormente que conservan la especificidad de unión (véase anteriormente). Tales mutantes de delección pueden comprender una o más delecciones, que implican la delección de por ejemplo 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 y hasta 15 o incluso 20 restos de aminoácidos. Las delecciones pueden residir en posición C terminal Nor o en forma interna. Preferentemente, la región variable comprende tres CDR intactas. Sin embargo, es concebible que en casos particulares aún sea posible suficiente unión a antígeno con solamente dos o incluso una CDR. En tales casos se prefiere que una de las CDR
15 restantes sea CDR3.

La identidad de secuencia puede determinarse usando el programa Bestfit® (Wisconsin Sequence Analysis Package, Versión 8 para Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). Bestfit® usa el algoritmo de homología local de Smith y Waterman para encontrar el mejor segmento de homología entre dos secuencias (Advances in Applied Mathematics 2: 482-489 (1981)). Cuando se usa Bestfit® o cualquier otro programa de alineamiento de secuencias para determinar si una secuencia en particular es, por ejemplo, 95% idéntica a una secuencia de referencia, los parámetros se establecen, por supuesto, de modo que el porcentaje de identidad se calcule sobre la longitud completa de la secuencia de nucleótidos de referencia y que se permitan huecos de homología de hasta el 5% del número total de nucleótidos en la secuencia de referencia. La identidad entre una primera secuencia y una segunda secuencia, también denominada alineamiento de secuencia global, se determina usando el programa informático FASTDB basado en el algoritmo de Brutlag y colaboradores (Comp. App. Biosci. 6: 237-245 (1990)). En un alineamiento de secuencia las secuencias de consulta y objeto son ambas secuencias de ADN. Una secuencia de ARN puede compararse convirtiendo U a T. El resultado de dicho alineamiento de secuencia global está en porcentaje de identidad. Son parámetros preferidos usados en un alineamiento FASTDB de secuencias de ADN para calcular el porcentaje de identidad: Matriz=Unitaria, k-tuple=4, Penalización de Emparejamiento Erróneo=1, Penalización de Unión=30, longitud de Grupo de Selección Aleatoria=0, Puntuación de Corte=1, Penalización de Hueco=5, Penalización de Tamaño de Hueco 0,05, Tamaño de Ventana=500 o la longitud de la secuencia de nucleótidos objeto, la que sea más corta.

35 El fragmento de proteína que comprende la secuencia VHH aplicada en la presente invención puede producirse con una célula productora de anticuerpos, que puede ser un linfocito de sangre periférica (PBL). Pueden obtenerse linfocitos de sangre periférica de sangre de un camello inmunizado. La técnica más habitual para separar leucocitos es mezclar sangre con un compuesto que agrega los eritrocitos, aumentando de este modo su tasa de sedimentación. La sedimentación de leucocitos se ve afectada solo ligeramente y puede recogerse de la parte superior del tubo cuando los eritrocitos se han sedimentado (Thorsby E. & Bratlie A., 1970, Histocompatibility Testing 1970, ed. P.I. Terasaki, p.655 Munksgaard, Copenhagen). Se prefiere un procedimiento convencional de acuerdo con un protocolo proporcionado, por ejemplo, por AXIS-SHIELD PoC AS (Oslo, Noruega) usando Lymphoprep™.

45 Los PBL pueden exponerse a un antígeno para el que es específica la región variable de la inmunoglobulina. Dicha exposición puede realizarse *in vitro* (es decir, por ejemplo en el cultivo celular) o *in vivo* (en un animal). Existen diversos protocolos en la técnica, para exponer PBL a antígenos (véase por ejemplo Harlow, E. y Lane, D. 1988 Antibodies, A Laboratory Manual, capítulo 5, Cold Spring Harbor, Nueva York).

50 Además, la proteína de fusión aplicada en la presente invención puede alterarse. El método de la presente invención comprende además la etapa adicional de modificar la primera molécula de ácido nucleico que codifica la región variable dentro de CDR1, CDR2 y/o CDR3. Esto puede hacerse con la intención de conservar la especificidad pero mejorar la afinidad, por ejemplo. Como alternativa, puede cambiarse la especificidad. El término "modificar" significa mutagenizar o cambiar la secuencia de nucleótidos que codifica la región variable. Las primeras secuencias mutantes pueden contener una o más delecciones, sustituciones o adiciones como se ha indicado anteriormente.

60 CDR1, CDR2 y/o CDR3 pueden modificarse por mutagénesis aleatoria de la primera molécula de ácido nucleico. La complejidad de las bibliotecas de fagos se ha limitado con frecuencia a 10^8 por la baja eficacia de la transformación de ADN en bacterias. Por lo tanto una biblioteca comprende solamente un número muy pequeño de todas las posibles secuencias. Esta es una de las principales razones por la que los ligandos seleccionados de bibliotecas de fragmentos de anticuerpos aleatorias son con frecuencia de baja afinidad. Un modo eficaz de superar esta limitación e identificar ligandos con propiedades mejoradas es a través de un proceso de "maduración de epítomos". Las secuencias de CDR que reconocen antígenos seleccionadas de la biblioteca inicial se mutan parcialmente para generar una población de variantes. La exploración de esta biblioteca presentando un antígeno permite identificar ligandos que son mejores que la secuencia candidata original. Esto puede realizarse por ejemplo por mutagénesis dirigida que intercambia nucleótidos únicos en las regiones CDR (referencia Clackson T. y Lowman H.B., 2004,

Phage Display: A practical Approach, Oxford University Press Inc., Nueva York). Otras técnicas son sintetizar las regiones CDR para intercambiar nucleótidos únicos que conducen a sustitución de un único aminoácido. Las CDR sintetizadas artificiales se insertarían en el armazón del fragmento de anticuerpo en las posiciones particulares por ligación usando restricción clásica o "corte y empalme de dos caras por extensión solapante" descrito en Horton R. *et al.*, 1989, *Gene* 77:61-68. Además, se prefiere usar mutagénesis aleatoria. Esto puede realizarse mediante mutagénesis dirigida a oligonucleótido descrita por Zoller, M.J. 1992, *Curr. Opin. in Biotechnology* 3:348-354. Se describe un desarrollo adicional usando trinucleótido fosforamiditas en una síntesis de ADN de fase sólida en Virnekas T. *et al.*, 1994, *Nucl. Acids Res.* 22, 5600-5607.

La primera molécula de ácido nucleico que codifica la región variable se selecciona poniendo en contacto la región variable con un compañero de unión (un antígeno). Se puede realizar selección, y en particular detección del dominio variable basándose en la proteína de fusión que comprende el (poli) péptido fluorescente o cromóforo. Como alternativa, también puede realizarse selección de regiones variables particulares en ausencia del (poli) péptido fluorescente o cromóforo. Por ejemplo, antes o después de fusión, puede ensayarse un grupo de regiones variables con respecto a su especificidad o afinidad por un antígeno particular. Los ensayos adecuados incluyen FACS, ELISA, RIA, EIA; MS-MALDI etc.

En muchos casos la región variable se expresará en una biblioteca de expresión como se describe en el presente documento posteriormente en detalle. Se conocen bien por los expertos en la materia métodos para explorar bibliotecas. Las bibliotecas se exploran típicamente usando un antígeno, o molécula de interés, para el que es deseable seleccionar un compañero de unión. Típicamente, el antígeno se une a una superficie sólida o un marcador específico, tal como biotina. El antígeno (o molécula de interés) se incuba con una biblioteca de la invención. Los polipéptidos o miembros de biblioteca que se unen al antígeno se separan después de los que no, usando cualquiera de varios métodos diferentes. Estos métodos implican etapas de lavado, seguidas de etapas de elución. El lavado puede realizarse, por ejemplo, con PBS, o tampones que contienen detergente. La elución puede realizarse con varios agentes, dependiendo del tipo de biblioteca. Por ejemplo, puede usarse un ácido, una base, bacterias o una proteasa cuando la biblioteca es una biblioteca de presentación de fagos.

Para facilitar la identificación y aislamiento del fragmento de reconocimiento de antígenos fluorescente o unido a antígeno, el fragmento de reconocimiento de antígenos fluorescente también puede modificarse por ingeniería genética como una proteína de fusión para incluir marcadores de selección (por ejemplo, marcadores epitópicos). Pueden usarse después anticuerpos reactivos con los marcadores de selección presentes en las proteínas o restos de fusión que se unen a los marcadores para aislar el complejo de antígeno-fragmento de reconocimiento de antígeno fluorescente mediante el epítipo o marcador.

Otros dominios facilitadores de purificación y detección incluyen, por ejemplo, péptidos quelantes metálicos tales como tramos de polihistidina y módulos de histidina-triptófano que permiten la purificación en metales inmovilizados, dominios de proteína A que permiten purificación en inmunoglobulina inmovilizada, o el dominio utilizado en el sistema de purificación de extensión/afinidad FLAGS (Immunex Corp, Seattle WA). Puede usarse cualquier epítipo con un anticuerpo de alta afinidad correspondiente, por ejemplo, un marcador myc (véase, por ejemplo, Kieck, 1997, *Protein Eng.* 10:1303-1310) o un marcador His (Pharmacia). Véase también Maier, 1998, *Anal. Biochem.* 259:68-73; Muller, 1998, *Anal. Biochem.* 259:54-61.

La proteína de fusión resultante puede seleccionarse en una biblioteca de expresión. La expresión puede realizarse como se ha indicado anteriormente. El sistema de expresión puede implicar la presentación del dominio variable. La biblioteca de expresión puede ser, por ejemplo, una biblioteca de expresión de levadura, bacteriana o de mamífero.

Además de bibliotecas de presentación de fagos, que se describen en detalle posteriormente, también pueden usarse bibliotecas de presentación de epítipos análogas. Por ejemplo, los métodos de la invención también pueden usar bibliotecas de presentación en superficie de levadura (véase, por ejemplo, Boder, 1997, *Nat. Biotechnol.*, 15553-15557), que pueden construirse usando tales vectores como el vector de expresión de levadura pYDI. Otros sistemas de presentación potenciales incluyen vectores de presentación de mamíferos y bibliotecas de *E. coli*.

También pueden usarse formatos de biblioteca de presentación *in vitro* conocidos para los expertos en la materia, por ejemplo, bibliotecas de presentación de ribosomas y bibliotecas de presentación de ARNm. En estas tecnologías de selección *in vitro*, se preparan proteínas usando traducción sin células y unida físicamente a su ARNm codificante después de traducción *in vitro*. En la metodología típica para generar estas bibliotecas, se transcribe ADN que codifica las secuencias para seleccionar *in vitro* y se traduce en un sistema sin células.

En bibliotecas de presentación de ribosomas (véase, por ejemplo, Mattheakis *et al.*, 1994, *Proc. Natl. Acad. Sci USA* 91:9022-9026; Hanes y Pluckthrun, 1997, *Proc. Natl. Acad. Sci USA*, 94:4937-4942) el enlace entre el ARNm que codifica el fragmento de reconocimiento de antígenos fluorescente de la invención y el ligando es el propio ribosoma. La construcción de ADN se diseña de modo que no se incluya ningún codón de parada en el ARNm transcrito. Por lo tanto, el ribosoma de traducción se detiene al final del ARNm y la proteína codificada no se libera. La proteína codificada puede plegarse en su estructura correcta mientras aún está unida al ribosoma. El complejo de ARNm, ribosoma y proteína se usa después directamente para selección contra una diana inmovilizada. El ARNm de

complejos ribosómicos unidos se recupera por disociación de los complejos con EDTA y se amplifica por RT-PCR.

Los métodos y bibliotecas basados en tecnología de presentación de ARNm, también denominados en el presente documento presentación de puromicina, se describen, por ejemplo en las Patentes de Estados Unidos N° 6.261.804; 6.281.223; 6.207.446 y 6.214.553. En esta tecnología, se fusiona en primer lugar un enlazador de ADN unido a puromicina con el extremo 3' de ARNm. La proteína se traduce después *in vitro* y el ribosoma se detiene en el punto de unión de ARN-ADN. La puromicina, que se asemeja al aminoacil ARNt, entra en el sitio ribosómico A y acepta el polipéptido naciente. La proteína traducida se une por tanto covalentemente con su ARNm codificante. Las moléculas fusionadas pueden después purificarse y explorarse con respecto a actividad de unión. Las secuencias de ácido nucleico que codifican ligandos con actividad de unión pueden obtenerse después, por ejemplo, usando RT-PCR.

Los fragmentos y secuencias de reconocimiento de antígenos fluorescentes, por ejemplo, enlazador de ADN para conjugación con puromicina, pueden unirse por métodos bien conocidos por los expertos en la materia y se describen, por ejemplo, en las Patentes de Estados Unidos N° 6.261.804; 6.281.223; 6.207.446 y 6.214.553.

Otras tecnologías implican el uso de proteínas virales (por ejemplo, proteína A) que se unen covalentemente en sí mismas a los genes que las codifican. Se crean proteínas de fusión que se unen al fragmento de reconocimiento de antígeno fluorescente con la secuencia de proteína A, proporcionando de este modo un mecanismo para unir los ligandos de unión a los genes que los codifican.

Los sistemas de presentación de plásmidos se basan en la fusión de proteínas presentadas a proteínas de unión a ADN, tales como el represor lac (véase, por ejemplo, Gates *et al.*, 1996, J. Biol., 255:373-386; 1996, Methods Enzymol. 267:171-191). Cuando el operador lac también está presente en el plásmido, la proteína de unión a ADN se une a él y puede purificarse conjuntamente con el plásmido. Pueden crearse bibliotecas unidas a la proteína de unión a ADN, y explorarse tras lisis de las bacterias. Los plásmidos/proteínas deseados se rescatan por transfección o amplificación.

También puede realizarse selección en una biblioteca de presentación de fagos.

La construcción de bibliotecas de presentación de fagos aprovecha la capacidad del bacteriófago para presentar péptidos y proteínas en sus superficies, es decir, en sus cápsidas. Con frecuencia, se usan fagos filamentosos tales como M13, fd o f1. El fago filamentosos contiene ADN monocatenario rodeado de múltiples copias de genes que codifican proteínas de la cubierta mayores y menores, por ejemplo, pIII. Las proteínas de cubierta se presentan en la superficie externa de la cápsida. Se transcriben conjuntamente secuencias de ADN insertadas en fase con genes de proteínas de la cápsida para generar proteínas de fusión o fragmentos de proteínas presentados en la superficie del fago. Las bibliotecas de fagos pueden presentar por lo tanto péptidos representativos de la diversidad de las secuencias insertadas. Significativamente, estos péptidos pueden presentarse en conformaciones plegadas "de forma natural". Los fragmentos de reconocimiento de antígenos fluorescentes expresados en bibliotecas de presentación de fagos pueden unirse después a moléculas diana, es decir, pueden interaccionar específicamente con moléculas compañeras de unión tales como antígenos, por ejemplo, (Petersen, 1995, Mol. Gen. Genet. 249: 425-31), receptores de superficie celular (Kay, 1993, Gene 12859-65), y proteínas extracelulares e intracelulares (Gram, 1993, J. Immunol. Methods, 161: 169-76).

El concepto de usar fagos filamentosos, tales como M13 o fd, para presentar péptidos en superficies de cápsida de fagos se introdujo por primera vez por Smith, 1985, Science 228:1315-1317. Se han presentado péptidos en superficies de fagos para identificar muchos ligandos potenciales (véase, por ejemplo, Cwirla, 1990, Proc. Natl. Acad. Sci. USA, 87:6378-6382). Hay numerosos sistemas y métodos para generar bibliotecas de presentación de fagos descritos en la bibliografía científica y de patentes, véase, por ejemplo, Sambrook y Russell, Molecule Cloning: A Laboratory Manual, 3ª edición, Cold Spring Harbor Laboratory Press, capítulo 18, 2001; Phage, Display of Peptides and Proteins: A Laboratory Manual, Academic Press, San Diego, 1996; Cramer, 1994, Eur J. Biochem. 226:53-58; de Kruif, 1995, Proc. Natl. Acad. Sci. USA, 92:3938-3942; McGregor, 1996, Mol Biotechnol, 6:155-162; Jacobsson, 1996, Biotechniques, 20: 1070-1076; Jespers, 1996, Gene, 173:179-181; Jacobsson, 1997, Microbiol Res., 152:121-128; Fack, 1997, J. Immunol. Methods, 206:43-52; Rossenu, 1997, J. Protein Chem., 16:499-503; Katz, 1997, Annu. Rev. Biophys. Biomol. Struct., 26: 27-45; Rader, 1997, Curr. Opin. Biotechnol., 8:503-508; Griffiths, 1998, Curr. Opin. Biotechnol., 9:102-108.

Típicamente, se insertan ácidos nucleicos exógenos que codifican las secuencias proteicas para presentar en un gen de proteína de cubierta, por ejemplo gen III o gen VIII del fago. Las proteínas de fusión resultantes se presentan en la superficie de la cápsida. La proteína VIII está presente en aproximadamente 2700 copias por fago, en comparación con de 3 a 5 copias para la proteína III (Jacobsson (1996), mencionado anteriormente). Pueden usarse vectores de expresión multivalentes, tales como fagémidos, para manipulación de las secuencias de ácido nucleico que codifican la biblioteca de fragmentos de reconocimiento de antígenos fluorescentes y producción de partículas de fagos en bacterias (véase, por ejemplo, Felici, 1991, J. Mol. Biol., 222:301-310).

Con frecuencia se emplean vectores fagémidos para construir la biblioteca de fagos. Estos vectores incluyen el origen de replicación de ADN del genoma de un bacteriófago filamentosos monocatenario, por ejemplo, M13 o f1 y

requieren la aportación de las otras proteínas de fagos para crear un fago. Este se proporciona habitualmente por un fago auxiliar que es menos eficaz en empaquetamiento en partículas de fago. Puede usarse un fagémido del mismo modo que un vector plasmídico ortodoxo, pero también pueden usarse para producir partículas de bacteriófagos filamentosas que contengan copias monocatenarias de segmentos clonados de ADN.

No es necesario que la proteína o el fragmento proteico presentado sea una proteína de fusión entre el (poli) péptido que comprende la región variable y la proteína del sistema de presentación. Por ejemplo, el (poli) péptido que comprende la región variable (es decir, el fragmento de reconocimiento del antígenos fluorescente) puede unirse a una proteína de cubierta debido a una interacción no covalente, por ejemplo, una interacción de unión de superenrollamiento, tal como una unión jun/fos, o una interacción covalente mediada por cisteínas (véase, por ejemplo, Cramer *et al.*, 1994, Eur. J. Biochem., 226:53-58) con o sin interacciones no covalentes adicionales. Morphosys ha descrito un sistema de presentación en el que se pone una cisteína en el extremo C terminal del scFv o Fab, y se pone otra en el extremo N terminal de g3p. Los dos se ensamblan en el periplasma y se produce presentación sin un gen o proteína de fusión.

No es necesario que la proteína de cubierta sea endógena. Por ejemplo, pueden incorporarse proteínas de unión a ADN en el genoma de fago/fagémido (véase, por ejemplo, McGregor y Robins, 2001, Anal. Biochem., 294:108-117.). Cuando la secuencia reconocida por dichas proteínas también está presente en el genoma, la proteína de unión a ADN se incorpora en el fago/fagémido. Esta puede actuar como una proteína de vector de presentación. En algunos casos se ha mostrado que la incorporación de proteínas de unión de ADN en la cubierta del fago puede producirse independientemente de la presencia de la señal de ADN reconocida.

También puede usarse otro fago. Por ejemplo, pueden emplearse vectores T7, vector T4, vectores T2 o vectores lambda en los que el producto presentado en la partícula de fago madura se libera por lisis celular.

Otra metodología es la tecnología de fago infeccioso selectivo (SIP), que posibilita la selección *in vivo* de pares de proteína-ligando de interacción. Un "fago infeccioso selectivo" consiste en dos componentes independientes. Por ejemplo, una partícula de fago filamentoso recombinante se hace no infecciosa reemplazando sus dominios N terminales de proteína del gen 3 (g3p) con una proteína de interés, por ejemplo, un antígeno. El ácido nucleico que codifica el antígeno puede insertarse de modo que se exprese. El segundo componente es una molécula "adaptadora" en la que el fragmento de reconocimiento de antígeno fluorescente se une a los dominios N terminales de g3p que están ausentes en la partícula de fago. Se restaura la infecciosidad cuando la proteína presentada (por ejemplo, un fragmento de reconocimiento de antígenos fluorescente) se une al antígeno. Esta interacción une los dominios N terminales ausentes de g3p con la partícula de presentación de fagos. La propagación de fagos se hace estrictamente dependiente de la interacción proteína-ligando. Véase, por ejemplo, Spada, 1997, J. Biol. Chem. 378:445-456; Pedrazzi, 1997, FEBS Lett. 41 5:289-293; Hennecke, 1998, Protein Eng. 11:405-410.

La presente invención también se refiere a una proteína de fusión que comprende una primera secuencia (poli) peptídica que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* (incluyendo cualquier camello o dromedario) y una segunda secuencia (poli) peptídica, que es una proteína detectable, preferentemente derivable de una proteína detectable, por ejemplo fluorescente, cromófora o fosforescente, en la que dicha (a) primera secuencia (poli) peptídica está codificada por la secuencia de ácido nucleico de SEQ ID NOs: 1 o 31 o codificada por una secuencia de ácido nucleico con al menos 95% de identidad de secuencia con SEQ ID NO: 11 y (b) segunda secuencia (poli) peptídica, si deriva de una proteína fluorescente o cromófora, pueden ser (i) la proteína verde fluorescente derivable de *Aequorea victoria* codificada por la secuencia de ácido nucleico de SEQ ID NO: 7, o un mutante fluorescente o fragmento del mismo; (ii) la proteína roja fluorescente derivable de *Discosoma* (DsRed) codificada por la secuencia de ácido nucleico de SEQ ID NO: 9, o un mutante fluorescente o fragmento del mismo; o (iii) un homólogo funcional de (i) o (ii) con al menos 80% de identidad de secuencia, más preferentemente al menos 85%, aún más preferentemente al menos 90%, aún más preferentemente al menos 95% y más preferentemente con al menos 98% de identidad de secuencia; en el que dicha primera secuencia (poli) peptídica está localizada en dirección N terminal de dicha segunda secuencia (poli) peptídica, estando separadas opcionalmente dichas secuencias por un enlazador de al menos un resto de aminoácido. El término "fragmento" se refiere a fragmentos funcionales capaces de unirse específicamente con el antígeno.

Las proteínas de fusión son proteínas quiméricas que consisten en secuencias derivadas de al menos dos moléculas diferentes. De acuerdo con las enseñanzas de la presente invención, la región codificante que codifica la región variable de una inmunoglobulina está fusionada en fase con la región codificante que codifica uno o más (poli) péptidos detectables, tales como (poli) péptidos fluorescentes. Puede realizarse fusión por cualquier técnica conocida por el experto en la materia, siempre que dé como resultado la fusión en fase de dichas regiones codificantes. Convencionalmente, la generación de una proteína de fusión de dos proteínas o dominios separados se basa en el "corte y empalme de dos caras por extensión solapante" descrito en Horton R., *et al.*, 1989, Gene 77:61-68. Los fragmentos que codifican los dominios únicos o proteínas se generan en dos reacciones de PCR primarias separadas. Los cebadores internos para las reacciones de PCR primarias contienen una región complementaria significativa, de aproximadamente 20 pb, que permite la fusión de los dos fragmentos de dominio en la segunda PCR. Como alternativa, las regiones codificantes pueden fusionarse haciendo uso de sitios de restricción que puede ser de origen natural o introducirse por tecnología de ADN recombinante.

En una realización preferida de la presente invención la secuencia del segundo (poli) péptido de la proteína de fusión comprende los restos 1 a 239 de SEQ ID NO: 6 o 1 a 226 de SEQ ID NO: 8 o un mutante fluorescente o fragmento del mismo.

5 En una realización más preferida de la presente invención el mutante de la proteína roja fluorescente es mRFP1 como se muestra en SEQ ID NO: 17 o una proteína o (poli) péptido codificado por la secuencia de una cualquiera de SEQ ID NOs: 17 y 24 a 27.

10 En una realización aún más preferida de la presente invención, la proteína de fusión tiene la secuencia de una cualquiera de SEQ ID NOs: 18, 32, 20 o 22 o está codificada por un ácido nucleico idéntico a o que comprende la secuencia de una cualquiera de SEQ ID NOs: 19, 33, 21 o 23.

15 El cromocuerpo anti-GFP puede expresarse de forma estable en células como un monómero activo y tiene acceso a todos los compartimentos y estructuras subcelulares ensayados. No se detectaron agregados, como se describe para un gran número de intracuerpos expresados de forma intracelular como scFvs²⁶. Esta funcionalidad intracelular del dominio VHH derivado de llama sugiere que su estabilidad intrínseca ($\Delta G = 30 \text{ kJ/mol}$)²⁷ incluso en ausencia del enlace disulfuro C22-C92 conservado es suficientemente alta para plegarse de forma apropiada. La ausencia de un par extra de cisteínas en las CDR en el dominio de unión a epítipo de HCAb derivados de llama podría ser por lo tanto una ventaja natural frente a dominios VHH de dromedarios, cuando sea esencial un enlace disulfuro entre bucles de aparición frecuente para el reconocimiento de antígenos^{9,28}. El cromocuerpo de GFP, en el que una de las cisteínas conservadas en la región VHH se han mutado a serina muestra no solo una mayor estabilidad de la molécula, sino sorprendentemente también una mejor accesibilidad de la molécula dentro de la célula.

25 En otra realización preferida de la presente invención, la proteína de fusión de la presente invención comprende además una secuencia diana seleccionada del grupo que consiste en señal de localización nuclear (NLS), secuencia de importación al retículo endoplásmico, secuencia de importación a la mitocondria. Un ejemplo de una NLS es la secuencia de péptidos PKKKRKV (señal de localización nuclear (NLS) del antígeno T grande de SV40, D. Kalderon *et al.*, 1984, Cell 39:499) que es capaz de dirigir las proteínas heterólogas al núcleo. Un ejemplo de la "secuencia de importación al retículo endoplásmico" es el péptido MMSFVSLLLVGLFWATEAEQLTRCVFQ (señal de localización en RE de cadena ligera de inmunoglobulina, Blobel G. & B. Dobberstein, 1975 J.Cell.Biol. 67:835-851) que es capaz de dirigir proteínas heterólogas al lumen del RE. Un ejemplo de una "secuencia de importación mitocondrial" es el péptido MLSLRQSIRFFRPATRTLCSRYLL (Neupert W. 1997 Annu. Rev. Biochem. 66:863-917) que es capaz de dirigir proteínas heterólogas a la mitocondria.

35 La presente invención también se refiere a un fragmento de la proteína de fusión de la presente invención capaz de unirse específicamente a su epítipo, comprendiendo dicho fragmento o consistiendo en: marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, codificado por la molécula de ácido nucleico de una cualquiera de SEQ ID NOs: 11 o codificado por una molécula de ácido nucleico con al menos 95% y más preferentemente al menos 98% de identidad de secuencia o un fragmento de la misma. Se prefiere que los cambios con respecto a la secuencia de referencia SEQ ID NO: 1 se localicen fuera de las CDR. En una realización de la presente invención, sin embargo, se introducen mutaciones particulares en al menos una posición de la CDR para modular la afinidad del fragmento por GFP. "Modular la afinidad" significa aumentar o reducir la constante de disociación K_D del complejo de GFP y fragmento.

45 Como se ilustra en el Ejemplo 6 de la presente invención, dicho fragmento es particularmente eficaz en inmunoprecipitación de GFP. Los resultados descritos en el presente documento son bastante sorprendentes puesto que este fragmento, aunque no contiene la parte Fc de los anticuerpos convencionales, aún es capaz de asociarse con la proteína A. Los experimentos mostrados en el presente documento demuestran que este fragmento es altamente eficaz en ensayos de inmunoprecipitación convencionales. Además este fragmento no solamente inmunoprecipita su antígeno (como se demuestra en la Figura 7B) sino también proteínas de interacción (como se muestra en la Figura 8). El bajo peso molecular y el estado monomérico de este fragmento de anticuerpo permite una manipulación sencilla en enfoques bioquímicos, comenzando con la expresión de alto rendimiento en una célula hospedadora como *E. coli* y una purificación de una etapa (mostrada en la Figura 6A). A diferencia de los anticuerpos convencionales este fragmento de anticuerpo, debido a su naturaleza, no interfiere en aplicaciones como SDS-PAGE y/o una transferencia de western. Los anticuerpos convencionales muestran una cadena ligera y pesada desnaturalizada, que con frecuencia interfiere con la detección del antígeno y/o proteínas coprecipitadas por inmunodetección o análisis de espectrometría de masas.

60 En otra realización preferida de la presente invención, la proteína de la presente invención comprende además (a) un marcador seleccionado del grupo que consiste en marcador His, marcador Strep, sitio de reconocimiento para biotinylation y opcionalmente (b) el sitio de reconocimiento para una proteasa. También se describe un marcador Hyc y marcador GST. Hablando en general, puede usarse cualquier epítipo con un anticuerpo de alta afinidad correspondiente como un marcador. Se prefieren particularmente marcador myc (véase, por ejemplo, Kiecke, 1997, Protein Eng. 10: 1303-1310), marcador His (Farmacia) o marcador GST (Farmacia) o marcador Strep (véase por ejemplo Skerra & Schmidt (1999) Biomolecular Engineering 16: 79-86). Uno de los enfoques más habitualmente usados es marcar la proteína recombinante en uno de los extremos con la enzima glutatión S-transferasa (GST) o

un fragmento de la misma, que tiene alta afinidad por su ligando natural glutatión (que es el tripéptido γ -glutamilcisteinilglicina, habitualmente abreviado GSH). La proteína se purifica usando una resina con GSH unida covalentemente. Después de la elución de las proteínas contaminantes con tampón, la proteína marcada se eluye con una solución de GSH. Otro método popular es marcar la proteína en uno de los extremos con 6-10 restos de His que confieren a la proteína recombinante la capacidad de unirse con una resina de Ni^{2+} . Después de la elución de proteínas contaminantes con tampón, la proteína marcada se separa por competencia de la columna Ni^{2+} con una solución de tampón que contiene imidazol (recuérdese que la cadena lateral de His contiene un anillo de imidazol).

La presente invención también se refiere a una molécula de ácido nucleico que codifica la proteína de fusión de la presente invención o un fragmento de la misma. Dicha molécula de ácido nucleico puede ser una molécula de ARN o ADN.

La presente invención también se refiere a un vector de expresión que comprende la molécula de ácido nucleico de la presente invención. El vector de expresión puede ser un vector de expresión eucariota o procariota, preferentemente un vector de expresión de mamífero.

Un vector de expresión de mamífero típico contiene el elemento promotor, que media en el inicio de la transcripción de ARNm, la secuencia codificante de proteína y señales requeridas para la terminación de la transcripción y poliadenilación del transcrito. Además, también pueden incluirse elementos tales como origen de replicación, gen de resistencia a fármacos, reguladores (como parte de un promotor inducible). El promotor *lac* es un promotor inducible típico, útil para células procariotas, que puede inducirse usando el análogo de lactosa *isopropilol-b-D-galactósido*. ("IPTG"). Para expresión recombinante, el fragmento de anticuerpo puede ligarse entre por ejemplo la señal líder de PelB, que dirige la proteína recombinante en el periplasma y el gen III en un fagémido llamado pHEN4 (descrito en Ghahroudi *et al*, 1997, FEBS Letters 414521-526). Los elementos adicionales podrían incluir potenciadores, secuencias Kozak y secuencias intermedias flanqueadas por sitios donadores y aceptores para corte y empalme de ARN. Puede conseguirse transcripción altamente eficaz con los promotores temprano y tardío de SV40, las repeticiones terminales largas (LTR) de retrovirus, por ejemplo, RSV, HTLV1, VIH1, y el promotor temprano del citomegalovirus (CMV). Sin embargo, también pueden usarse elementos celulares (por ejemplo, el promotor de actina humano). Los vectores de expresión adecuados para su uso en la práctica de la presente invención incluyen, por ejemplo, vectores tales como pSVL y pMSG (Farmacia, Uppsala, Suecia), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) y pBC12MI (ATCC 67109). Las células hospedadoras de mamífero que podrían usarse incluyen células Hela, 293, H9 y Jurkat humanas, células NIH3T3 y C127 de ratón, Cos 1, Cos 7 y CV1, células QC1-3 de codorniz, células L de ratón y células de ovario de hámster chino (CHO). Como alternativa, el (poli) péptido recombinante puede expresarse en líneas celulares estables que contienen la construcción génica integrada en un cromosoma. La cotransfección con un marcador seleccionable tal como dhfr, gpt, neomicina, higromicina permite la identificación y aislamiento de las células transfectadas. El ácido nucleico transfectado también puede amplificarse para expresar grandes cantidades del (poli) péptido codificado. El marcador de DHFR (dihidrofolato reductasa) es útil para desarrollar líneas celulares que porten varios cientos o incluso varios miles de copias del gen de interés. Otro marcador de selección útil es la enzima glutamina sintasa (GS) (Murphy *et al*. 1991, Biochem J. 227: 277-279; Bebbington *et al*. 1992, Bio/Technology 10: 169-175). Usando estos marcadores, las células de mamífero se cultivan en medio selectivo y se seleccionan las células con la mayor resistencia. Como se ha indicado anteriormente, los vectores de expresión preferentemente incluirán al menos un marcador seleccionable. Tales marcadores incluyen dihidrofolato reductasa, G418 o resistencia a neomicina para cultivo de células eucariotas y genes de resistencia a tetraciclina, kanamicina o ampicilina para cultivar en *E. coli* y otras bacterias. Los ejemplos representativos de hospedadores apropiados incluyen, pero sin limitación, células bacterianas, tales como células de *E. coli*, *Streptomyces* y *Salmonella typhimurium*; células fúngicas, tales como células de levadura; células de insecto tales como células de *Drosophila* S2 y *Spodoptera* Sf9; células animales tales como células CHO, COS, 293 y melanoma de Bowes; y células vegetales. Se conocen en la técnica medios y condiciones de cultivo apropiados para las células hospedadoras anteriormente descritas.

El polipéptido expresado de forma recombinante puede contener restos de aminoácidos adicionales para aumentar la estabilidad o para modificar la dirección de la proteína. Por ejemplo, puede añadirse una región de aminoácidos adicionales, particularmente aminoácidos cargados, al extremo N terminal del polipéptido para mejorar la estabilidad y persistencia en la célula hospedadora, durante la purificación, o durante la manipulación y almacenamiento posteriores. Además, pueden añadirse restos peptídicos al polipéptido para facilitar la purificación. Tales regiones pueden retirarse antes de la preparación final del polipéptido. La adición de restos peptídicos a polipéptidos para inducir secreción o excreción, para mejorar la estabilidad y para facilitar la purificación, entre otros, son técnicas familiares y rutinarias en este campo. Una proteína de fusión preferida comprende una región heteróloga de inmunoglobulina que es útil para estabilizar y purificar proteínas. Por ejemplo, el documento EP-A-0 464 533 (homólogo Canadiense 2045869) describe proteínas de fusión que comprenden diversas partes de región constante de moléculas de inmunoglobulina junto con otra proteína humana o parte de la misma. En muchos casos, la parte de Fc en una proteína de fusión es verdaderamente ventajosa para su uso en terapia y diagnóstico y por lo tanto da como resultado, por ejemplo, propiedades farmacocinéticas mejoradas (documento EP-A 0 232262). Por otro lado, para algunos usos sería deseable poder suprimir la parte Fc después de que la proteína de fusión se haya expresado, detectado y purificado de la manera ventajosa descrita. En el descubrimiento de fármacos, por ejemplo, se han fusionado proteínas humanas, tales como hIL-5, con partes Fc para el fin de realizar ensayos de exploración

de alto rendimiento para identificar antagonistas de hIL-5. Véase, D. Bennett *et al.*, J. Molecular Recognition 8: 52-58 (1995) y K. Johanson *et al.*, J. Biol. Chem. 270: 9459-9471 (1995). Las proteínas de fusión y proteínas de la presente invención pueden recuperarse y purificarse de cultivos de células recombinantes por métodos bien conocidos incluyendo precipitación con sulfato de amonio o etanol, extracción ácida, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad y/o cromatografía de hidroxipatita. Más preferentemente, se usa cromatografía líquida de proteínas rápida (FPLC) para purificación, sin embargo, también puede emplearse cromatografía líquida de alto rendimiento ("HPLC") para purificación.

La presente invención también se refiere a una célula hospedadora que comprende la molécula de ácido nucleico de la presente invención y/o el vector de expresión de la presente invención. La célula hospedadora puede ser una célula hospedadora eucariota o procariota. Preferentemente las células hospedadoras para usar para expresión del (poli) péptido, proteína o proteína de fusión de la presente invención se seleccionan del grupo que consiste en células bacterianas incluyendo *E. coli*: X11 blue, BL21, JM 109; células eucariotas inferiores incluyendo levadura *Saccharomyces cerevisiae*, *Pichia Pastoris*; o cualquier célula o cepa descrita en Sambrook y Russel, 2001, Molecular Cloning: A Laboratory Manual 3ª edición, Cold Spring Harbor Laboratory Press, Apéndice 3.

También se describen en el presente documento proteínas de fusión de acuerdo con la presente invención o fragmentos de las mismas, que pueden obtenerse de la célula hospedadora de la presente invención o del método de la presente invención. La proteína recombinante puede recuperarse de células o sobrenadante de cultivo celular por cualquier método adecuado conocido por el experto en la materia. Se describen métodos típicos en Bollag, D.M. y Edelstein, S.J., 1991 (véase Protein Methods Wiley-Liss, Nueva York, Methods in Enzymology, 1990, Guide to Protein Purification, Vol. 182, Academic Press, Nueva York).

La presente memoria descriptiva también describe un método para generar una biblioteca de proteínas detectables, tales como fluorescentes, capaces de unión a antígeno, que comprende las etapas del método como se ha descrito anteriormente. También se describe una biblioteca generada por este método. Dichas bibliotecas permiten una exploración directa y rápida de nuevas proteínas detectables, tales como fluorescentes, capaces de unión a antígeno. Puede evitarse el enfoque caro y que requiere mucho tiempo que implica la inmunización animal tradicional. Además, se necesita menos antígeno y pueden evitarse los efectos adversos de antígenos biológicos durante la inmunización. Estas nuevas proteínas capaces de unión a antígeno están inmediatamente listas para su uso, sin ningún procedimiento experimental adicional como reticulación química con, por ejemplo, fluoróforos o enzimas cromogénicas o estructuras moleculares para detección o cuantificación. Los fragmentos de anticuerpo fluorescentes (cromocuerpos) identificados en la exploración inicial pueden mejorarse adicionalmente en su especificidad y afinidad por mutagénesis posterior y ciclos de selección adicionales.

Se describe también en el presente documento un método para purificar la proteína o proteína de fusión de la presente invención, que comprende (a) poner en contacto una solución que comprende la proteína con un compuesto capaz de unirse específicamente a la proteína, en la que antes o después de la unión, el compuesto se une a un soporte sólido; (b) lavar el soporte sólido de la etapa (a) para retirar constituyentes unidos de forma no específica; (c) eluir la proteína de fusión.

Un "compuesto capaz de unirse específicamente a la proteína" puede ser un anticuerpo o fragmento de anticuerpo o un antígeno para el que es específica la proteína la fusión de la presente invención. Los soportes sólidos que pueden emplearse de acuerdo con la invención incluyen material de filtro, microplacas, obleas, placas de microtitulación y perlas, por nombrar algunas. La unión significa unión covalente o no covalente. Las proteínas y proteínas de fusión de la presente invención pueden solubilizarse en tampones de proteínas convencionales tales como PBS. 1xPBS con un pH de 7,5 es el tampón preferido, pero también pueden usarse otros tampones (véase Sambrook *et al.*, Molecular Cloning: A Laboratory Manual 3ª edición, Cold Spring Harbor Laboratory Press, Apéndice A1.2, 2001). Estos otros tampones pueden seleccionarse por el experto en la materia usando métodos convencionales conocidos en la técnica. La etapa de lavado puede realizarse a diferente concentración salina para modificar o modular la especificidad de la unión proteica. Por ejemplo en 1xPBS que contiene NaCl o KCl 150 mM - 1 M. El procedimiento completo se basa en un protocolo convencional para, por ejemplo, inmunoprecipitación conjunta proporcionada por el Kit de Inmunoprecipitación Conjunta ProFound™ Pierce, Rockford, Estados Unidos. También puede usarse cualquier otro protocolo descrito para purificación proteica o identificación de interacciones proteicas y purificación proteica usando anticuerpos (Referencias: Methods in Enzymology, 1990, Guide to Protein Purification, Vol. 182, Academic Press, Nueva York; Harlow, E. y Lane, D. 1988 Antibodies, A Laboratory Manual, Capítulo 11, Cold Spring Harbor, Nueva York; Hermanson, G.T., 1992, Immobilized Affinity Ligands Techniques, Academic Press, Inc., San Diego, CA; Sambrook y Russel, 2001, Molecular Cloning: A Laboratory Manual 3ª edición, Cold Spring Harbor Laboratory Press, Apéndice A1.2). En una realización preferida de la presente invención, la proteína de fusión contiene una secuencia marcadora como se ha definido anteriormente. Por ejemplo puede usarse cualquier epítipo con un anticuerpo de alta afinidad correspondiente como un marcador, por ejemplo, un marcador myc (véase, por ejemplo, Kieckhefer, 1997, Protein Eng. 10: 1303-1310) o un marcador His (Pharmacia). Véase también Maier, 1998, Anal. Biochem. 259: 68-73; Muller, 1998, Anal. Biochem. 259: 54-61. En una realización más preferida de la presente invención, dicho compuesto capaz de unirse específicamente a la proteína de fusión es un compuesto que se une específicamente a la secuencia marcadora. Se han descrito varios marcadores en la memoria descriptiva

de la presente invención, todos estos marcadores pueden emplearse.

También se describe en el presente documento un método para identificar una proteína detectable, tal como fluorescente, capaz de unirse específicamente con un antígeno, que comprende las etapas de: (a) proporcionar una biblioteca creada con el método descrito anteriormente; (b) explorar la biblioteca con un antígeno; (c) seleccionar un miembro de la biblioteca que se una al antígeno. La exploración de la biblioteca puede realizarse por cualquiera de los métodos de exploración descritos en el presente documento. Preferentemente la biblioteca se explora por presentación de fagos (referencia y protocolo: Clackson T. y Lowman H.B., 2004, Phage Display: A Practical Approach, Oxford University Press Inc., Nueva York).

La presente memoria descriptiva también describe un método para detectar la presencia, cantidad o localización subcelular de una estructura antigénica de interés en una célula, que comprende las etapas de: (a) poner en contacto la célula con la proteína de fusión como se ha descrito anteriormente en condiciones que permitan que la proteína de fusión se una a dicha estructura antigénica; (b) lavar la proteína de fusión no unida de la célula; (c) exponer la célula a luz correspondiente a la longitud de onda de excitación de la proteína de fusión; y (d) detectar la energía, tal como fluorescencia emitida de la célula, en el que la energía detectada tal como fluorescencia indica la presencia de la estructura en la superficie celular. El pequeño tamaño del dominio de unión a antígeno permite la detección de antígenos de otro modo inaccesibles en una superficie celular. Además, permite una mayor resolución de estructuras biológicas complejas. Mediante la visualización directa del antígeno en una superficie celular usando el cromocuerpo, puede tener lugar una detección estequiométrica con respecto a cuantificación lineal.

La presente memoria descriptiva también describe un método para detectar la presencia o cantidad de una estructura antigénica de interés, que comprende las etapas de: (a) poner en contacto una muestra sospechosa de contener la estructura antigénica de interés con la proteína de fusión de la presente invención en condiciones que permitan que la proteína de fusión se una con la proteína de interés; (b) retirar la proteína de fusión no unida de la muestra; (c) exponer la muestra a luz correspondiente a la longitud de onda de excitación de la proteína de fusión; y (d) detectar la energía (tal como fluorescencia, fosforescencia, etc.) emitida desde la muestra, en el que la energía detectada indica la presencia o cantidad de la estructura antigénica. Este enfoque bioquímico permite una detección directa de un antígeno usando la fluorescencia de la proteína de fusión unida. A diferencia de métodos convencionales que implican detección de anticuerpos secundarios o reacciones de seguimiento enzimáticas éste proporciona un método sencillo y rápido. Por lo tanto este método es ideal para estudios de exploración de extractos celulares o soluciones proteicas líquidas con respecto a la presencia del antígeno de interés. La correlación directa de la fluorescencia de la proteína de fusión unida y el antígeno da como resultado una cuantificación sencilla y directa del antígeno de interés en la sonda. Preferentemente la energía detectada es energía emitida desde una proteína fluorescente, fosforescente o cromogénica.

La presente memoria descriptiva describe que la proteína de fusión o la estructura antigénica contenida en la muestra se acoplan a un soporte sólido.

El método comprende la etapa adicional de realizar cromatografía antes de exponer la muestra a luz correspondiente a la longitud de onda de excitación de la proteína de fusión.

La presente memoria descriptiva describe que dicha estructura antigénica se selecciona de proteína, modificación de proteína, cofactor, compuesto molecular pequeño, ADN y ARN. La memoria descriptiva se refiere además a un método para purificar el fragmento de la proteína de fusión como se ha descrito anteriormente, que comprende las etapas de: (a) poner en contacto una solución que comprende dicho fragmento con un anticuerpo o un compuesto capaz de unirse específicamente a la secuencia marcadora, en el que antes o después de la unión, el anticuerpo o compuesto de unión marcador se une al soporte sólido; (b) lavar el soporte sólido de la etapa (a) para retirar constituyentes unidos de forma no específica; (c) eluir la proteína.

Se describe además en el presente documento un método para acoplar el fragmento purificado con un soporte inmovilizado, que comprende las etapas de: (a) poner en contacto una solución que comprende el fragmento proteico con un anticuerpo o un compuesto capaz de unirse específicamente con dicha proteína; (b) unir el anticuerpo o compuesto unido antes de o después de la unión de dicha proteína con un soporte sólido; (c) lavar el soporte sólido de la etapa (b) para retirar constituyentes unidos de forma no específica.

La presente memoria descriptiva también describe un método para purificar una estructura génica de interés, que comprende a) poner en contacto una muestra que contiene dicha estructura antigénica con l. una proteína de fusión dirigida a dicha estructura antigénica, comprendiendo dicha proteína de fusión una primera secuencia (poli) peptídica que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* y una segunda secuencia (poli) peptídica, que es una proteína detectable, preferentemente derivable de una proteína fluorescente o cromófora, en la que dicha (1.) primera secuencia (poli) peptídica está compuesta de marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, codificada por la secuencia de ácido nucleico de SEQ ID NO: 2 o codificada por una secuencia de ácido nucleico con al menos 70% de identidad de secuencia o un fragmento de la misma; y (2.) segunda secuencia (poli) peptídica es una proteína detectable, preferentemente (i) la proteína verde fluorescente derivable de *Aequorea victoria* codificada por la secuencia de ácido nucleico de SEQ ID NO: 7, o un mutante fluorescente o fragmento de la

misma; (ii) la proteína roja fluorescente derivable de *Discosoma* (DsRed) codificada por la secuencia de ácido nucleico de SEQ ID NO: 9, o un mutante fluorescente o fragmento del mismo; o (iii) un homólogo funcional de (i) o (ii) con al menos el 80% de identidad de secuencia, preferentemente al menos 85%, aún más preferentemente al menos 90%, aún más preferentemente al menos 95% y más preferentemente con al menos 98% de identidad de secuencia; en la que dicha primera secuencia (poli) peptídica se localiza en dirección N terminal o C terminal de dicha segunda secuencia (poli) peptídica, estando separadas opcionalmente dichas secuencias por un enlazador de al menos un resto de aminoácido, o II. un (poli) péptido que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* compuesto de marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, codificado por la secuencia de ácido nucleico de SEQ ID NO: 2 o codificada por una secuencia de ácido nucleico con al menos 70% de identidad de secuencia o un fragmento de la misma; en el que la proteína de fusión está unida a un soporte sólido; b) opcionalmente lavar el soporte sólido de la etapa a) para retirar constituyentes unidos de forma no específica; c) eluir la estructura antigénica.

El término "purificar" como se usa a lo largo de la presente memoria descriptiva resume técnicas usadas para aislar una estructura antigénica. El término comprende, pero sin limitación, purificación por inmunoprecipitación, cromatografía de exclusión por tamaño o afinidad, y similares y cualquier combinación de las mismas.

La presente memoria descriptiva describe que dicha región variable comprende marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3, y está codificada por la secuencia de ácido nucleico de SEQ ID NO: 2 o está codificada por una secuencia de ácido nucleico con al menos 70%, más preferentemente 80%, más preferentemente al menos 85%, aún más preferentemente al menos 90%, aún más preferentemente al menos 95% y más preferentemente al menos 98% de identidad de secuencia o un fragmento de la misma. Preferentemente dicho fragmento es un fragmento funcional, es decir un fragmento capaz de unión a antígeno específica. La "unión específica" puede describirse, por ejemplo, con respecto a reactividad cruzada. Preferentemente, los agentes de unión específicos son proteínas (proteínas de fusión o fragmentos de las mismas) que tienen una constante de disociación o K_D de menos de 10^{-12} M en particular 5×10^{-13} M, 10^{-13} M, 5×10^{-14} M, 10^{-14} M, 5×10^{-15} M y 10^{-15} M.

También se describe que la secuencia de (a) CDR1 consiste en los restos mostrados en SEQ ID NO: 3; (b) CDR2 consiste en los restos mostrados en SEQ ID NO: 4; y (c) CDR3 consiste en los restos mostrados en SEQ ID NO: 5.

Se describe además que dicha primera secuencia (poli) peptídica de dicha proteína de fusión o el (poli) péptido que comprende la región variable de un anticuerpo de cadena pesada de *Camelidae* tiene la secuencia de una cualquiera de SEQ ID NOs: 10, 30, 12 o 14 o está codificada por la secuencia de una cualquiera de SEQ ID NOs: 11, 31, 13 o 15.

La presente memoria descriptiva describe que dicha segunda secuencia (poli) peptídica comprende los restos de 1 a 239 de SEQ ID NO: 6 o 1 a 226 de SEQ ID NO: 8 o un mutante fluorescente o fragmento del mismo.

La presente memoria descriptiva también describe que dicho mutante de la proteína roja fluorescente es mRFP1 como se codifica por SEQ ID NO: 17 o cualquiera de los mutantes de DsRed fluorescentes codificados por una cualquiera de SEQ ID NOs: 17 y 24 a 27 o por un fragmento de los mismos. Puede usarse cualquier mutante fluorescente derivado de DsRed de acuerdo con las enseñanzas en el presente documento. En particular, las proteínas de fusión generadas de acuerdo con las enseñanzas del presente documento pueden comprender una segunda secuencia (poli) peptídica fluorescente que comprende cualquiera de los mutantes de DsRed anteriormente mencionados o mutantes fluorescentes o fragmentos de los mismos.

La presente memoria descriptiva también describe que la proteína de fusión de la presente invención comprende además una secuencia de dirección seleccionada del grupo que consiste en señal de localización nuclear (NLS), secuencia de importación al retículo endoplásmico, secuencia de importación mitocondrial.

La presente memoria descriptiva se refiere a una proteína de fusión que tiene la secuencia de una cualquiera de SEQ ID NOs: 18, 20 o 22 o que se codifica por una molécula de ácido nucleico que comprende la secuencia de una cualquiera de SEQ ID NOs: 19, 21 o 23.

La presente invención también se refiere a un método para precipitar específicamente GFP o una proteína marcada con GFP u otros antígenos celulares en una muestra, que comprende la etapa de poner en contacto la muestra con el cromocuerpo anti-GFP de la presente invención que es específico para GFP (SEQ ID NOs: 18 y 32) de la presente invención o un fragmento del mismo. Este método también puede usarse para detectar las proteínas o antígenos posteriores. En una realización preferida de la presente invención, dicho fragmento comprende o consiste en los restos de aminoácidos como se muestran en SEQ ID NOs: 10 o 30.

En una realización preferida de la presente invención, el cromocuerpo anti-GFP o fragmento del mismo comprende además un marcador seleccionado del grupo que consiste en marcador de His, marcador de Strep, sitio de reconocimiento para biotilación; y opcionalmente el sitio de reconocimiento para una proteasa.

En otra realización preferida de la presente invención, se extraen (es decir se precipitan) GFP o una proteína marcada con GFP u otros antígenos celulares y proteínas de interacción poniendo en contacto una muestra con la

proteína purificada de la presente invención o un fragmento unido a un soporte sólido. Las proteínas precipitadas conjuntamente pueden después analizarse por métodos convencionales incluyendo análisis de espectrometría de masas.

5 En una realización más preferida que se refiere a cualquier realización de la presente invención, la proteína de fusión o fragmento se une a un soporte sólido.

En una realización más preferida de la presente invención dicho soporte sólido es proteína A. La proteína A en sí misma puede acoplarse a resinas habituales tales como Sepharose.

10 En una realización más preferida de la presente invención dicha proteína de fusión o fragmentos se une a dicho soporte sólido. La expresión "unido a" comprende acoplado covalentemente. En general, las proteínas pueden acoplarse a matrices mediante grupos amino primarios.

15 En una realización aún más preferida de la presente invención dicho soporte sólido es Sepharose. Las matrices o soportes sólidos son por ejemplo Sepharose 4 activada por NHS (Amersham), medio de afinidad activado por Affi-Gel 10 o Affi-Gel 15 (BIO-RAD) o Gel de Acoplamiento AminoLink Plus (Pierce). También se describe en el presente documento un kit que comprende la proteína de la presente invención, la molécula de ácido nucleico de la presente invención, el vector de expresión de la presente invención, la célula hospedadora de la presente invención y/o la biblioteca de la presente invención y opcionalmente instrucciones para su uso.

Las Figuras muestran:

25 **Figura 1:** Perfil esquemático de un anticuerpo IgG convencional en comparación con un anticuerpo IgG de cadena pesada derivado de llama (adaptado de ref. 11). Los dominios de unión a antígeno de un IgG convencional (**a**) están compuestos de las cadenas ligera (VL) y pesada (VH) variables. Por el contrario, en anticuerpo IgG de cadena pesada (**b**) de *Camelidae* está desprovisto de la cadena ligera (VL & CL) y carece también del dominio CH1. La unión a antígenos se lleva a cabo por tres regiones determinantes de complementariedad (CDR) localizadas en el dominio variable de la cadena pesada (VHH). Para generar un cromocuerpo fluorescente el VHH de pequeño tamaño se fusiona con una proteína roja fluorescente monomérica (mRFP1) y se expresa en células de mamífero. (**c**) La estructura potencial del cromocuerpo se modeló basándose en las estructuras cristalinas conocidas de un VHH¹¹ de llama y el mRFP1³¹. Dentro de esta estructura CDR1 se muestra en amarillo, CDR2 en cian, CDR3 en verde, mRFP1 en rojo y el enlazador flexible de 24 aminoácidos se muestra como una línea de puntos negra.

35 **Figura 2:** Alineamiento de siete secuencias VHH derivadas de llama (líneas 1-4) y camello (líneas 5-7). Desde arriba a abajo: cAb-GFP4 (cAb de llama alpaca inducido contra GFP, este trabajo); RR6-R2 (cAb de llama inducido contra colorante rojo RR6³²); RR6-R9 (cAb de llama inducido contra colorante rojo RR6³³); hCG-H14 (cAb de llama inducido contra hCG³⁴); cAb-Lys3 (cAb de camello inducido contra lisozima²⁸); cAb-TT1 (cAb de camello inducido contra toxoide del tétanos¹²), AMYL-D10 (cAb de camello inducido contra amilasa porcina³⁵). Los restos de CDR están coloreados en amarillo, cian y verde para CDR 1 (H1), 2 y 3 respectivamente. Las sustituciones de aminoácidos de VHH de camélido características en marco 2 están destacadas en azul. Las cisteínas se muestran en negrita. La numeración es de acuerdo con Kabat y Wu¹⁸. Debajo se indican los restos de aminoácidos idénticos (*), muy similares (:) y similares (.) .

40 **Figura 3:** El cromocuerpo anti-GFP se expresa como una proteína monomérica estable en células de mamífero. (**a**) representación esquemática del cromocuerpo anti-GFP. (**b**) se prepararon extractos celulares totales de células 293T que expresaban cromocuerpo anti-GFP o células tratadas con simulación 72 horas después de la transfección. Se analizaron cantidades crecientes de proteína por transferencia de western y detección con antisuero anti-mRFP1. Carril 1; 30 µg de extracto proteico de células transfectadas con simulación, carril 2-4, 10, 20 y 30 µg de extracto proteico de células que expresan cromocuerpo anti-GFP. El tamaño predicho de la proteína quimérica es 41 kDa (panel superior). Como un control de carga, la transferencia se volvió a incubarse con un anticuerpo contra PCNA (panel inferior). (**c**) Análisis de filtración en gel de extractos de células de mamífero que expresan el cromocuerpo. Se lisaron 1×10^7 células 293T que expresaban cromocuerpo anti-GFP en presencia de NP-40 0,5% y se sometieron a filtración en gel en una columna de Superose-12. Las fracciones se analizaron por SDS-PAGE y transferencia de western usando un anticuerpo contra mRFP1. El cromocuerpo anti-GFP se eluye de la columna en fracciones pico correspondientes a una masa molecular aparente de ~ 40 kDa. Las flechas indican la elución de patrones de calibración. (**d**) La expresión del cromocuerpo anti-GFP en células HeLa en ausencia de cualquier antígeno muestra una distribución difusa de la proteína por todo el citoplasma y el núcleo (panel superior), que es comparable a la distribución de mRFP1 no fusionada en el mismo tipo celular (panel inferior).

Figura 4: Dirección del cromocuerpo anti-GFP a sitios de unión de proteínas de fusión de GFP en diferentes compartimentos y estructuras celulares en células vivas. **(a)** Representación esquemática de las proteínas de fusión. **(b-e)** Se transfectaron de forma transitoria células HeLa con las construcciones de fusión de GFP indicadas y el plásmido de expresión de cromocuerpo anti-GFP (con la excepción de **(e)** en el que se usaron células HeLa que expresaban de forma estable H2B-GFP). Se muestran imágenes representativas de células vivas. **(b)** Los cromocuerpos anti-GFP se colocalizan con GFP- β -actina en filamentos de actina citoesqueléticos. **(c)** Los cromocuerpos anti-GFP entran en el núcleo y se unen a GFP-laminina B1 incorporada en la lámina nuclear. **(d)** Los cromocuerpos anti-GFP se unen a una proteína regulada por ciclo celular, GFP-PCNA, en focos de replicación. Se tomó una serie de lapso de tiempo de una célula que atravesaba la fase S. Se muestra una imagen de la célula en mitad de la fase S. **(e)** Los cromocuerpos anti-GFP se unen a la histona H2B-GFP incorporada en cromatina. La célula transfectada con el cromocuerpo anti-GFP está en prometafase y son visibles los cromosomas separados. Las flechas indican células no transfectadas en meta e interfase que no expresan el cromocuerpo. La ausencia de fluorescencia roja confirma que no hay escape inespecífico a través del canal de GFP. Barras de escala, 5 μ m.

Figura 5: Localización de antígenos con cromocuerpos. **(a)** Localización de una proteína de cromatina a lo largo de la mitosis. Imagen de lapso de tiempo de una célula HeLa que expresa de forma estable histona H2B-GFP transfectada con cromocuerpo anti-GFP. Se muestran fotogramas seleccionados de esta serie temporal. En el momento en que se inició la captura de imágenes (0 horas), esta célula estaba en fase G2 tardía. **(b)** Rastreo de un componente de la maquinaria de replicación a lo largo de la fase S hasta G2. Imágenes de lapso de tiempo de células HeLa cotransfectadas que expresan GFP-PCNA y cromocuerpo anti-GFP. Se muestran fotogramas seleccionados de esta serie. Al comienzo de la captura de imágenes (0 horas) las células estaban en fase S temprana a media. Barras de escala, 5 μ m.

Figura 6: Purificación de cAb-GFP4-His₆ recombinante **(A)**. El fragmento de anticuerpo cAb-GFP4 expresado recombinante puede purificarse en una etapa de cromatografía. Se resuspendieron 5 ml de extracto de proteína de *E. coli* soluble en tampón de unión (1x PBS, NaCl 500 mM, imidazol 20 mM) y se cargó en una columna que contenía 1 ml de resina NiNTA. Después de lavado exhaustivo, la proteína unida se eluyó añadiendo por etapas 1 ml de tampón de elución (1x PBS, NaCl 500 mM, imidazol 200 mM). Se sometieron 2% de sobrenadante, sedimento después de centrifugación, flujo continuo y fracción de elución a SDS-PAGE y se tiñeron con azul brillante de coomassie. La fracción principal de cAb-GFP4 eluye en las fracciones 2 y 3, que se agruparon y dializaron frente a 1x PBS y se ajustaron a una concentración de 1 μ g/ μ l. **(B)** Análisis de filtración en gel del fragmento de anticuerpo cAb-GFP purificado. Se sometieron 10 μ g de cAb-GFP4 purificado a filtración en gel en una columna de Superose-12. Las fracciones se analizaron por SDS-PAGE y transferencia de western usando un anticuerpo contra el marcador de histidina C terminal (His₆). El fragmento de anticuerpo cAb-GFP4 eluye desde la columna en fracciones pico correspondientes a una masa molecular aparente de - 15 kDa. Las flechas indican la elución de patrones de calibración.

Figura 7: Inmunoprecipitación de proteínas marcadas con GFP usando el fragmento de anticuerpo cAb-GFP4 purificado **(A)** representación esquemática de las proteínas de fusión. **(B)** Se lisaron Células 293T tratadas con simulación o que expresaban construcción de fusión de GFP en presencia de NP-40 0,5%. Los extractos de proteína soluble se sometieron a inmunoprecipitación añadiendo 20 μ g, respectivamente 10 μ g de cAb-GFP4 purificado acoplado con perlas de agarosa de proteína A. Las fracciones únicas que contenían entrada, flujo continuo, lavado y sedimento de perlas se analizaron por SDS-PAGE seguido de tinción con coomassie o análisis de inmunotransferencia. Las proteínas sobreexpresadas eran parcialmente detectables en la fracción de entrada. En la fracción de perlas las proteínas precipitadas eran detectables como una banda clara junto con el fragmento de anticuerpo cAb-GFP4 (indicado con flechas). Para confirmar que estas bandas eran las proteínas de fusión con GFP esperadas, respectivamente el fragmento de anticuerpo, los geles se transfirieron y se incubaron con anticuerpos contra GFP y el marcador de histidina C terminal (His₆). Las señales de anticuerpos revelan la precipitación casi cuantitativa del antígeno, mientras que no aparecen bandas proteicas no específicas en el gel que muestra las células tratadas con simulación. Como control para la especificidad de la inmunoprecipitación, las manchas de transferencia se volvieron a explorar con un anticuerpo contra β actina. Solo aparecen señales correspondientes en los carriles de entrada y flujo continuo, lo que demuestra que el cAb-GFP4 precipita su antígeno con una alta especificidad.

Figura 8: Análisis de inmunoprecipitación conjunta. Para confirmar que el fragmento de anticuerpo cAb-GFP4 no solo extrae su epítipo sino que también proteínas que interaccionan, se realizó análisis de inmunoprecipitación conjunta expresando un dominio marcado con GFP de Dnmt1 (aa 158-629), que se sabe que interacciona con PCNA. Se llevó a cabo inmunoprecipitación como se ha

descrito y la inmunotransferencia se analizó con anticuerpos contra GFP y PCNA. Aunque el GFP-Dnmt1₍₁₅₈₋₆₂₉₎ precipitado es claramente visible como una banda fuerte en el peso molecular correspondiente, se precipitó conjuntamente también una fracción significativa del PCNA endógeno.

5 **Figura 9:** Reconocimiento de estructuras subcelulares por cromocuerpos inducidos contra epítomos endógenos. **(A)** Se transfectaron transitoriamente células de mioblasto murino (C2C12) con el plásmido de expresión de cromocuerpo anti-citoqueratina 8. Después de cinco horas las células se sometieron a microscopía de células vivas. Para obtener una mayor resolución las células se fijaron, se tiñeron con DAPI para detectar ADN y se analizaron con microscopía confocal. Se muestra una imagen representativa de un lapso de tiempo de células vivas. La fluorescencia roja en filamentos citoplasmáticos indica el reconocimiento de fibras de citoqueratina. **(B)** Se transfectaron transitoriamente células HeLa con el plásmido de expresión de cromocuerpo anti-laminina. Como se describe en 9 (A) las células que expresaban cromocuerpo anti-laminina se analizaron por microscopía confocal. La fluorescencia roja del cromocuerpo muestra una clara estructura circundante del núcleo, lo que indica la tinción de la lámina nuclear. Barras de escala, 5 μm .

10 **Figura 10:** Comparación de la capacidad de precipitación del agente de unión de GFP con anticuerpos de GFP mono y policlonales. Se incubaron extractos de células 293T productoras de GFP con 1 μg de agente de unión de GFP (cAbGFP4) o 2 μg de anticuerpos de GFP mono o policlonales y se precipitaron inmunocomplejos con proteína A/G sepharose. Se separaron las alícuotas de entrada (I), flujo continuo (F) y fracción unida (B) por SDS-PAGE seguido de tinción con coomassie (parte superior) o análisis de inmunotransferencia usando un anticuerpo anti-GFP (parte inferior). Están marcados con flechas GFP precipitado, cadenas ligera (lc) y pesada (hc) desnaturalizadas de los IgG y el agente de unión de GFP (cAbGFP4).

15 **Figura 11:** Purificación basada en columna de GFP. Se cargaron extractos proteicos de células 293T productoras de GFP en una columna que contenía agente de unión de GFP (cAbGFP4) acoplado directamente con Sepharose 10. Se eluyeron las proteínas unidas y se separaron alícuotas de entrada (I), flujo continuo (F) y fracción unida (B) por SDS-PAGE y se tiñeron con coomassie.

20 **Figura 12:** Comparación de la expresión de cAbGFP4 y cAbGFP4 (Cys92 Ser) en células de mamífero
 25 Panel superior: detección de la distribución intracelular de cAbGFP4 (original) y cAbGFP4 (Cys 92 Ser) en células 293T humanas en ausencia de antígeno. Panel inferior: detección de la funcionalidad en células 293T humanas

- 30 (a) señal de PCNA marcado con GFP
- 35 (b) señal de cAbGFP4 (original) (panel superior) y cAbGFP4 (Cys 92 Ser) (panel inferior)
- 40 (c) la fusión de ambas señales indica accesibilidad mucho mejor de cAbGFP4 (Cys 92 Ser) al núcleo.

45 **Figura 13:** Actividad y detección intracelular de agente de unión a GFP recombinante (cAbGFP4) acoplado a Cy5.
 50 Inmunotinción de células HeLa que expresan GFP-PCNA con un dominio VHH específico de GFP (cAbGPF4, agente de unión a GFP) acoplado con Cy5-NHS (Amersham). Las células se fijaron con paraformaldehído 3,7. Barra de escala 5 μm .

- 55 (a) señal de GFP-PCNA
- 60 (b) señal de Cy5 acoplado al agente de unión de GFP
- 65 (c) tinción con DAPI
- (d) fusión de (a) y (b)

Los ejemplos ilustran la invención:

Ejemplo 1: Identificación y caracterización de un VHH específico de GFP

Para ensayar la viabilidad de crear anticuerpos fluorescentes que puedan expresarse y localizarse en células vivas, los inventores eligieron GFP como la molécula diana. GFP ya se ha fusionado con una diversidad de proteínas con localización subcelular y movilidad bien caracterizadas y por lo tanto proporciona un antígeno bien conocido y "visible" que puede compararse directamente con anticuerpos fluorescentes. La diversidad de fusiones de GFP

disponibles permite a los inventores ensayar este enfoque en varios compartimientos subcelulares.

Se aislaron linfocitos de una llama alpaca (*Lama pacos*) inmunizada con GFP recombinante purificada. El repertorio de las regiones variables de la cadena pesada se amplificó por PCR y se clonó en un vector de presentación de fagos. Se obtuvo una biblioteca de VHH de 10^6 clones individuales en células TG1 de *Escherichia coli* a partir de las que se recuperó un VHH con especificidad por GFP (cAb-GFP4) después de la presentación de fagos y tres ciclos de selección. La secuencia del agente de unión mostró las sustituciones de aminoácidos en la región marco-2 (Figura 1D) que son características para un VHH de llama: Y37, E44, R45 (numeración de acuerdo con Kabat y Wu¹⁸). La naturaleza de estos aminoácidos en combinación con el resto de arginina en la posición 35 anula la interacción con un posible dominio VL, y su carácter hidrófilo hace al dominio soluble en medio acuoso. La CDR3 del cAb-GFP4 con solamente seis aminoácidos es inusualmente corta para un VHH. La ausencia de restos de cisteína adicionales además del C22 y C92 conservados es una característica común entre VHH de llamas que los distingue de los VHH de dromedario¹⁹.

Para investigar la especificidad de unión del cAb-GFP4, los inventores construyeron un plásmido de expresión bacteriana marcado con histidina (His₆) C terminal y purificaron el fragmento de anticuerpo recombinante soluble de células WK6 de *Escherichia coli*. Como se esperaba, el cAb-GFP4 se expresó en gran medida y produjo 0,7-1 mg de VHH soluble por cada 200 ml de cultivo bacteriano inducido con IPTG. El cAb-GFP4 se eluyó como un monómero con un peso molecular de 15 kDa en análisis de filtración en gel, lo que es coherente con el tamaño predicho a partir de su secuencia (datos no mostrados). Las mediciones por resonancia de plasmón superficial indicaron una velocidad cinética de asociación rápida de $7,68 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ y una disociación lenta de $1,74 \times 10^{-4} \text{ s}^{-1}$ para la interacción de cAb-GFP4 y el antígeno de GFP. La constante de disociación baja (Kd = 0,23 nM) calculada a partir de estos valores es coherente con un anticuerpo madurado por afinidad que reconoce su antígeno afin.

Ejemplo 2: Generación y expresión de una versión fluorescente del cAb-GFP4 en células de mamífero

Para investigar la dirección de antígenos intracelulares por el fragmento de anticuerpo cAb-GFP4 en células vivas, los inventores crearon un anticuerpo "visible", denominado cromocuerpo, fusionando la región codificante de la proteína roja fluorescente monomérica (mRFP1) en su extremo C terminal (Figura 3a). El cromocuerpo anti-GFP (SEQ ID NO: 18) se codifica por la secuencia de ácido nucleico de SEQ ID NO: 19. Resulta importante que tanto GFP como mRFP1 no interactúan entre sí, lo que se analizó por estudios de colocalización en células que coexpresaban GFP y mRFP1 solos o como construcciones de fusión (datos no mostrados). Los máximos de excitación y emisión distinguibles del antígeno verde fluorescente y el anticuerpo rojo fluorescente (eGFP: 498 nm / 516 nm; mRFP1: 584 nm / 607 nm) permitieron la comparación directa de la distribución celular, movilidad y dinámica de antígeno y cromocuerpo.

En primer lugar los inventores ensayaron la expresión y distribución del cromocuerpo en células de mamífero. Se transfectaron células 293T humanas con un vector de expresión que codificaba el cromocuerpo anti-GFP rojo fluorescente y se sometieron posteriormente los extractos celulares totales a análisis de inmunotransferencia usando un antisuero anti-mRFP1 policlonal. En estas células los inventores pudieron detectar una proteína de 40 kDa, que se correspondía al tamaño esperado del cromocuerpo y no estaba presente en células no transfectadas (Figura 3b). La ausencia de productos de degradación incluso a las 72 horas después de la transfección sugiere que el cromocuerpo anti-GFP es muy estable en células de mamífero. El análisis de filtración en gel muestra además que el fragmento de anticuerpo fluorescente está presente como un monómero cuando se expresa en células humanas (Figura 3C).

Para determinar la distribución subcelular del cromocuerpo anti-GFP, se fijaron células HeLa que expresaban el cromocuerpo, se tiñeron con DAPI (para destacar el ADN) y se analizaron por microscopía confocal. En un periodo de 24 horas las células presentaron la fluorescencia roja del cromocuerpo distribuida por todo el citoplasma y el núcleo (Figura 3d, panel superior), lo que es comparable con la distribución de mRFP1 solamente (Figura 3d, panel inferior). La distribución dispersada del cromocuerpo claramente demuestra el acceso a compartimientos subcelulares y la ausencia de agregación visible.

Ejemplo 3: El cromocuerpo anti-GFP fluorescente reconoce y se dirige a su epítipo en compartimientos celulares diferentes

Los inventores investigaron después la afinidad y especificidad del cromocuerpo anti-GFP fluorescente *in vivo* analizando su capacidad para reconocer y unirse a proteínas de fusión de GFP en diferentes compartimientos subcelulares y estructuras en células vivas. Se muestra una visión de conjunto esquemática de las fusiones de GFP usadas en la Figura 4a.

Para ensayar un epítipo típico localizado en el citoplasma los inventores usaron GFP-β-actina. Esta proteína de fusión se incorpora en filamentos de actina crecientes y permite la visualización de estructuras que contienen actina en células fijadas y vivas²⁰. Se cotransfectaron células HeLa con GFP-β-actina y vectores de expresión de cromocuerpo anti-GFP y 24 horas después se analizaron por microscopía de células en vivo. La Figura 4b muestra una imagen confocal representativa de una célula transfectada doble en la que la fluorescencia verde y roja es

detectable en el citoesqueleto. A partir de esto los inventores concluyen que GFP- β -actina se incorpora correctamente en los filamentos de actina, mientras que la colocalización con la fluorescencia roja resulta de la dirección de GFP- β -actina por el cromocuerpo anti-GFP. Este hallazgo demuestra que el cromocuerpo anti-GFP reconoce eficazmente y se dirige a su epítipo localizado en el citoplasma incluso cuando es parte de estructuras altamente organizadas.

Para analizar el acceso y actividad del cromocuerpo específico de GFP en otros compartimentos subcelulares, los inventores realizaron estudios de distribución y colocalización usando un componente marcado con GFP de la envoltura nuclear, GFP-laminina B1, como un epítipo. Las lamininas son un componente principal de la lámina nuclear, una estructura fibrosa que reviste la superficie nucleoplasmática de la membrana nuclear. Estudios recientes han mostrado el ensamblaje correcto de GFP-laminina B1 en la lámina nuclear de células de mamífero²¹. Las células HeLa transfectadas dobles mostraron una clara colocalización de GFP-laminina B1 con cromocuerpo anti-GFP (Figura 4C) lo que demuestra que el cromocuerpo puede entrar en el núcleo, permanece en una conformación funcional y reconoce eficazmente y se une a su antígeno específico.

Para investigar la unión de antígeno dentro del nucleoplasma los inventores eligieron el antígeno nuclear de células proliferativas fusionado con GFP (GFP-PCNA). PCNA es un componente central de la maquinaria de replicación de ADN²². GFP-PCNA, de forma similar a su homólogo endógeno, se localiza en sitios de replicación de ADN^{23,24}. Como se ha descrito anteriormente los inventores coexpresaron GFP-PCNA y el cromocuerpo anti-GFP en células HeLa. La sección media confocal mostrada en la Figura 4d revela una tinción conjunta clara de un patrón de replicación de fase S medio en una célula HeLa por GFP-PCNA y el cromocuerpo anti-GFP.

Finalmente, los inventores investigaron la unión del cromocuerpo anti-GFP con una proteína de cromatina marcada con GFP. Para este fin los inventores eligieron una línea celular HeLa estable que expresaba una versión verde fluorescente de la histona H2B (H2B-GFP). Se sabe que esta proteína de fusión se ensambla en nucleosomas como la H2B natural en sí misma²⁵. Después de la transfección y expresión del cromocuerpo anti-GFP los inventores detectaron la colocalización del cromocuerpo rojo con su epítipo verde en una célula en profase (Figura 4e). Las células no transfectadas de esta imagen actúan como un control negativo y muestran que en estas condiciones experimentales no se produce escape de fluorescencia entre el canal verde y el rojo.

Estos resultados muestran que la eficacia de unión del cromocuerpo anti-GFP parece ser independiente del contexto estructural y la localización subcelular del epítipo. Para cuantificar el grado de colocalización se calculó un coeficiente de solapamiento usando el programa adicional de colocalización del software ImageJ que produce de media un solapamiento del 90%-97% de las intensidades de fluorescencia de antígeno y cromocuerpo. A este respecto, los inventores no pueden distinguir si la unión con el antígeno se produce ya en el citoplasma o justo después de que el epítipo entre en su posición apropiada. Sin embargo los datos de los inventores muestran que las proteínas marcadas con GFP observadas se clasifican correctamente y se integran en las estructuras celulares en presencia del cromocuerpo. Juntos estos datos demuestran que el cromocuerpo anti-GFP reconoce eficazmente y se une su epítipo en el citoplasma así como en compartimientos subnucleares.

Ejemplo 4: El cromocuerpo anti-GFP fluorescente localiza su epítipo a lo largo de diferentes etapas del ciclo celular

A continuación, los inventores investigaron si el cromocuerpo anti-GFP puede usarse para localizar proteínas cuya localización cambia a lo largo del ciclo celular y realizar análisis de lapso de tiempo que localiza dos proteínas nucleares diferentes, la proteína histona H2B y la proteína de replicación PCNA. Como se ha descrito anteriormente H2B-GFP se ensambla en nucleosomas sin afectar al ciclo celular y permite captura de imágenes de alta resolución de cromosomas mitóticos y cromatina de interfase, revelando diversos estados de condensación de cromatina en células vivas²⁵. Los inventores transfectaron células HeLa H2B-GFP con el vector de expresión codificante de cromocuerpo anti-GFP y siguieron células transfectadas y no transfectadas a lo largo del ciclo celular. El análisis de lapso de tiempo (Figura 5a) mostró la colocalización del cromocuerpo anti-GFP con H2B-GFP persistente a lo largo de la mitosis. Estos resultados muestran que la expresión del cromocuerpo anti-GFP no afecta a la incorporación de H2B-GFP en cromatina ni a la progresión del ciclo celular.

Como una segunda diana los inventores seleccionaron PCNA, que constituye un reto especial para microscopía de células vivas ya que es un componente esencial y central de la maquinaria de replicación. GFP-PCNA, como el PCNA endógeno en sí mismo se concentra en focos de replicación en el núcleo durante la fase S temprana a tardía y muestra un patrón difuso en G1 y G2^{23,24}. Los inventores siguieron las células que expresaban GFP-PCNA y el cromocuerpo anti-GFP desde la fase S temprana hasta G2 tomando grupos de imágenes tridimensionales confocales cada 20 minutos. Los resultados se resumen en la Figura 5b y se muestran completamente en la información complementaria, video en línea. Tanto GFP-PCNA como el cromocuerpo mostraron cambios idénticos en su distribución subnuclear a lo largo de las fases S y G2. A partir de la asociación altamente específica del cromocuerpo anti-GFP con GFP-PCNA y la colocalización observada a lo largo de la fase S los inventores concluyeron que el cromocuerpo anti-GFP localiza GFP-PCNA en el núcleo a lo largo del ciclo celular sin alterar la replicación de ADN y progresión a fase S.

Estos análisis de lapso de tiempo muestran que pueden localizarse componentes de cromatina integrales como H2B así como componentes esenciales de la maquinaria de replicación por un cromocuerpo específico sin afectar a la

progresión y viabilidad del ciclo celular.

Ejemplo 5: Métodos usados en los Ejemplos 1 a 4

5 A. Expresión de GFP, purificación e inmunización de llamas

Se transformaron células *Escherichia coli* BL21DE3 en pRSETB-GFPS65T³⁰ (proporcionado amablemente por Roger Y. Tsien, UCSD) y se purificó GFP65T marcado con (His⁶) sobre-expresado usando cromatografía de afinidad metálica-iónica de acuerdo con las instrucciones del fabricante (Talon, Clontech, CA, Estados Unidos). Se inmunizó una llama alpaca (*Lama pacos*) con GFP purificada recombinante en adyuvante Gerbu de acuerdo con el siguiente esquema: día 0, 250 µg de GFP; día 7, 14, 21, 28 y 35, 100 µg de GFP; día 42, se recogió una muestra de sangre de 150 ml.

15 B. Construcción de biblioteca de VHH y selección del agente de unión de GFP

Se diluyó sangre heparinizada (36 ml) con RPMI precalentado y se estratificó en Lymphoprep[™] (AXIS-Shield, Oslo, Noruega) para purificar las células PBL de acuerdo con las instrucciones del fabricante. Se aisló un total de 2×10^7 PBL y se almacenaron a -80 °C en alícuotas de 6×10^6 células. El ARNm se extrajo de 6×10^6 linfocitos y se sintetizó ADNc con transcriptasa inversa RNasaH SuperscriptII (Invitrogen, CA, Estados Unidos) usando un cebador oligo-dt. La primera PCR en el molde de ADNc se realizó usando los cebadores CALL001 (5'-GTC CTG GCT GCT CTT CTA CA AGG-3') y CALL002 (5'-GGT ACG TGC TGT TGA ACT GTT CC-3') que hibridan en la secuencia líder y en el exón CH2 de las cadenas pesadas de todas las IgG de llama, respectivamente. Los productos de PCR que carecen de las secuencias CH1 (es decir, fragmentos con tamaños entre 650 y 750 pb) se purificaron a partir de un gel de agarosa usando el kit de extracción en gel de PCR QIAquick (Qiagen GmbH, Hilden, Alemania). Se realizó una PCR anidada con una mezcla equimolar de los cebadores SM017 y SM018 (5'-CCA GCC GGC CAT GGC TCA GGT GCA GCT GGT GGA GTC TGG-3' y 5'-CCA GCC GGC CAT GGC TGA TGT GCA GCT GGT GGA GTC TGG-3', respectivamente) y cebador CALL002, y el producto de PCR se volvió a purificar a partir de gel de agarosa como se ha descrito anteriormente. Los genes de VHH finalmente se volvieron a amplificar con los cebadores A4 corto (5'-CAT GCC ATG ACT CGC GGC CAC GCC CAT GGC-3') y 38 (5'-GGA CTA GTG CGG CCG CTG GAG ACG GTG ACC TGG GT-3') y se digirieron con enzimas de restricción NcoI y NotI para obtener extremos de ADN adhesivos. El fragmento se purificó con QIAquick, se ligó en el vector pHEN4¹² cortado con las mismas enzimas y la mezcla de ligación se usó para transformar células *Escherichia coli* TG1. Después de crecimiento durante una noche en placas de LB/ampicilina, las colonias bacterianas se separaron por raspado de las placas en LB, se lavaron en el mismo medio y se almacenaron en LB/glicerol al 15% a -80 °C hasta su uso posterior. Se usó una alícuota representativa de esta biblioteca para inocular LB/ampicilina hasta que las células alcanzaron la fase de crecimiento exponencial antes de la infección con fagos auxiliares M13K07 para expresar el VHH clonado. La biblioteca de VHH presentado en fagos se sometió a selección con respecto a la presencia de agentes de unión de GFP en GFP revestido de fase sólida (0,1 µg de GFP/100 µl por pocillo) durante tres ciclos consecutivos. Después del tercer ciclo de selección, se seleccionaron colonias individuales y se indujo expresión de su proteína periplásmica soluble con IPTG 1 mM. El VHH recombinante extraído del periplasma se ensayó con respecto a reconocimiento de antígenos en un ELISA.

45 C. Expresión y purificación del fragmento de anticuerpo de dominio único

El gen de VHH del clon que tuvo puntuación positiva en ELISA (cAb-GFP4) se volvió a clonar en el vector de expresión pHEN6 y se usó para transformar células de *Escherichia coli* WK6. La producción y purificación a gran escala siguió el protocolo descrito en van Koningbruggen *et al.*¹⁵

50 D. Mediciones de afinidad.

Se realizaron mediciones de afinidad mediante adición de concentraciones diferentes de GFP, que variaban de 500 nM a 7,5 nM, a VHH con cola de his purificado unido en un biochip de ácido níquel-nitrilo triacético (Biacore International AB, Uppsala, Suecia) de acuerdo con la descripción del fabricante. Los parámetros cinéticos de unión K_{on} , K_{off} y K_D se determinaron con el software BIAevaluation (versión 3.0).

55 E. Construcciones de expresión de mamíferos

La construcción plasmídica que codifica una fusión traduccional de cAb-GFP4 y mRFP1 derivó por amplificación por PCR de la región codificante de cAb-GFP4 con los cebadores gfp4#F (5'-GGG GGC TCG AGC CGG CCA TGG CCG ATG TGC AG-3') y gfp4 #RC (5'-GGG GGA ATT CCT TGA GGA GAC GGT GAC-3'). El producto de PCR se purificó como se ha descrito y se digirió con enzimas de restricción XhoI y EcoRI y se ligó en un vector pEYFP-N1 modificado (Clontech, CA, Estados Unidos) en el que la secuencia YFP se había reemplazado por la región codificante de mRFP1³¹ (proporcionado amablemente por Roger Y. Tsien, UCSD). El cromocuerpo anti-GFP (SEQ ID NO: 18) se codifica por la secuencia de ácido nucleico de SEQ ID NO: 19.

Las construcciones plasmídicas que codifican fusiones traduccionales de GFP fueron las siguientes: GFP-β-actina

(Clontech, CA, Estados Unidos), GFP-laminina B¹²¹ (proporcionada amablemente por Jan Ellenberg, EMBL), GFP-PCNA²³.

F. Cultivo y transfección de células de mamíferos

Se cultivaron células 293T, células HeLa y células HeLa que expresaban de forma estable H2B-GFP²⁵ (proporcionadas amablemente por Kevin F. Sullivan, The Scripps Research Institute) en DMEM complementado con FCS al 10%. Las células 293T se transfectaron con ADN plasmídico usando el reactivo TransFectin™ (Bio-Rad Laboratories, Hercules, CA, Estados Unidos) de acuerdo con las directrices del fabricante y se incubaron durante una noche, 48 horas o 72 horas respectivamente antes de realizar las inmunotransferencias. Para microscopía se cultivaron células HeLa hasta 50-70% de confluencia en cubreobjetos de vidrio de 18 x 18 mm, cubreobjetos de vidrio redondos de 40 mm o en cubreobjeto de vidrio con cámara Lab-Tek™ (Nunc-GmbH, Wiesbaden, Alemania) y se transfectaron con las construcciones de expresión indicadas usando el reactivo de transfección Polyplus jetPEI™ (BIOMOL GmbH, Hamburgo, Alemania) de acuerdo con las instrucciones del fabricante. Después de 4-6 horas el medio de transfección se cambió a medio de cultivo nuevo y las células se incubaron después durante otras 24 horas antes de realizar microscopía de células vivas o fijación con formaldehído 3,7% en PBS durante 10 minutos a temperatura ambiente. Las células fijadas se permeabilizaron con Triton X-100 0,2% en PBS durante 3 minutos, se contratiñeron con DAPI y se montaron en Vectashield (Vector Laboratories, CA, Estados Unidos).

G. Análisis de transferencia de western

Se separaron cantidades crecientes de proteína de extractos celulares totales de células 293T transfectadas con simulación o que expresaban la construcción cAb-GFP4-mRFP1 en un SDS-PAGE al 12% y después se transfirieron por electroforesis a membrana de nitrocelulosa (Bio-Rad Laboratories, CA, Estados Unidos). La membrana se bloqueó con leche al 3% en PBS y se incubó durante una noche a 4 °C con un anticuerpo policlonal de conejo anti-mRFP1. Después de lavar con PBS que contenía Tween-20 0,1%, las manchas de transferencia se incubaron con anticuerpo anti-IgG de conejo conjugado con peroxidasa de rábano rústico. Las bandas inmunorreactivas se visualizaron con el kit de detección de transferencia de western ECL plus (Amersham Biosciences, NJ, Estados Unidos). Como un control de carga, las membranas se volvieron a explorar con anticuerpo anti-PCNA.

H. Filtración en gel

Se sometieron extractos de células 293T que expresaban el cromocuerpo anti-GFP a análisis de filtración en gel. Brevemente, se homogeneizaron 1×10^7 células en 500 μ l de tampón de lisis (Tris/HCl 20 mM pH 7,5, NaCl 150 mM, EDTA 0,5 mM, PMSF 2 mM, NP40 0,5%). Después de una etapa de centrifugación (10 minutos, 20.000xg, 4 °C) se cargó el sobrenadante transparente en una columna de Superose-12 (Amersham Pharmacia Biotech, NJ, Estados Unidos) y se sometió a cromatografía a un caudal de 0,4 ml/minuto en tampón de columna (Tris/HCl 20 mM pH 7,5, NaCl 150 mM, EDTA 0,5 mM). Se analizaron fracciones (500 μ l cada una) por SDS-PAGE y las proteínas se tiñeron con azul brillante de Coomassie R-250 o se exploraron adicionalmente por transferencia de western seguido de incubación con un anticuerpo contra mRFP1 como se ha descrito anteriormente. Como patrones de calibración se usó albúmina de suero bovino (66 kDa), anhidrasa carbónica (29,5 kDa) y citocromo c (12,5 kDa).

I. Microscopía

Se analizaron células vivas o fijadas que expresaban proteínas fluorescentes usando un microscopio confocal Leica TCS SP2 AOBs equipado con un objetivo de inmersión en aceite 63x/1,4 NA Plan-Apochromat. Los fluoróforos se excitaron con un láser de diodo de 405 nm, un láser de Ar de 488 nm, un láser de estado sólido bombeado con diodo (DPSS) de 561 nm. Los grupos de imágenes confocales de células vivas o fijadas se registraron típicamente con un tamaño de marco de 512 x 512 píxeles, un tamaño de píxel de 70-160 nm, un tamaño de etapa z de 280 nm y el agujero de alfiler abierto a 1 unidad Airy. Después se realizó una proyección de intensidad máxima de varias secciones z medias usando ImageJ (versión 1.34, <http://rsb.info.nih.gov/ij/>). Para observación de células in vivo a largo plazo se montaron cubreobjetos de vidrio de 40 mm de diámetro en cámaras de células vivas FCS2 (Biotech, Butler, PA, Estados Unidos) y se mantuvieron a 37 °C. Se adquirieron secciones ópticas lumínicas con un microscopio de barrido láser confocal Zeiss LSM410 usando la línea de láser de Ar de 488 nm y la línea de láser de HeNe de 543 nm. Se tomaron tres secciones z medias a intervalos de 1 μ m y el agujero de alfiler se abrió a 2 unidades Airy a los intervalos de tiempo indicados. Las células se siguieron hasta 12 horas. La deriva del foco a lo largo del tiempo se compensó con un macro, que usa el reflejo en la interfaz del cubreobjetos con el medio como referencia. Después de la adquisición de imágenes, se realizó una proyección de las tres secciones z a partir de cada punto temporal. Para análisis de colocalización se usó el programa adicional de "colocalización-descubridor" (versión 1.1) para ImageJ escrito por C. Laummonerie.

Ejemplo 6: Ensayos de inmunoprecipitación usando un dominio VHH

Prácticamente cada laboratorio de biología molecular y celular ha generado o usado construcciones de fusión con GFP. En la actualidad las construcciones de fusión están disponibles para la mayoría, si no todas las proteínas conocidas. Las fusiones de GFP son el método elegido para estudiar la localización de proteínas en células vivas.

Los inventores han desarrollado ahora herramientas que ofrecen una posibilidad rápida y eficaz para estudiar qué proteínas o factores celulares se asocian con estas proteínas de fusión de GFP.

Para expresión recombinante, los inventores construyeron un plásmido de expresión bacteriana marcado con histidina (His₆) C terminal y purificaron el fragmento de anticuerpo recombinante soluble cAb-GFP4 de células de *E. coli* BL21. El fragmento antigénico pequeño cAb-GFP4 se expresa como una proteína soluble a un rendimiento de 5 mg/litro de cultivo de *E. coli* inducido. La Figura 6A muestra la purificación del cAb-GFP4-His₆ de 50 ml de cultivo bacteriano inducido por IPTG realizando cromatografía de afinidad inmovilizada (IMAC) usando una columna de resina de NiNTA. El cAb-GFP4 purificado se eluyó como un monómero de aproximadamente 15 kDa en análisis de filtración en gel (Figura 6 B).

Para ensayar la capacidad del fragmento de cAb-GFP4 purificado para reconocer y precipitar su GFP de antígeno en la combinación con diferentes construcciones de fusión (Figura 7 A), los inventores realizaron ensayos de inmunoprecipitación. El cAb-GFP4 se acopló con proteína A inmovilizada. Se ha descrito que el dominio variable de IgG humano del subgrupo III (V_HIII) se une a la proteína A de superantígeno bacteriano^{1a, 2a}. La proteína A reconoce un epítipo no lineal dentro del dominio V_H y se ha propuesto para purificación específica de proteína V_H plegada correctamente solamente^{3a}. Puesto que la cadena pesada variable (VHH) de *Camelidae* evolucionó dentro de la V_H subgrupo III, también se ha propuesto que los dominios VHH inducidos en *Camelidae* se unen a proteína A. De forma consecuente con esto, se ha indicado la purificación de fragmentos VHH expresados recombinantes usando proteína A agarosa^{4a}.

Para experimentos de inmunoprecipitación se añadió cAb-GFP4 purificado al extracto de células solubles de células 293T tratadas con simulación, que expresaban el dominio TS de Dnmt1 fusionado con GFP (Dnmt1 TS₂₅₆₋₆₂₉-GFP), GFP-PCNA o GFP solamente. Para precipitación se añadió proteína A agarosa, y se analizaron las proteínas precipitadas por tinción de Coomassie o análisis de inmunotransferencia.

Tanto la tinción de Coomassie como el análisis de inmunotransferencia muestran una precipitación eficaz de proteínas marcadas con GFP usando el fragmento de anticuerpo cAb-GFP4 acoplado a proteína A. La GFP puede estar en el extremo C terminal (Dnmt1 TS₂₅₆₋₆₂₉-GFP) o el extremo N terminal (GFP-PCNA) de la proteína de fusión. El análisis de inmunotransferencia revela además que las construcciones de fusión de GFP se precipitan de una manera casi cuantitativa. Esta precipitación es altamente específica ya que cAb-GFP4 acoplado a proteína A no precipita β-actina usada como un control. Los inventores concluyeron a partir de esto que el cAb-GFP4 expresado recombinante reconoce eficazmente y precipita su antígeno y por lo tanto es una herramienta valiosa para enfoques bioquímicos usando proteínas de fusión marcadas con GFP.

Además los inventores demostraron que el fragmento cAb-GFP4 también puede usarse para inmunoprecipitación conjunta. La Figura 8 muestra la precipitación conjunta de PCNA endógeno con un dominio que interacciona con PCNA descrito de Dnmt1 fusionado con GFP (GFP-Dnmt1₁₅₈₋₆₂₉).

En resumen, los inventores demuestran en este documento el uso eficaz de un fragmento de anticuerpo derivado de alpaca inducido contra GFP. Los inventores quieren enfatizar que este fragmento, ya que tiene un bajo peso molecular (15 kDa) es una nueva herramienta extraordinaria para analizar proteínas marcadas con GFP usando técnicas bioquímicas. Proporciona una serie de ventajas excepcionales, ya que puede expresarse en cantidades altas en *E. coli* y purificarse en una etapa. Como han podido mostrar los inventores, precipita su antígeno de una manera cuantitativa a partir de extractos celulares solubles. No hay preferencia por proteínas marcadas con GFP en el extremo N o C terminal, lo que es coherente con las observaciones *in vivo* de los inventores (véase anteriormente). A diferencia de los anticuerpos convencionales usados para inmunoprecipitación (mono o policlonal) el fragmento cAb-GFP4 solamente aparece como una banda única en el gel y no hay interferencia con la detección de proteínas por encima de 15 kDa. La expresión sencilla y eficaz de agentes de unión funcionales en *E. coli* y mamíferos hace posible además combinarlo con otros elementos funcionales. En particular, puede usarse una fusión de secuencia de biotilación para purificación y/o detección usando productos de estreptavidina disponibles en el mercado.

Ejemplo 7: Métodos usados en el Ejemplo 6

A. Expresión de GFP, purificación e inmunización de llamas

Se transformaron células *Escherichia coli* BL21 DE3 con pRSETB-GFPS65T^{5a} (proporcionadas amablemente por Roger Y. Tsien, UCSD) y se purificó GFPs65T marcado con (His₆) sobreexpresado usando cromatografía de afinidad metálica-iónica de acuerdo con las instrucciones del fabricante (Talon, Clontech, CA, Estados Unidos). Se inmunizó una llama alpaca (*Lama pacos*) con GFP purificada recombinante en adyuvante Gerbu de acuerdo con el siguiente esquema: día 0, 250 µg de GFP; día 7, 14, 21, 28 y 35, 100 µg de GFP; día 42, se recogió una muestra sanguínea de 150 ml.

B. Construcción de biblioteca de VHH y selección del agente de unión de GFP.

Se diluyó sangre heparinizada (36 ml) con RPMI precalentado y se estratificó en Lymphoprep™ (AXIS-Shield, Oslo, Noruega) para purificar las células PBL de acuerdo con las instrucciones del fabricante. Se aislaron un total de 2×10^7 PBL y se almacenaron a -80°C en alícuotas de 6×10^6 células. El ARNm se extrajo de 6×10^6 linfocitos y se sintetizó ADNc con transcriptasa inversa RNasaH- SuperscriptII (Invitrogen, CA, Estados Unidos) usando un cebador oligo-dT. La primera PCR en el molde de ADNc se realizó usando cebadores CALL001 (5'-GTC CTG GCT GCT CTT CTA CA AGG-3') y CALL002 (5'-GGT ACG TGC TGT TGA ACT GTT CC-3') que hibridan en la secuencia líder y en el exón CH2 de las cadenas pesadas de todas las IgG de llama, respectivamente. Los productos de PCR que carecían de las secuencias CH1 (es decir fragmentos con tamaños entre 650 y 750 pb) se purificaron de un gel de agarosa usando el kit de extracción en gel de PCR QIAquick (Qiagen-GmbH, Hilden, Alemania). Se realizó una PCR anidada con una mezcla equimolar de cebadores SM017 y SM018 (5'-CCA GCC GGC CAT GGC TCA GGT GCA GCT GGT GGA GTC TGG-3' y 5'-CCA GCC GGC CAT GGC TGA TGT GCA GCT GGT GGA GTC TGG-3', respectivamente) y cebador CALL002, y el producto de PCR se volvió a purificar a partir de gel de agarosa como se ha descrito anteriormente. Los genes de VHH finalmente se volvieron a amplificar con los cebadores A4 corto (5'-CAT GCC ATG ACT CGC GGC CAC GCC GGC CAT GGC-3') y 38 (5'-GGA CTA GTG CGG CCG CTG GAG ACG GTG ACC TGG GT-3') y se digirieron con enzimas de restricción NcoI y NotI para obtener extremos de ADN adhesivos. El fragmento se purificó con QIAquick, se ligó en el vector pHEN4^{4a} cortado con las mismas enzimas y la mezcla de ligación se usó para transformar células *Escherichia coli* TG1. Después de crecimiento durante una noche en placas de LB/ampicilina, las colonias bacterianas se pasaron por raspado de las placas a LB, se lavaron en el mismo medio y se almacenaron en LB/glicerol 15% a -80°C hasta su uso posterior. Se usó una alícuota representativa de esta biblioteca para inocular LB/ampicilina hasta que las células alcanzaron la fase de crecimiento exponencial antes de la infección con fagos auxiliares M13K07 para expresar el VHH clonado. La biblioteca de VHH presentado en fagos se sometió a selección con respecto a la presencia de agentes de unión de GFP en GFP revestida de fase sólida (0,1 μg de GFP / 100 μl por pocillo) durante tres ciclos consecutivos. Después del tercer ciclo de selección, se seleccionaron colonias individuales y se indujo expresión de su proteína periplásmica soluble con IPTG 1 mM. El VHH recombinante extraído del periplasma se ensayó con respecto a reconocimiento de antígenos en un ELISA.

C. Expresión y purificación del fragmento de anticuerpo de dominio único

El gen de VHH del clon que tuvo puntuación positiva en ELISA (cAb-GFP4) se volvió a clonar en el vector de expresión pHEN6 y se usó para transformar células de *Escherichia coli* BI21. La producción y purificación a gran escala siguió el protocolo descrito en van Koningbruggen *et al.*^{15a}.

D. Cultivo y transfección de células de mamífero

Se cultivaron células 293T en DMEM complementado con FCS 10%. Las células 293T se transfectaron con ADN plasmídico usando el reactivo TransFectin™ (Bio-Rad Laboratories, Hercules, CA, Estados Unidos) de acuerdo con las directrices del fabricante y se incubaron durante una noche, 48 horas o 72 horas respectivamente antes de realizar inmunoprecipitaciones.

E. Inmunoprecipitación

Se homogeneizaron 2×10^7 células 293T tratadas con simulación o que expresaban de forma transitoria Dnmt1 (310-629)-GFP, GFP-PCNA o GFP en 200 μl de tampón de lisis (Tris/HCl 20 mM pH 7,5, NaCl 150 mM, EDTA 0,5 mM, PMSF 2 mM, NP40 0,5%). Después de una etapa de centrifugación (10 minutos, 20.000xg, 4°C) el sobrenadante transparente se ajustó con tampón de dilución (Tris/HCl 20 mM pH 7,5, NaCl 150 mM, EDTA 0,5 mM, PMSF 2 mM) a 500 μl . Se añadieron 50 μl a tampón de muestra que contenía SDS (denominado entrada). Se añadieron 10 μl , respectivamente 20 μl de cAb-GFP4-His₆ (concentración 1 $\mu\text{g}/\mu\text{l}$) y se incubaron durante dos horas en un rotor vertical a 4°C en la habitación fría. Se lavaron 25 μl de perlas de proteína A agarosa (Amersham Pharmacia, NJ, Estados Unidos) 3 veces en 500 μl de tampón de dilución antes de la adición. La incubación continuó durante 1 hora. Después de una etapa de centrifugación (2 minutos, 5.000xg, 4°C) el sedimento de las perlas se resuspendió en tampón de dilución que contenía NaCl 300 mM. El sobrenadante se indicó como flujo continuo. La etapa de lavado se realizó dos veces incluyendo cambio de las cubetas. Se retiraron 50 μl (10%) denominados como lavado. Después de la última etapa de lavado, el sedimento de las perlas se incubó en tampón de muestra que contenía SDS 2x y se hirvió durante 10 minutos a 95°C .

F. Análisis de transferencia de western.

Después de hervir las perlas, los sobrenadantes solubles así como las fracciones de entrada, flujo continuo y lavado se separaron en un SDS-PAGE al 12% y después se tiñeron con Coomassie o se transfirieron por electroforesis a membrana de nitrocelulosa (Bio-Rad Laboratories, CA, Estados Unidos). La membrana se bloqueó con leche al 3% en PBS y se incubó durante una noche a 4°C con los anticuerpos correspondientes. Después de lavar con PBS que contenía Tween-20 0,1%, las manchas de transferencia se incubaron con el anticuerpo secundario apropiado conjugado con peroxidasa de rábano rusticano. Se visualizaron bandas inmunorreactivas con el Kit de Detección de

Transferencia de Western ECL plus (Amersham Biosciences, NJ, Estados Unidos).

G. Filtración en gel

5 Se sometieron 10 μ g del fragmento de anticuerpo cAb-GFP-His₆ purificado a análisis de filtración en gel. Se cargaron 100 μ l en una columna de Superose-12 (Amersham Pharmacia Biotech, NJ, Estados Unidos) y se sometió a cromatografía a un caudal de 0,4 ml/minuto en tampón de columna (Tris/HCl 20 mM pH 7,5, NaCl 150 mM, EDTA 0,5 mM). Las fracciones (500 μ l cada una) se analizaron por SDS-PAGE, se transfirieron y se exploraron con un anticuerpo contra un marcador de His6 C terminal (Invitrogen). Como patrones de calibración se usaron albúmina de suero bovino (66 kDa), anhidrasa carbónica (29,5 kDa) y citocromo c (12,5 kDa).

Ejemplo 9: Los cromocuerpos fluorescentes expresados *in vivo* se dirigen eficazmente a proteínas endógenas

15 Para investigar si los cromocuerpos expresados de forma intercelular también pueden reconocer epítomos endógenos los inventores generaron cromocuerpos contra antígenos citoplasmáticos y nucleares.

Se indujeron anticuerpos de cadena única contra proteínas nucleares y citoplasmáticas en camello. Se amplificaron los dominios VHH de dos agentes de unión fuertes contra citoqueratina-8 y laminina Dm0 por PCR con los cebadores VHH (BgIII)#F (5'-GGG GAG ATC TCC GGC CAT GGC TCA GGT GCA G-3') y gfp4#RC (5'-GGG GGA ATT CCT TGA GGA GAC GGT GAC-3'). El producto de PCR se purificó como se ha descrito y se digirió con las enzimas de restricción BgIII y EcoRI y se ligó en un vector pEYFP-N1 modificado (Clontech, CA, Estados Unidos) en el que la secuencia de YFP se había reemplazado por la región codificante de mRFP1 [31] (proporcionada amablemente por Roger Y. Tsien, UCSD).

25 Para analizar un antígeno citoplásmico los inventores fusionaron un VHH contra citoqueratina-8 con mRFP1 para generar un cromocuerpo anti-citoqueratina-8. Este cromocuerpo anti-citoqueratina-8 (SEQ ID NO: 20) está codificado por la secuencia de ácido nucleico de SEQ ID NO: 21. Se transfectaron mioblastos murinos con un vector de expresión que codificaba el cromocuerpo anti-citoqueratina-8. En un periodo de 12 a 24 horas fueron detectables estructuras filamentosas en el citoplasma de células C2C12 que expresaban el cromocuerpo anti-citoqueratina-8 por la fluorescencia roja. En consecuencia, las células se fijaron, se contratiñeron para el ADN con DAPI y se analizaron por microscopía confocal para obtener una mayor resolución (Figura 9A).

35 A pesar de un fondo de fluorescencia sustancial pudieron detectarse estructuras filamentosas únicas. El alto nivel de cromocuerpo dispersado podría estar provocado por una relación modificada entre antígeno y cromocuerpo. Como la citoqueratina-8 se expresa en mioblastos murinos, el nivel de expresión es más bien bajo, lo que da como resultado un exceso de cromocuerpos no unidos. No obstante, en comparación con la distribución dispersada de un cromocuerpo en ausencia de un antígeno (Figura 3D, panel superior) y también de mRFP1 solo (Figura 3D, panel inferior) la fluorescencia roja del cromocuerpo muestra un patrón acentuado en el citoplasma que refleja las estructuras filamentosas de citoqueratina-8.

40 Para analizar un epítipo nuclear endógeno los inventores generaron un cromocuerpo anti-laminina. Por lo tanto, los inventores fusionaron un fragmento de anticuerpo dirigido contra laminina con mRFP1. El cromocuerpo anti-laminina (SEQ ID NO: 22) está codificado por la secuencia de ácido nucleico de SEQ ID NO: 23. En este caso los inventores aprovecharon un fragmento de unión a anticuerpo de cadena única recientemente identificado inducido contra laminina Dm0 de *Drosophila*. Debido a la alta conservación los inventores anticiparon un posible reconocimiento del ortólogo de mamífero. Para estudios de localización se transfectaron células HeLa con el vector de expresión de cromocuerpo anti-laminina. Las células HeLa que expresaban el cromocuerpo anti-laminina se fijaron en un periodo de 12-24 horas después de la transfección, se tiñeron con DAPI y se sometieron a microscopía confocal (Figura 9B). Se muestra una célula representativa en la que está destacada la estructura de laminina típica, que rodea el núcleo, por la fluorescencia roja del cromocuerpo. Solo era observable un bajo fondo de cromocuerpo distribuido dispersado. Esto demuestra un reconocimiento específico alto del epítipo por el cromocuerpo anti-laminina que es comparable a la dirección de laminina B1-GFP por el cromocuerpo anti-GFP como se muestra en la Figura 4C.

55 Estos datos demuestran que los epítomos endógenos como la citoqueratina-8 localizada en el citoplasma así como la proteína laminina nuclear pueden visualizarse específica y eficazmente por cromocuerpos expresados de forma intracelular. Las proteínas celulares usadas en este estudio son solamente ejemplos, cualquier otra proteína, por ejemplo celular, puede detectarse del mismo modo usando cromocuerpos específicos (o fluorocuerpos). Resulta importante que los ejemplos demuestran que no debe esperarse limitación con respecto a la localización subcelular, en particular compartimentalización.

60 Ejemplo 10: Comparación de la expresión de cAbGFP4 y cAbGFP4 (Cys92 Ser) en células de mamífero

65 Se mutó cAbGFP4 para aumentar la solubilidad y para reducir la tendencia a agregación de la proteína debido a la formación de puentes disulfuro inespecíficos. Por lo tanto la cisteína en la posición 98 de SEQ ID NO: 18 se alteró a serina (SEQ ID NO: 32).

(A) Detección de la distribución intracelular

5 Ambas proteínas se expresaron en células 293T. Como se representa en la Figura 12 (panel superior), cAbGFP4 (Cys 92 Ser) está distribuido más uniformemente dentro de la célula humana en comparación con el cAbGFP4 original sin estar presentes antígenos. Resulta notable que el núcleo es accesible a un porcentaje mucho mayor de cAbGFP4 (Cys 92 Ser) en comparación con cAbGFP4. Esto indica una mayor solubilidad así como la ausencia de agregación inespecífica, puesto que una de las dos cisteínas presentes en la molécula se mutó para evitar la formación de puentes disulfuro no deseados.

10 (B) Detección de la funcionalidad en células 293T humanas

La Figura 12 (panel inferior) muestra claramente que en presencia de GFP-PCNA como antígeno en células humanas, cAbGFP4 (Cys92Ser) se une a su antígeno de forma más eficaz que la molécula cAbGFP4 original. La cantidad de molécula libre es mucho menor en el primer caso.

15 Ejemplo 11: Comparación de la capacidad de precipitación del agente de unión de GFP con anticuerpos de GFP mono y policlonales

20 Se incubaron extractos proteicos de células 293T productoras de GFP con 1 µg de agente de unión de GFP (cAbGFP4) o 2 µg de anticuerpos de GFP mono o policlonales y se precipitaron inmunocomplejos con proteína A/G sepharose. Se separaron alícuotas de fracción de entrada (I), flujo continuo (F) y unida (B) por SDS-PAGE. La Figura 10 ilustra los resultados después de tinción con Coomassie (parte superior) o análisis de inmunotransferencia usando un anticuerpo anti-GFP (parte inferior). La GFP precipitada, cadenas pesada (hc) y ligera (lc) desnaturizadas de las IgG y el agente de unión de GFP (cAbGFP4) están marcados por flechas. Se muestra claramente que la precipitación/purificación de la GFP de antígeno con el agente de unión a GFP es mucho más específica que con anticuerpos monoclonales o policlonales dirigidos a GFP. No son visibles otras bandas distintas de las correspondientes a GFP y el agente de unión a GFP en el gel.

30 Ejemplo 12: Purificación basada en columna de GFP

35 Se cargaron extractos proteicos de células 293T productoras de GFP en una columna que contenía agente de unión a GFP (cAbGFP4) acoplado covalentemente a Sepharose 10 (Amersham). Las proteínas unidas se eluyeron y se separaron las alícuotas de fracción de entrada (I), flujo continuo (F) y unido (B) por SDS-PAGE y se tiñeron con Coomassie. Como se representa en la Figura 11 la proteína de interés, en este caso GFP, podría enriquecerse y purificarse.

Ejemplo 13: Actividad y detección intracelular de agente de unión a GFP recombinante (cAbGFP4) acoplado a Cy5

40 El agente de unión a GFP recombinante (cAbGFP4) se purificó por cromatografía de afinidad inmovilizada (IMAC) y la concentración proteica se ajustó a 1 mg/ml. Se acopló 1 ml de la solución proteica al NHS-éster Cy5 fluorescente (Amersham) de acuerdo con las instrucciones del fabricante. La proteína acoplada se sometió a filtración en gel en una columna Superdex-75 (Amersham) intercambiando el tampón de acoplamiento (NaHCO₃ 0,1 M) a 1xPBS. Para inmunotinciones se usó el agente de unión de GFP-Cy5 purificado en concentración de 50 a 100 ng/ml. La Figura 13 muestra una inmunotinción de células HeLa que expresan GFPPCNA con un dominio VHH específico de GFP (cAbGPF4, agente de unión de GFP) acoplado con Cy5-NHS (Amersham). Las células se fijaron con paraformaldehído 3,7. Barra de escala 5 µm. La fusión de las dos señales que surgen de GFP-PCNA y la emisión de Cy5 de la fusión agente de unión de GFP-Cy5 en (a) y (b) muestra la colocalización de ambas proteínas (Figura 13(d)). Esto demuestra la idoneidad de colorantes fluorescentes para las aplicaciones de la presente invención.

50

Referencias

- 5 1. Biocca, S., Neuberger, M.S. y Cattaneo, A. Expression and targeting of intracellular antibodies in mammalian cells. *Embo J* 9, 101-108 (1990).
2. Cattaneo, A. y Biocca, S. The selection of intracellular antibodies. *Trends Biotechnol.* 17, 115-121 (1999).
- 10 3. Biocca, S., Pierandrei-Amaldi, P. y Cattaneo, A. Intracellular expression of anti-p21 ras single chain Fv fragments inhibits meiotic maturation of xenopus oocytes. *Biochem: Biophys. Res. Commun.* 197, 422-427 (1993).
4. Biocca, S., Neuberger, M.S. y Cattaneo, A. Expression and targeting of intracellular antibodies in mammalian cells. *Embo J.* 9, 101-108 (1990).
- 15 5. Marasco, W.A., Chen, S., Richardson, J.H., Ramstedt, U. y Jones, S.D. Intracellular antibodies against HIV-1 envelope protein for AIDS gene therapy. *Hum. Gene. Ther.* 9, 1627-1642 (1998).
- 20 6. Cardinale, A., Lener, M., Messina, S., Cattaneo, A. y Biocca, S. The mode of action of Y13-259 scFv fragment intracellularly expressed in mammalian cells. *FEBS Lett.* 439, 197-202 (1998).
7. Kontermann, R.E. Intrabodies as therapeutic agents. *Methods* 34, 163-170 (2004).
- 25 8. Hamers-Casterman, C. *et al.* Naturally occurring antibodies devoid of light chains. *Nature* 363, 446-448 (1993).
9. Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J.A. y Hamers, R. Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng.* 7, 1129-1135 (1994).
- 30 10. Sheriff, S. y Constantine, K.L. Redefining the minimal antigen-binding fragment. *Nat. Struct. Biol.* 3, 733-736 (1996).
11. Muyldermans, S. Single domain camel antibodies: current status. *J. Biotechnol.* 74, 277-302 (2001).
- 35 12. Arbabi Ghahroudi, M., Desmyter, A., Wyns, L., Hamers, R. y Muyldermans, S. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.* 414, 521-526 (1997).
13. van der Linden, R.H. *et al.* Comparison of physical chemical properties of llama VHH antibody fragments and mouse monoclonal antibodies. *Biochim. Biophys. Acta* 1431, 37-46 (1999).
- 40 14. Lauwereys, M. *et al.* Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *Embo J.* 17, 3512-3520 (1998).
- 45 15. van Koningsbruggen, S. *et al.* Llama-derived phage display antibodies in the dissection of the human disease oculopharyngeal muscular dystrophy. *J. Immunol. Methods* 279, 149-161 (2003).
16. Cortez-Retamozo, V. *et al.* Efficient tumor targeting by single-domain antibody fragments of camels. *Int. J. Cancer* 98, 456-462 (2002).
- 50 17. Jobling, S.A. *et al.* Immunomodulation of enzyme function in plants by single-domain antibody fragments. *Nat. Biotechnol.* 21, 77-80 (2003).
18. Kabat, E.A. y Wu, T.T. Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. *J. Immunol.* 147, 1709-1719 (1991).
- 55 19. Conrath, K.E., Wernery, U., Muyldermans, S. y Nguyen, V.K. Emergence and evolution of functional heavy chain antibodies in Camelidae. *Dev. Comp. Immunol.* 27, 87-103 (2003).
- 60 20. Westphal, M. *et al.* Microfilament dynamics during cell movement and chemotaxis monitored using a GFP-actin fusion protein. *Curr. Biol.* 7, 176-183 (1997).
- 65 21. Daigle, N. *et al.* Nuclear pore complexes form immobile networks and have a very low turnover in live mammalian cells. *J. Cell. Biol.* 154, 71-84 (2001).

22. Warbrick, E. The puzzle of PCNA's many partners. *Bioessays* 22, 997-1006 (2000).
23. Leonhardt, H. *et al.* Dynamics of DNA replication factories in living cells. *J. Cell. Biol.* 149, 271-280 (2000).
- 5 24. Sporbert, A., Gahl, A., Ankerhold, R., Leonhardt, H. y Cardoso, M.C. DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. *Mol. Cell* 10, 1355-1365 (2002).
- 10 25. Kanda, T., Sullivan, K.F. y Wahl, G.M. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr. Biol.* 8, 377-385 (1998).
26. Cardinale, A., Filesi, I., Mattei, S. y Biocca, S. Intracellular targeting and functional analysis of single-chain Fv fragments in mammalian cells. *Methods* 34, 171-178 (2004).
- 15 27. Saerens, D. *et al.* Identification of a Universal VHH Framework to Graft Non-canonical Antigen-binding Loops of Camel Single-domain Antibodies. *J. Mol. Biol.* In Press (2005).
28. Desmyter, A. *et al.* Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat. Struct. Biol.* 3, 803-811 (1996).
- 20 29. Riechmann, L. y Muyldermans, S. Single domain antibodies: comparison of camel VH and camelised human VH domains. *J. Immunol. Methods* 231, 25-38 (1999).
- 30 30. Heim, R., Cubitt, A.B. y Tsien, R.Y. Improved green fluorescence. *Nature* 373, 663-664 (1995).
31. Campbell, R.E. *et al.* A monomeric red fluorescent protein. *Proc. Natl Acad. Sci. U S A* 99, 7877-7882 (2002).
- 30 32. Spinelli, S. *et al.* Camelid heavy-chain variable domains provide efficient combining sites to haptens. *Biochemistry* 39, 1217-1222 (2000).
- 35 33. Spinelli, S. *et al.* Domain swapping of a llama VHH domain builds a crystal-wide beta-sheet structure. *FEBS Lett.* 564, 35-40 (2004).
34. Spinelli, S. *et al.* The crystal structure of a llama heavy chain variable domain. *Nat. Struct. Biol.* 3, 752-757 (1996).
- 40 35. Desmyter, A. *et al.* Three camelid VHH domains in complex with porcine pancreatic alpha-amylase. Inhibition and versatility of binding topology. *J. Biol. Chem.* 277, 23645-23650 (2002).

Referencias adicionales

- 45 1a. Sasso, E.H., Silverman, G.J. y Mannik, M. Human IgA and IgG F(ab')₂ that bind to staphylococcal protein A belong to the VHIII subgroup. *J Immunol* 147, 1877-1883 (1991).
- 50 2a. Riechmann, L. y Davies, J. Backbone assignment, secondary structure and protein A binding of an isolated, human antibody VH domain. *J Biomol NMR* 6, 141-152 (1995).
- 3a. Riechmann, L. y Muyldermans, S. Single domain antibodies: comparison of camel VH and camelised human VH domains. *J. Immunol. Methods* 231, 25-38 (1999).
- 55 4a. Arbabi Ghahroudi, M., Desmyter, A., Wyns, L., Hamers, R. y Muyldermans, S. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.* 414, 521-526 (1997).
- 5a. Heim, R., Cubitt, A.B. y Tsien, R.Y. Improved green fluorescence. *Nature* 373, 663-664 (1995).
- 60 6a. van Koningsbruggen, S. *et al.* Llama-derived phage display antibodies in the dissection of the human disease oculopharyngeal muscular dystrophy. *J. Immunol. Methods* 279, 149-161 (2003).

LISTADO DE SECUENCIAS

<110> Ludwig-Maximilians-Universität Manchen

5 <120> Dirección y localización de antígenos en células vivas

<130> L 2417 EP/1

<160> 33

10 <170> PatentIn versión 3.3

<210> 1

<211> 117

<212> PRT

15 <213> secuencia artificial

<220>

<223> dominio variable genérico

20 <220>

<221> característica misc

<222> (29)..(37)

<223> Xaa puede ser cualquier aminoácido de origen natural

25 <220>

<221> característica misc

<222> (52)..(68)

<223> Xaa puede ser cualquier aminoácido de origen natural

30 <220>

<221> característica misc

<222> (101)..(106)

<223> Xaa puede ser cualquier aminoácido de origen natural

35 <400> 1

Met	Ala	Asp	Val	Gln	Leu	Val	Glu	Ser	Gly	Gly	Ala	Leu	Val	Gln	Pro
1				5					10					15	
Gly	Gly	Ser	Leu	Arg	Leu	Ser	Cys	Ala	Ala	Ser	Gly	Xaa	Xaa	Xaa	Xaa
			20					25					30		
Xaa	Xaa	Xaa	Xaa	Xaa	Trp	Tyr	Arg	Gln	Ala	Pro	Gly	Lys	Glu	Arg	Glu
		35					40					45			
Trp	Val	Ala	Xaa												
	50					55					60				
Xaa	Xaa	Xaa	Xaa	Arg	Phe	Thr	Ile	Ser	Arg	Asp	Asp	Ala	Arg	Asn	Thr
65					70					75					80
Val	Tyr	Leu	Gln	Met	Asn	Ser	Leu	Lys	Pro	Glu	Asp	Thr	Ala	Val	Tyr
				85					90					95	
Tyr	Cys	Asn	Val	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Trp	Gly	Gln	Gly	Thr	Gln
			100						105				110		
				Val	Thr	Val	Ser	Ser							
									115						

ES 2 401 359 T3

<210> 2
 <211> 351
 <212> ADN
 <213> secuencia artificial
 5
 <220>
 <223> dominio variable genérico, una secuencia codificante de SEQ ID NO: 1
 <220>
 <221> característica misc
 <222> (85)..(111)
 <223> n es a, c, g o t
 10
 <220>
 <221> característica misc
 <222> (154)..(204)
 <223> n es a, c, g o t
 15
 <220>
 <221> característica misc
 <222> (301)..(318)
 <223> n es a, c, g o t
 20
 <400> 2
 25
 atggccgatg tgcagctggt ggagctctggg ggagccttgg tgcagccggg ggggtctctg 60
 agactctcct gtgcagcctc tggannnnnn nnnnnnnnnn nnnnnnnnnn ntggtaccgc 120
 caggctccag ggaaggagcg cgagtgggtc gcgnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
 nnnnnnnnnn nnnnnnnnnn nnnncgattc accatctcca gagacgacgc caggaatacg 240
 gtgtatctgc aaatgaacag cctgaaacct gaggacacgg ccgtgtatta ctgtaatgtc 300
 nnnnnnnnnn nnnnnnnntg gggccagggg acccagggtca ccgtctcctc a 351
 <210> 3
 <211> 9
 <212> PRT
 <213> secuencia artificial
 30
 <220>
 <223> CDR 1
 35
 <400> 3
 Phe Pro Val Asn Arg Tyr Ser Met Arg
 1 5
 <210> 4
 <211> 17
 <212> PRT
 <213> secuencia artificial
 40
 <220>
 <223> CDR 2
 45
 <400> 4

ES 2 401 359 T3

Gly Met Ser Ser Ala Gly Asp Arg Ser Ser Tyr Glu Asp Ser Val Lys
1 5 10 15

Gly

5 <210> 5
<211> 6
<212> PRT
<213> secuencia artificial

10 <220>
<223> CDR 3

<400> 5

Asn Val Gly Phe Glu Tyr
1 5

15 <210> 6
<211> 238
<212> PRT
<213> Aequorea victoria

20 <400> 6

ES 2 401 359 T3

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
35 40 45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
50 55 60

Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
65 70 75 80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85 90 95

Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100 105 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115 120 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
130 135 140

Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145 150 155 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180 185 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
210 215 220

Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225 230 235

<210> 7
<211> 717
<212> ADN
<213> Aequorea victoria
<400> 7

ES 2 401 359 T3

```

atgagtaaag gagaagaact tttcaactgga gttgtcccaa ttcttggtga attagatggt      60
gatgttaatg ggcacaaatt ttctgtcagt ggagaggggtg aaggatgatgc aacatacggg      120
aaacttacc c ttaaatttat ttgcactact ggaaaactac ctgttccatg gccaacactt      180
gtcactactt tctcttatgg tgttcaatgc ttttcaagat acccagatca tatgaaacag      240
catgactttt tcaagagtgc catgcccga ggttatgtac aggaaagaac tatatttttc      300
aaagatgacg ggaactacaa gacacgtgct gaagtcaagt ttgaaggtga tacccttggt      360
aatagaatcg agttaaaagg tattgathtt aaagaagatg gaaacattct tggacacaaa      420
ttggaataca actataactc acacaatgta tacatcatgg cagacaaaca aaagaatgga      480
atcaaagtta acttcaaaat tagacacaac attgaagatg gaagcgttca actagcagac      540
cattatcaac aaaatactcc aattggcgat ggccctgtcc ttttaccaga caaccattac      600
ctgtccacac aatctgccct ttcgaaagat cccaacgaaa agagagacca catggtcctt      660
cttgagtttg taacagctgc tgggattaca catggcatgg atgaactata caaataa      717

```

5

<210> 8
 <211> 225
 <212> PRT
 <213> Discosoma

10

<400> 8

ES 2 401 359 T3

Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
1 5 10 15

Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
20 25 30

Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val
35 40 45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
50 55 60

Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
65 70 75 80

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
85 90 95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
100 105 110

Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
130 135 140

Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
145 150 155 160

Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
165 170 175

Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
180 185 190

Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
195 200 205

Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe
210 215 220

Leu
225

<210> 9
<211> 678
<212> ADN
<213> Discosoma

5

<400> 9

atgaggtctt ccaagaatgt tatcaaggag ttcattgaggt ttaaggttcg catggaagga 60
 acggtcaatg ggcacgagtt tgaaatagaa ggcgaaggag aggggaggcc atacgaaggc 120
 cacaataccg taaagcttaa ggtaaccaag gggggacctt tgccatttgc ttgggatatt 180
 ttgtcaccac aatttcagta tggaagcaag gtatatgtca agcaccttgc cgacatacca 240
 gactataaaa agctgtcatt tcctgaagga tttaaattggg aaaggggtcat gaactttgaa 300
 gacgggtggcg tcgttactgt aaccaggat tccagtttgc aggatggctg tttcatctac 360
 aaggtcaagt tcattggcgt gaactttcct tccgatggac ctgttatgca aaagaagaca 420
 atgggctggg aagccagcac tgagcgtttg taccctcgtg atggcgtgtt gaaaggagag 480
 attcataagg ctctgaagct gaaagacggt ggtcattacc tagttgaatt caaaagtatt 540
 tacatggcaa agaagcctgt gcagctacca gggactact atgttgactc caaactggat 600
 ataacaagcc acaacgaaga ctatacaatc gttgagcagt atgaaagaac cgaggggacgc 660
 caccatctgt tcctttag 678

5 <210> 10
 <211> 117
 <212> PRT
 <213> secuencia artificial

10 <220>
 <223> Fragmento VHH de cromocuerpo fluorescente contra GFP

<400> 10

Met Ala Asp Val Gln Leu Val Glu Ser Gly Gly Ala Leu Val Gln Pro
 1 5 10 15

Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Val Asn
 20 25 30

Arg Tyr Ser Met Arg Trp Tyr Arg Gln Ala Pro Gly Lys Glu Arg Glu
 35 40 45

Trp Val Ala Gly Met Ser Ser Ala Gly Asp Arg Ser Ser Tyr Glu Asp
 50 55 60

Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asp Ala Arg Asn Thr
 65 70 75 80

Val Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr
 85 90 95

Tyr Cys Asn Val Asn Val Gly Phe Glu Tyr Trp Gly Gln Gly Thr Gln
 100 105 110

Val Thr Val Ser Ser
 115

15

ES 2 401 359 T3

<210> 11
 <211> 351
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> secuencia que codifica fragmento VHH de cromocuerpo fluorescente de SEQ ID NO: 10

<400> 11

10

```

atggccgatg tgcagctggt ggagtctggg ggagccttgg tgcagccggg ggggtctctg      60
agactctcct gtgcagcctc tggattcccc gtcaatcgct atagtatgag gtggtaccgc      120
caggctccag ggaaggagcg cgagtgggtc gcgggtatga gtagtgctgg tgatcgttca      180
agttatgaag actccgtgaa gggccgattc accatctcca gagacgacgc caggaatagc      240
gtgtatctgc aatgaacag cctgaaacct gaggacacgg ccgtgtatta ctgtaatgtc      300
aatgtgggct ttgagtactg gggccagggg acccaggtca ccgtctcctc a                351
    
```

<210> 12
 <211> 123
 <212> PRT
 <213> secuencia artificial

15

<220>
 <223> Fragmento VHH de cromocuerpo fluorescente contra laminina

20

<400> 12

```

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

Gly Glu Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Asn Phe Ala
          20                25      30

Ser Tyr Ser Met Thr Trp Val Arg Gln Ala Pro Gly Lys Glu Leu Asp
          35                40      45

Trp Val Ala Thr Ile Ser Phe Ser Gly Gly Ser Thr Tyr Tyr Asp Asp
 50                55                60

Ser Val Lys Gly Arg Ala Thr Ile Ser Arg Asp Asn Ala Lys Ser Thr
 65                70                75      80

Val Tyr Leu Gln Leu Asn Ser Leu Lys Thr Glu Asp Thr Gly Met Tyr
          85                90                95

Tyr Cys Thr Lys Gly Arg Val Ser Pro Pro Ile Pro Asp Gly Leu Ser
          100               105                110

Arg Gly Gln Gly Thr Gln Val Thr Val Ser Ser
          115                120
    
```

25

ES 2 401 359 T3

<210> 13
 <211> 369
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> secuencia que codifica fragmento VHH de cromocuerpo fluorescente de SEQ ID NO: 12

10

<400> 13
 atggctcagg tgcagctggt ggagtctggg ggaggcttgg tgcagcctgg ggagtctctc 60
 agactctcct gtgcagcctc tggattcaac ttcgcaagtt attccatgac ctgggtccgc 120
 caggctccag ggaaggaact cgactgggtc gcaacaataa gtttcagtgg tgggtccact 180
 tactatgatg actccgtgaa gggccgggcc acaatctcca gagacaacgc caagagcaca 240
 gtgtatctcc agttaaacag cctgaaaact gaggacacgg gaatgtatta ctgtacaaaa 300
 ggcagagtat ccccgcccat tcccgatggc ctttctcgag gccaggggac ccaggtcacc 360
 gtctcctca 369

15

<210> 14
 <211> 128
 <212> PRT
 <213> secuencia artificial

20

<220>
 <223> Fragmento VHH de cromocuerpo fluorescente contra citoqueratina-8
 <400> 14

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

25

ES 2 401 359 T3

5
<210> 15
<211> 384
<212> ADN
<213> secuencia artificial

<220>
<223> secuencia que codifica fragmento VHH de cromocuerpo fluorescente de SEQ ID NO: 14

10
<400> 15

```
atggctcagg tacagctgca ggagtctgga ggaggcttgg tgcaggcagg ggggtctctg      60
acactctcct gtacatactc tggcttgact ttgatgatt ataacatggg ctggttccgc      120
cagggccag ggaaggagcg cgagagagtt tcggcgatta gctttcgtgg catcacctat      180
tatgtagact ccgtaaggg ccgattcacc atctccaggg ataacgcaa gaagacgctg      240
tatctgcaaa tgaacggcct gacacctgac gatacggcca cgtattactg tgcgggtagt      300
cgcttcttaa gtcctttgt acgcgatggt gacacgaaat tgattaacga ttggggccag      360
gggaccagcag tcaccgtctc ctca      384
```

15
<210> 16
<211> 229
<212> PRT
<213> secuencia artificial

<220>
<223> mRFP1

20
<400> 16

ES 2 401 359 T3

Met Ala Ser Ser Glu Asp Val Ile Lys Glu Phe Met Arg Phe Lys Val
 1 5 10 15

Arg Met Glu Gly Ser Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
 20 25 30

Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr Ala Lys Leu Lys Val
 35 40 45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
 50 55 60

Phe Gln Tyr Gly Ser Lys Ala Tyr Val Lys His Pro Ala Asp Ile Pro
 65 70 75 80

Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
 85 90 95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
 100 105 110

Leu Gln Asp Gly Glu Phe Ile Tyr Lys Val Lys Leu Arg Gly Thr Asn
 115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
 130 135 140

Ala Ser Thr Glu Arg Met Tyr Pro Glu Asp Gly Ala Leu Lys Gly Glu
 145 150 155 160

Ile Lys Met Arg Leu Lys Leu Lys Asp Gly Gly His Tyr Asp Ala Glu
 165 170 175

Val Lys Thr Thr Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Ala
 180 185 190

Tyr Lys Thr Asp Ile Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
 195 200 205

Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly Arg His Ser Thr Gly
 210 215 220

Ala Gly Leu Tyr Lys
 225

ES 2 401 359 T3

<210> 17
 <211> 690
 <212> ADN
 <213> secuencia artificial
 5 <220>
 <223> mRFP1
 <400> 17

 atggcctcct ccgaggacgt catcaaggag ttcattgcgct tcaagggtgcg catggagggc 60
 tccgtgaacg gccacgagtt cgagatcgag ggcgagggcg agggccgccc ctacgagggc 120
 acccagaccg ccaagctgaa ggtgaccaag ggcggccccc tgccttcgc ctgggacatc 180
 ctgtcccctc agttccagta cggctccaag gcctacgtga agcaccocgc cgacatcccc 240
 gactacttga agctgtcctt ccccgagggc ttcaagtggg agcgcgtgat gaacttcgag 300
 gacggcggcg tggtgaccgt gaccaggac tcctccctgc aggacggcga gttcatctac 360

 aagggtgaagc tgcgcggcac caacttcccc tccgacggcc ccgtaatgca gaagaagacc 420
 atgggctggg aggcctccac cgagcggatg taccocgagg acggcgcocct gaagggcgag 480
 atcaagatga ggctgaagct gaaggacggc ggccactacg acgccgaggt caagaccacc 540
 tacatggcca agaagcccgt gcagctgccc ggcgcctaca agaccgacat caagctggac 600
 atcacctccc acaacgagga ctacaccatc gtggaacagt acgagcgcgc cgagggccgc 660
 10 cactccaccg gcgcccggcct gtacaagtaa 690

 <210> 18
 <211> 368
 <212> PRT
 15 <213> secuencia artificial

 <220>
 <223> cromocuerpo anti-GFP
 20 <400> 18

ES 2 401 359 T3

Met Ala Asp Val Gln Leu Val Glu Ser Gly Gly Ala Leu Val Gln Pro
 1 5 10 15

Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Val Asn
 20 25 30

Arg Tyr Ser Met Gly Trp Tyr Arg Gln Ala Pro Gly Lys Glu Arg Glu
 35 40 45

Trp Val Ala Gly Met Ser Ser Ala Gly Asp Arg Ser Ser Tyr Glu Asp
 50 55 60

Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asp Ala Arg Asn Thr
 65 70 75 80

Val Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr
 85 90 95

Tyr Cys Asn Val Asn Val Gly Phe Glu Tyr Trp Gly Gln Gly Thr Gln
 100 105 110

Val Thr Val Ser Ser Arg Asn Ser Ile Ser Ser Leu Ser Ile Pro Ser
 115 120 125

Thr Val Pro Arg Ala Arg Asp Pro Pro Val Asp Met Ala Ser Ser Glu
 130 135 140

Asp Val Ile Lys Glu Phe Met Arg Phe Lys Val Arg Met Glu Gly Ser
 145 150 155 160

Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro

ES 2 401 359 T3

				165						170						175
Tyr	Glu	Gly	Thr	Gln	Thr	Ala	Lys	Leu	Lys	Val	Thr	Lys	Gly	Gly	Pro	
			180					185					190			
Leu	Pro	Phe	Ala	Trp	Asp	Ile	Leu	Ser	Pro	Gln	Phe	Gln	Tyr	Gly	Ser	
		195					200					205				
Lys	Ala	Tyr	Val	Lys	His	Pro	Ala	Asp	Ile	Pro	Asp	Tyr	Leu	Lys	Leu	
	210					215					220					
Ser	Phe	Pro	Glu	Gly	Phe	Lys	Trp	Glu	Arg	Val	Met	Asn	Phe	Glu	Asp	
225					230					235					240	
Gly	Gly	Val	Val	Thr	Val	Thr	Gln	Asp	Ser	Ser	Leu	Gln	Asp	Gly	Glu	
				245					250					255		
Phe	Ile	Tyr	Lys	Val	Lys	Leu	Arg	Gly	Thr	Asn	Phe	Pro	Ser	Asp	Gly	
			260					265					270			
Pro	Val	Met	Gln	Lys	Lys	Thr	Met	Gly	Trp	Glu	Ala	Ser	Thr	Glu	Arg	
		275					280					285				
Met	Tyr	Pro	Glu	Asp	Gly	Ala	Leu	Lys	Gly	Glu	Ile	Lys	Met	Arg	Leu	
	290					295					300					
Lys	Leu	Lys	Asp	Gly	Gly	His	Tyr	Asp	Ala	Glu	Val	Lys	Thr	Thr	Tyr	
305					310					315					320	
Met	Ala	Lys	Lys	Pro	Val	Gln	Leu	Pro	Gly	Ala	Tyr	Lys	Thr	Asp	Ile	
				325					330					335		
Lys	Leu	Asp	Ile	Thr	Ser	His	Asn	Glu	Asp	Tyr	Thr	Ile	Val	Glu	Gln	
			340					345					350			
Tyr	Glu	Arg	Ala	Glu	Gly	Arg	His	Ser	Thr	Gly	Ala	Gly	Leu	Tyr	Lys	
		355					360					365				

<210> 19
 <211> 1107
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> cromocuerpo anti-GFP

10

<400> 19

atggccgatg tgcagctggt ggagctctggg ggagccttgg tgcagccggg ggggtctctg 60
 agactctcct gtgcagcctc tggattcccc gtcaatcgct atagtatggg gtgggtaccgc 120

ES 2 401 359 T3

caggctccag ggaaggagcg cgagtgggtc gcggtatga gtagtgctgg tgatcgttca 180
 agttatgaag actccgtgaa gggccgattc accatctcca gagacgacgc caggaatacg 240
 gtgtatctgc aatgaacag cctgaaacct gaggacacgg ccgtgtatta ctgtaatgtc 300
 aatgtgggct ttgagtactg gggccagggg acccaggtca ccgtctcttc aaggaattcg 360
 atatcaagct tatcgatacc gtcgacggta ccgcgggccc gggatccacc ggtcgacatg 420
 gcctcctccg aggacgtcat caaggagttc atgcgcttca aggtgcgcat ggagggctcc 480
 gtgaacggcc acgagttcga gatcgagggc gagggcgagg gccgccccta cgagggcacc 540
 cagaccgcca agctgaaggt gaccaagggc gccccctgc ccttcgectg ggacatcctg 600
 tcccctcagt tccagtacgg ctccaaggcc tacgtgaagc accccgccga catccccgac 660
 tacttgaagc tgccttccc cgagggttc aagtgggagc gcgtgatgaa cttcgaggac 720
 ggcggcgtgg tgaccgtgac ccaggactcc tccctgcagg acggcgagtt catctacaag 780
 gtgaagctgc gcggcaccaa cttcccctcc gacggccccg taatgcagaa gaagaccatg 840
 ggctgggagg cctccaccga gcggatgtac cccgaggacg gcgccctgaa gggcgagatc 900
 aagatgaggc tgaagctgaa ggacggcggc cactacgacg ccgaggtcaa gaccacctac 960
 atggccaaga agcccgtgca gctgcccggc gcctacaaga ccgacatcaa gctggacatc 1020
 acctcccaca acgaggacta caccatcgtg gaacagtacg agcgcgccga gggccgccac 1080
 tccaccggcg ccggcctgta caagtaa 1107

<210> 20
 <211> 374
 <212> PRT
 <213> secuencia artificial

<220>
 <223> cromocuerpo anti-citoqueratina-8

<400> 20

Met Ala Gln Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro
 1 5 10 15
 Gly Glu Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Asn Phe Ala
 20 25 30
 Ser Tyr Ser Met Thr Trp Val Arg Gln Ala Pro Gly Lys Glu Leu Asp
 35 40 45
 Trp Val Ala Thr Ile Ser Phe Ser Gly Gly Ser Thr Tyr Tyr Asp Asp
 50 55 60
 Ser Val Lys Gly Arg Ala Thr Ile Ser Arg Asp Asn Ala Lys Ser Thr
 65 70 75 80

ES 2 401 359 T3

Val Tyr Leu Gln Leu Asn Ser Leu Lys Thr Glu Asp Thr Gly Met Tyr
85 90 95

Tyr Cys Thr Lys Gly Arg Val Ser Pro Pro Ile Pro Asp Gly Leu Ser
100 105 110

Arg Gly Gln Gly Thr Gln Val Thr Val Ser Ser Arg Asn Ser Ile Ser
115 120 125

Ser Leu Ser Ile Pro Ser Thr Val Pro Arg Ala Arg Asp Pro Pro Val
130 135 140

Asp Met Ala Ser Ser Glu Asp Val Ile Lys Glu Phe Met Arg Phe Lys
145 150 155 160

Val Arg Met Glu Gly Ser Val Asn Gly His Glu Phe Glu Ile Glu Gly
165 170 175

Glu Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr Ala Lys Leu Lys
180 185 190

Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro
195 200 205

Gln Phe Gln Tyr Gly Ser Lys Ala Tyr Val Lys His Pro Ala Asp Ile
210 215 220

Pro Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg
225 230 235 240

Val Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser
245 250 255

Ser Leu Gln Asp Gly Glu Phe Ile Tyr Lys Val Lys Leu Arg Gly Thr
260 265 270

Asn Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp
275 280 285

Glu Ala Ser Thr Glu Arg Met Tyr Pro Glu Asp Gly Ala Leu Lys Gly
290 295 300

Glu Ile Lys Met Arg Leu Lys Leu Lys Asp Gly Gly His Tyr Asp Ala
305 310 315 320

Glu Val Lys Thr Thr Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly
325 330 335

ES 2 401 359 T3

Ala Tyr Lys Thr Asp Ile Lys Leu Asp Ile Thr Ser His Asn Glu Asp
 340 345 350

Tyr Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly Arg His Ser Thr
 355 360 365

Gly Ala Gly Leu Tyr Lys
 370

5 <210> 21
 <211> 1125
 <212> ADN
 <213> secuencia artificial

10 <220>
 <223> cromocuerpo anti-citoqueratina-8

<400> 21

```

atggctcagg tgcagctggt ggagtctggg ggaggcttgg tgcagcctgg ggagtctctc      60
agactctcct gtgcagcctc tggattcaac ttcgcaagtt attccatgac ctgggtccgc      120
caggctccag ggaaggaact cgactgggtc gcaacaataa gtttcagtgg tgggtccact      180
tactatgatg actccgtgaa gggccggggc acaatctcca gagacaacgc caagagcaca      240
gtgtatctcc agttaaacag cctgaaaact gaggacacgg gaatgtatta ctgtacaaaa      300
ggcagagtat ccccgcccat tcccgatggc ctttctcgag gccaggggac ccaggtcacc      360
gtctcctcaa ggaattcgat atcaagctta tcgataccgt cgacgggtacc gcgggcccgg      420
gatccaccgg tcgacatggc ctccctccgag gacgtcatca aggagttcat gcgcttcaag      480
gtgcgcatgg agggctccgt gaacggccac gagttcgaga tcgagggcga gggcgagggc      540
cgcccctacg agggcaccca gaccgccaag ctgaagggtga ccaagggcgg ccccctgccc      600
ttcgctggg acatcctgtc ccctcagttc cagtacggct ccaaggccta cgtgaagcac      660
cccgccgaca tccccgacta cttgaagctg tccttccccg agggcttcaa gtgggagcgc      720
gtgatgaact tcgaggacgg cggcgtggtg accgtgacct aggactcctc cctgcaggac      780
ggcgagttca tctacaaggt gaagctgcgc ggcaccaact tcccctccga cgccccgta      840
atgcagaaga agaccatggg ctgggaggcc tccaccgagc ggatgtacct cgaggacggc      900
gccctgaagg gcgagatcaa gatgaggctg aagctgaagg acggcggcca ctacgacgcc      960
gaggtcaaga ccacctacat ggccaagaag cccgtgcagc tgcccggcgc ctacaagacc     1020
gacatcaagc tggacatcac ctcccacaac gaggactaca ccatcgtgga acagtacgag     1080
cgcgcccagg gccgccaact caccggcgcc ggctgtaca agtaa                          1125
    
```

15 <210> 22
 <211> 379
 <212> PRT
 <213> secuencia artificial

20 <220>

ES 2 401 359 T3

<223>cromocuerpo anti-laminina-8

<400> 22

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

ES 2 401 359 T3

225		230		235		240
Phe Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val Val Thr		245		250		255
Val Thr Gln Asp Ser Ser Leu Gln Asp Gly Glu Phe Ile Tyr Lys Val		260		265		270
Lys Leu Arg Gly Thr Asn Phe Pro Ser Asp Gly Pro Val Met Gln Lys		275		280		285
Lys Thr Met Gly Trp Glu Ala Ser Thr Glu Arg Met Tyr Pro Glu Asp		290		295		300
Gly Ala Leu Lys Gly Glu Ile Lys Met Arg Leu Lys Leu Lys Asp Gly		305		310		315
Gly His Tyr Asp Ala Glu Val Lys Thr Thr Tyr Met Ala Lys Lys Pro		325		330		335
Val Gln Leu Pro Gly Ala Tyr Lys Thr Asp Ile Lys Leu Asp Ile Thr		340		345		350
Ser His Asn Glu Asp Tyr Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu		355		360		365
Gly Arg His Ser Thr Gly Ala Gly Leu Tyr Lys		370		375		

<210> 23
 <211> 1140
 <212> ADN
 <213> secuencia artificial

<220>
 <223> cromocuerpo anti-laminina-8

<400> 23

atgcccagg tgcagctgca ggagtctgga ggaggcttgg tgcaggcagg ggggtctctg	60
acactctcct gtacatactc tggcttgact tttgatgatt ataacatggg ctggttccgc	120
caggggccag ggaaggagcg cgagagagtt tcggcgatta gctttcgtgg catcacctat	180
tatgtagact ccgttaaggg ccgattcacc atctccaggg ataacgcca gaagacgctg	240
tatctgcaaa tgaacggcct gacacctgac gatacggcca cgtattactg tgcgggtagt	300
cgcttcttaa gtcccttgt acgcgatggt gacacgaaat tgattaacga ttggggccag	360
gggaccagg tcaccgtctc ctcaaggaat tcgatatcaa gcttatcgat accgtcgacg	420
gtaccggggg cccgggatcc accggtcgac atggcctcct ccgaggacgt catcaaggag	480

ES 2 401 359 T3

ttcatgcgct tcaaggtgcg catggagggc tccgtgaacg gccacgagtt cgagatcgag 540
 ggcgagggcg agggccgccc ctacgagggc acccagaccg ccaagctgaa ggtgaccaag 600
 ggcggccccc tgcccttcgc ctgggacatc ctgtcccctc agttccagta cggctccaag 660
 gcctacgtga agcaccocgc cgacatcccc gactacttga agctgtcctt ccccgagggc 720
 ttcaagtggg agcgcgtgat gaacttcgag gacggcggcg tggtgaccgt gaccaggac 780
 tcctccctgc aggacggcga gttcatctac aagggtgaagc tgcgcggcac caacttcccc 840
 tccgacggcc ccgtaatgca gaagaagacc atgggctggg aggcctccac cgagcggatg 900
 taccocgagg acggcgcctt gaagggcgag atcaagatga ggctgaagct gaaggacggc 960
 ggccactacg acgccgaggt caagaccacc tacatggcca agaagcccgt gcagctgccc 1020
 ggcgcctaca agaccgacat caagctggac atcacctccc acaacgagga ctacaccatc 1080
 gtggaacagt acgagcgcgc cgagggccgc cactccaccg gcgccggcct gtacaagtaa 1140

5 <210> 24
 <211> 702
 <212> ADN
 <213> secuencia artificial

10 <220>
 <223> mRFPmars
 <400> 24

atggcatcat cagaagatgt tattaagaa tttatgagat tcaaagttaa aatggaaggt 60
 tcagttaacg gtcatgaatt tgaaattgaa ggtgaaggtg aaggtagacc atatgaaggt 120
 acacaaacag caaaattaaa agttacaaaa ggtggtccat taccatttgc atgggatatt 180
 ctttcaccac aatttcaata tggtagtaaa gcatatgtta aacatccagc agatattcca 240
 gattatctta aactttcatt tcagaaggt ttaaattggg aaagaattat gaattttgaa 300
 gatggtggtg ttgtaacagt tactcaagat tcaacattac aagatggtga atttatttat 360
 aaagttaaat taagaggtag caattttcca tcagatggtc cagtaatgca aaaaaaact 420
 atgggttggg aagcatcaac agaaagaatg tatccagaag atggtgcact aaaaggtgaa 480
 attaaaatga gattaaaact taaggatggt ggtcattatg atgcagaagt taaaacaact 540
 tataaagcta aaaaacaagt tcaattacca ggtgcttata aaacagatat taaattagat 600
 attatttcac ataatgaaga ttatactatt gttgaacaat atgaaagagc agaaggtaga 660
 cattcaacag gtgcaggatc ccatcatcac caccatcatt aa 702

15 <210> 25
 <211> 711
 <212> ADN
 <213> secuencia artificial

20 <220>
 <223> mCherry
 <400> 25

ES 2 401 359 T3

atggtgagca agggcgagga ggataacatg gccatcatca aggagtcat gcgcttcaag 60
 gtgcacatgg agggctccgt gaacggccac gagttcgaga tcgagggcga gggcgagggc 120
 cgccccacg agggcaccca gaccgccaag ctgaagggtga ccaaggggtg ccccctgccc 180
 ttcgcctggg acatcctgtc ccctcagttc atgtacggct ccaaggccta cgtgaagcac 240
 cccgccgaca tccccgacta cttgaagctg tccttccccg agggcttcaa gtgggagcgc 300
 gtgatgaact tcgaggacgg cggcgtggtg accgtgaccc aggactcctc cctgcaggac 360
 ggcgagttca tctacaaggt gaagctgcgc ggcaccaact tcccctccga cggccccgta 420
 atgcagaaga agaccatggg ctgggaggcc tcctccgagc ggatgtaccc cgaggacggc 480
 gccctgaagg gcgagatcaa gcagaggctg aagctgaagg acggcggcca ctacgacgct 540
 gaggatcaaga ccacctaca ggccaagaag cccgtgcagc tgcccggcgc ctacaacgctc 600
 aacatcaagt tggacatcac ctcccacaac gaggactaca ccatcgtgga acagtacgaa 660
 cgcgccgagg gccgccactc caccggcggc atggacgagc tgtacaagta a 711

<210> 26
 <211> 681
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> MRaspberry

10

<400> 26

atggtgagca agggcgagga ggtcatcaag gagttcatgc gcttcaaggt gcgcatggag 60
 ggctccgtga acggccacga gttcgagatc gagggcgagg gcgagggccg cccctacgag 120
 ggcacccaga ccgccaagct gaaggtgacc aagggtggcc ccctgccctt cgcctgggac 180
 atcctgtccc ctcagtgcat gtacggctcc aagggctacg tgaagcacc cgccgacatc 240
 cccgactact tgaagctgtc cttccccgag ggcttcaagt gggagcgcgt gatgaacttc 300
 gaggacggcg gcgtggtgac cgtgacccag gactcctccc tgcaggacgg cgagttcatc 360
 tacaaggtga agctgcgcgg caccaactc ccctccgacg gccccgtaat gcagaagaag 420
 accatgggct gggaggcctc ctccgagcgg atgtaccccc aggacggcgc cctgaagggc 480
 gagatgaaga tgaggctgaa gctgaaggac ggcggccact acgacgccga ggtcaagacc 540
 acctacatgg ccaagaagcc cgtgcagctg cccggcgcct acaagaccga catcaagctg 600
 gacatcacct cccacaacga ggactacacc atcgtggaac agtacgagcg cgccgagggc 660
 cgccactcca ccggcgccta a 681

15

ES 2 401 359 T3

<210> 27
 <211> 681
 <212> ADN
 <213> secuencia artificial
 5
 <220>
 <223> mPlum
 10
 <400> 27
 atggtgagca agggcgagga ggtcatcaag gagttcatgc gcttcaagga gcacatggag 60
 ggctccgtga acggccacga gttcgagatc gagggcgagg gcgagggccg cccctacgag 120
 ggcacccaga ccgccaggct gaaggtgacc aaggggtggcc ccctgccctt cgctgggac 180
 atcctgtccc ctcatgatc gtacggctcc aaggcctacg tgaagcacc cgccgacatc 240
 cccgactact tgaagctgtc cttccccgag ggcttcaagt gggagcgcgt gatgaacttc 300
 gaggacggcg gcgtggtgac cgtgaccacg gactcctccc tgcaggacgg cgagttcatc 360
 tacaaggtga aggtgcgcgg caccaacttc ccctccgacg gccccgtaat gcagaagaag 420
 accatgggct gggaggcctc ctccgagcgg atgtaccccg aggacggcgc cctgaagggc 480
 gagatgaaga tgaggctgag gctgaaggac ggcggccact acgacgccga ggtcaagacc 540
 acctacatgg ccaagaagcc cgtgcagctg cccggcgcct acaagaccga catcaagctg 600
 gacatcacct cccacaacga ggactacacc atcgtggaac agtacgagcg cgccgagggc 660
 cgccactcca ccggcgccta a 681
 15
 <210> 28
 <211> 117
 <212> PRT
 <213> secuencia artificial
 20
 <220>
 <223> dominio variable genérico
 25
 <220>
 <221> característica misc
 <222> (29)..(37)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 30
 <220>
 <221> característica misc
 <222> (52)..(68)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 35
 <220>
 <221> característica misc
 <222> (101)..(106)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 <400> 28

ES 2 401 359 T3

Met Ala Asp Val Gln Leu Val Glu Ser Gly Gly Ala Leu Val Gln Pro
 1 5 10 15

Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Xaa Xaa Xaa Xaa
 20 25 30

Xaa Xaa Xaa Xaa Xaa Trp Tyr Arg Gln Ala Pro Gly Lys Glu Arg Glu
 35 40 45

Trp Val Ala Xaa
 50 55 60

Xaa Xaa Xaa Xaa Arg Phe Thr Ile Ser Arg Asp Asp Ala Arg Asn Thr
 65 70 75 80

Val Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr
 85 90 95

Tyr Ser Asn Val Xaa Xaa Xaa Xaa Xaa Xaa Trp Gly Gln Gly Thr Gln
 100 105 110

Val Thr Val Ser Ser
 115

5

10

15

20

25

<210> 29
 <211> 351
 <212> ADN
 <213> secuencia artificial

 <220>
 <223> dominio variable genérico, una secuencia codificante de SEQ ID NO: 28

 <220>
 <221> característica misc
 <222> (85)..(111)
 <223> n es a, c, g o t

 <220>
 <221> característica misc
 <222> (154)..(204)
 <223> n es a, c, g o t

 <220>
 <221> característica misc
 <222> (301)..(318)
 <223> n es a, c, g o t

 <400> 29

ES 2 401 359 T3

```

atggccgátg tgcagctggt ggagctctggg ggagccttgg tgcagccggg ggggtctctg      60
agactctcct gtgcagcctc tggannnnnn nnnnnnnnnn nnnnnnnnnn ntggtaccgc      120
caggctccag ggaaggagcg cgagtggggc gcnnnnnnnn nnnnnnnnnn nnnnnnnnnn      180
nnnnnnnnnn nnnnnnnnnn nnnncgattc accatctcca gagacgacgc caggaatacg      240
gtgtatctgc aatgaacag cctgaaacct gaggacacgg ccgtgtatta ctggaatgtc      300
nnnnnnnnnn nnnnnnnntg gggccagggg acccaggtca ccgtctcctc a              351
    
```

5 <210> 30
 <211> 117
 <212> PRT
 <213> secuencia artificial

10 <220>
 <223> fragmento VHH de cromocuerpo fluorescente contra GFP (mutante C92S)

```

<400> 30
Met Ala Asp Val Gln Leu Val Glu Ser Gly Gly Ala Leu Val Gln Pro
 1          5          10          15
Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Val Asn
          20          25          30
Arg Tyr Ser Met Arg Trp Tyr Arg Gln Ala Pro Gly Lys Glu Arg Glu
          35          40          45
Trp Val Ala Gly Met Ser Ser Ala Gly Asp Arg Ser Ser Tyr Glu Asp
 50          55          60
Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asp Ala Arg Asn Thr
65          70          75          80
Val Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr
          85          90          95
Tyr Ser Asn Val Asn Val Gly Phe Glu Tyr Trp Gly Gln Gly Thr Gln
          100          105          110
Val Thr Val Ser Ser
          115
    
```

15 <210> 31
 <211> 351
 <212> ADN
 <213> secuencia artificial
 <220>
 20 <223> secuencia que codifica fragmento VHH de cromocuerpo fluorescente de SEQ ID NO: 30
 <400> 31

ES 2 401 359 T3

```
atggccgatg tgcagctggt ggagtcctggg ggagccttgg tgcagccggg ggggtctctg    60
agactctcct gtgcagcctc tggattcccc gtcaatcgct atagtatgag gtggtaccgc    120
caggctccag ggaaggagcg cgagtgggtc gcgggtatga gtagtgctgg tgatcgttca    180
agttatgaag actccgtgaa gggccgattc accatctcca gagacgacgc caggaatagc    240
gtgtatctgc aatgaacag cctgaaacct gaggacacgg ccgtgtatta ctgcaatgtc    300
aatgtgggct ttgagtactg gggccagggg acccaggtca ccgtctctc a                351
```

<210> 32
<211> 368
<212> PRT
<213> secuencia artificial

5

<220>
<223> cromocuerpo anti-GFP (mutante C92S)

10

<400> 32

ES 2 401 359 T3

Met Ala Asp Val Gln Leu Val Glu Ser Gly Gly Ala Leu Val Gln Pro
1 5 10 15

Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Pro Val Asn
20 25 30

Arg Tyr Ser Met Arg Trp Tyr Arg Gln Ala Pro Gly Lys Glu Arg Glu
35 40 45

Trp Val Ala Gly Met Ser Ser Ala Gly Asp Arg Ser Ser Tyr Glu Asp
50 55 60

Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asp Ala Arg Asn Thr
65 70 75 80

Val Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr
85 90 95

Tyr Ser Asn Val Asn Val Gly Phe Glu Tyr Trp Gly Gln Gly Thr Gln
100 105 110

Val Thr Val Ser Ser Arg Asn Ser Ile Ser Ser Leu Ser Ile Pro Ser
115 120 125

Thr Val Pro Arg Ala Arg Asp Pro Pro Val Asp Met Ala Ser Ser Glu
130 135 140

Asp Val Ile Lys Glu Phe Met Arg Phe Lys Val Arg Met Glu Gly Ser
145 150 155 160

Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro
165 170 175

Tyr Glu Gly Thr Gln Thr Ala Lys Leu Lys Val Thr Lys Gly Gly Pro
180 185 190

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Phe Gln Tyr Gly Ser
195 200 205

Lys Ala Tyr Val Lys His Pro Ala Asp Ile Pro Asp Tyr Leu Lys Leu
210 215 220

ES 2 401 359 T3

Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val Met Asn Phe Glu Asp
 225 230 235 240

Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser Leu Gln Asp Gly Glu
 245 250 255

Phe Ile Tyr Lys Val Lys Leu Arg Gly Thr Asn Phe Pro Ser Asp Gly
 260 265 270

Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu Ala Ser Thr Glu Arg
 275 280 285

Met Tyr Pro Glu Asp Gly Ala Leu Lys Gly Glu Ile Lys Met Arg Leu
 290 295 300

Lys Leu Lys Asp Gly Gly His Tyr Asp Ala Glu Val Lys Thr Thr Tyr
 305 310 315 320

Met Ala Lys Lys Pro Val Gln Leu Pro Gly Ala Tyr Lys Thr Asp Ile
 325 330 335

Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr Thr Ile Val Glu Gln
 340 345 350

Tyr Glu Arg Ala Glu Gly Arg His Ser Thr Gly Ala Gly Leu Tyr Lys
 355 360 365

<210> 33
 <211> 1107
 <212> ADN
 <213> secuencia artificial

<220>
 <223> secuencia que codifica el cromocuerpo anti-GFP (mutante C92S)

<400> 33

atggccgatg tgcagctggt ggagtctggg ggagccttgg tgcagccggg ggggtctctg 60
 agactctcct gtgcagcctc tggattcccc gtcaatcgct atagtatgag gtggtaccgc 120
 caggtccag ggaaggagcg cgagtgggtc gcgggatga gtagtgctgg tgategttca 180
 agttatgaag actccgtgaa gggccgattc accatctcca gagacgagcg caggaatacg 240
 gtgtatctgc aatgaacag cctgaaacct gaggacacgg ccgtgtatta ctgcaatgtc 300
 aatgtgggct ttgagtactg gggccagggg acccaggtca ccgtctcctc aaggaattcg 360
 atatcaagct tatcgatacc gtcgacggtc ccgcgggccc gggatocacc ggtcgacatg 420
 gcctctccg aggacgtcat caaggagttc atgcgcttca aggtgcgcat ggagggctcc 480
 gtgaacggcc acgagttcga gatcgagggc gagggcgagg gccgccccta cgagggcacc 540

ES 2 401 359 T3

cagaccgcca	agctgaagg	gaccaagggc	ggccccctgc	ccttcgcctg	ggacatcctg	600
tcccctcagt	tccagtacgg	ctccaaggcc	tacgtgaagc	accccgccga	catccccgac	660
tacttgaagc	tgtccttccc	cgagggcttc	aagtgggagc	gcgtgatgaa	cttcgaggac	720
ggcggcgtgg	tgaccgtgac	ccaggactcc	tccctgcagg	acggcgagtt	catctacaag	780
gtgaagctgc	gcggcaccaa	cttcccctcc	gacggccccg	taatgcagaa	gaagaccatg	840
ggctgggagg	cctccaccga	gcggatgtac	cccgaggacg	gcgccctgaa	gggcgagatc	900
aagatgaggc	tgaagctgaa	ggacggcggc	cactacgacg	ccgaggtcaa	gaccacctac	960
atggccaaga	agcccgtgca	gctgcccggc	gcctacaaga	ccgacatcaa	gctggacatc	1020
acctcccaca	acgaggacta	caccatcgtg	gaacagtacg	agcgcgccga	gggccgccac	1080
tccaccggcg	ccggcctgta	caagtaa				1107

REIVINDICACIONES

- 5 1. Un polipéptido que comprende la región variable una cadena pesada de un anticuerpo de *Camelidae* y capaz de unirse específicamente a un epítipo de GFP, que consiste en o comprende: marco 1, CDR1, marco 2, CDR2, marco 3 y CDR3,
- 10 (a) que tiene la secuencia de una cualquiera de las SEQ ID NOs: 10 o 30 o está codificado por la molécula de ácido nucleico de una cualquiera de las SEQ ID NOs: 11 o 31; o
(b) que está codificado por una molécula de ácido nucleico con al menos un 95% de identidad de secuencia con la SEQ ID NO: 11.
- 15 2. El polipéptido de la reivindicación 1, que comprende además
- (a) un marcador seleccionado del grupo que consiste en marcador de His, marcador de Strep, sitio de reconocimiento para biotilación; y opcionalmente
(b) el sitio de reconocimiento para una proteasa.
- 20 3. El polipéptido de la reivindicación 1 o 2, que comprende además una secuencia (poli) peptídica que es una proteína detectable, preferentemente derivable de una proteína fluorescente o cromófora.
4. Una molécula de ácido nucleico que codifica el polipéptido de una cualquiera de las reivindicaciones 1 a 3.
5. Un vector de expresión que comprende la molécula de ácido nucleico de la reivindicación 4.
- 25 6. Una célula hospedadora que comprende la molécula de ácido nucleico de la reivindicación 4 y/o el vector de expresión de la reivindicación 5.
- 30 7. Un método para precipitar específicamente GFP, una proteína marcada con GFP o proteínas que interaccionan con la misma en una muestra, que comprende la etapa de poner en contacto la muestra con el polipéptido de la reivindicación 1.
8. El método de la reivindicación 7, en el que el polipéptido está unido a un soporte sólido.
9. El método de la reivindicación 8, en el que el soporte sólido es sepharose.

Figura 1

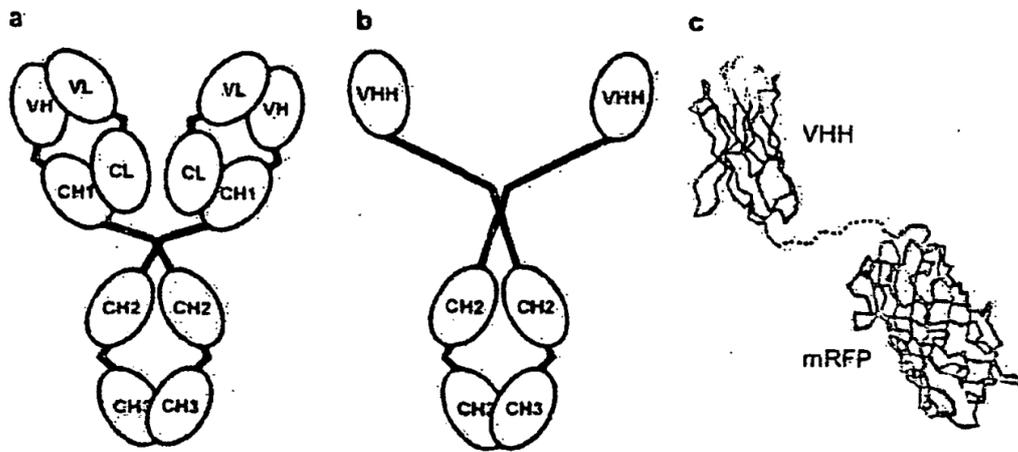


Figura 2

```

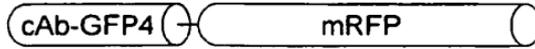
< marco 1 > < H1/CDR1 > < marco 2 > < CDR2 >
.....!10.....!20.....!30.....!40.....!50.....!60.
cAb-GFP4 QVQLVESGGALVQPGGSLRLSCAASG FPVN---RYSMR WYRQAPGKEREVVA GMSSAGDRSSYEDSVKG
RR6-R2 QVQLQESGGGLVQAGGSLRLSCAASG RAASGHGHYGMG WFRQVPGKEREVVA AIRWSGKETWYKDSVKG
RR6-R9 QVQLQESGGGLVQAGESLRLSCAASG NTFSG---GFMG WYRQAPGKQRELVA TINSRGI-TNYADSVKG
hCG-H14 DVQLQESGGGLVQAGGSLRLSCAASG RTGS---TYDMG WFRQAPGKERESVA AINWDSARTYYASSVRG
cAb-Lys3 DVQLQASGGGSVQAGGSLRLSCAASG YTIG---PYCMG WFRQAPGKEREVVA AINMGGGITYYADSVKG
cAb-TT1 EVQLQASGGGSVQAGGSLRLSCAASG GQTFDS--YAMA WFRQAPKGECELVA SIIG-DDNRNYADSVKG
AMYL-D10 QVQLVESGGGTVPAGGSLRLSCAASG NTLC---TYDMS WYRRAPGKGRDFVS GIDN-DGTTTTYVDSVAG
:*** **,* * **:****** . * **:****:: **: : * :*:.*

< marco 3 > < CDR3 >
....!70.....!80.....!90.....!100.....!110
cAb-GFP4 RFTISRDDARNTVYLMNSLKPEDTAVYYCNV N-----VGFY WGQGTQVTVSS
RR6-R2 RFTISRDNAKTTVYLMNSLKGEDTAVYYCAA RP-VRVADISL-----PVGFDY WGQGTQVTVSS
RR6-R9 RFTISRDNAKKTVYLEMNSLEPEDTAVYYCYT H-----YFRSY WGQGTQVTVSS
hCG-H14 RFTISRDNAKKTVYLMNSLKPEDTAVYTCGA G---EGGTWDS----- WGQGTQVTVSS
cAb-Lys3 RFTISQDNAKNTVYLLMNSLEPEDTAIYYCAA DSTIYASYECGHGLSTGGYGYDS WGQGTQVTVSS
cAb-TT1 RFTISRDNAKNTVYLMNDRNPEDTAVYYCAQ LGSAR-S----AMY-----C---- AGQGTQVTVSS
AMYL-D10 RFTISQGNAKNTAYLQMDSLKPDPTAMYYCKP S--LRYGLPG-----CPIIP WGQGTQVTVSS
***:*.:.:.*.** *::: :*:*.* :*****

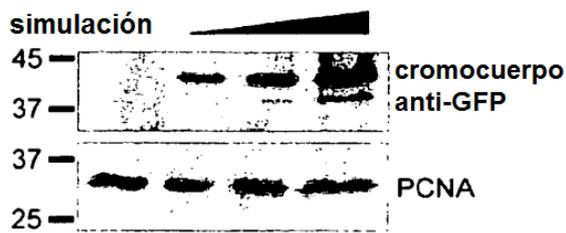
```

Figura 3

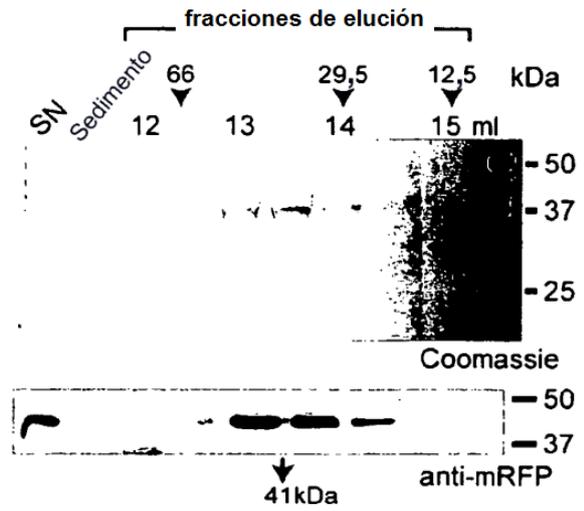
a cromocuerpo anti-GFP



b expresión en células 293T



c análisis de filtración en gel



d

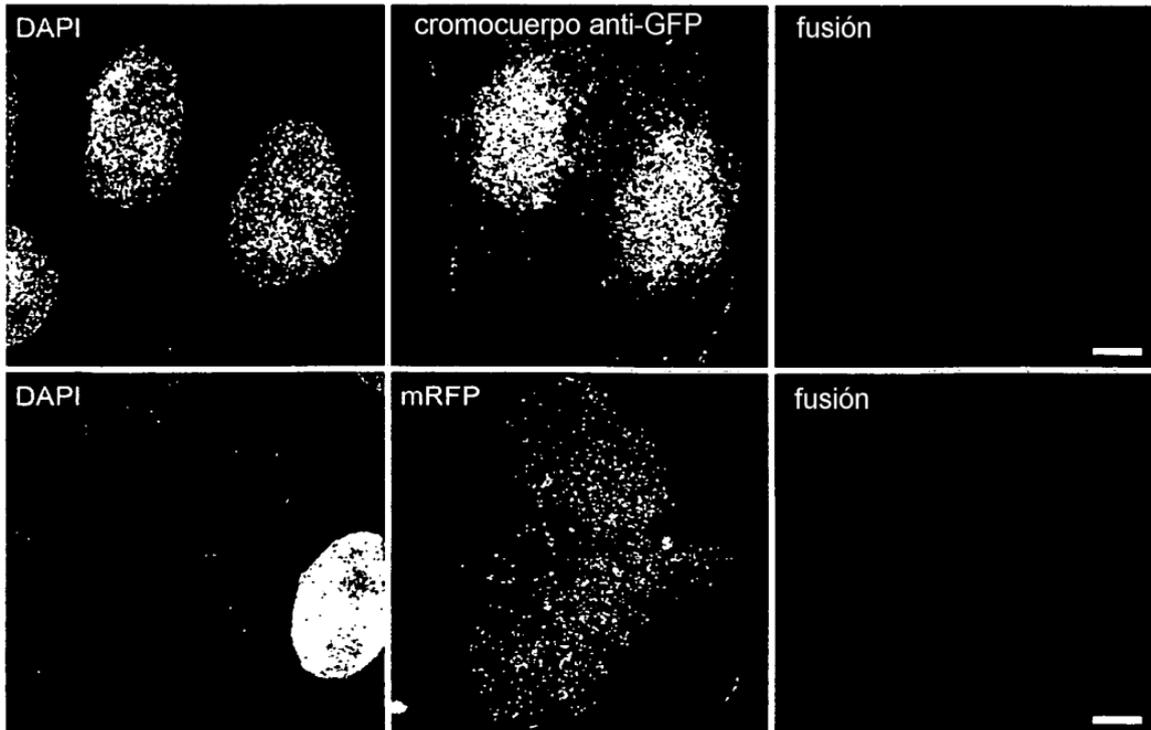


Figura 4

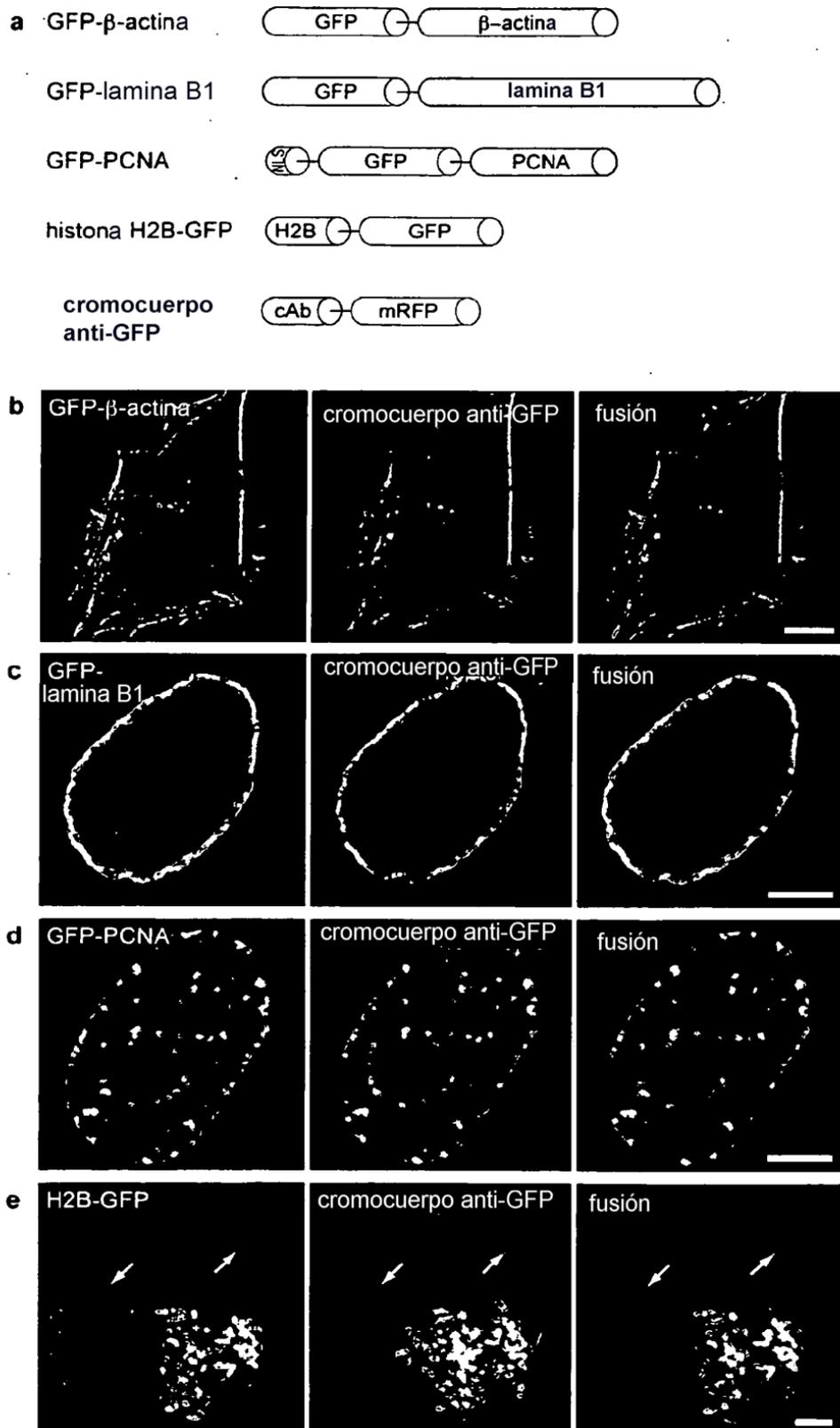


Figura 5

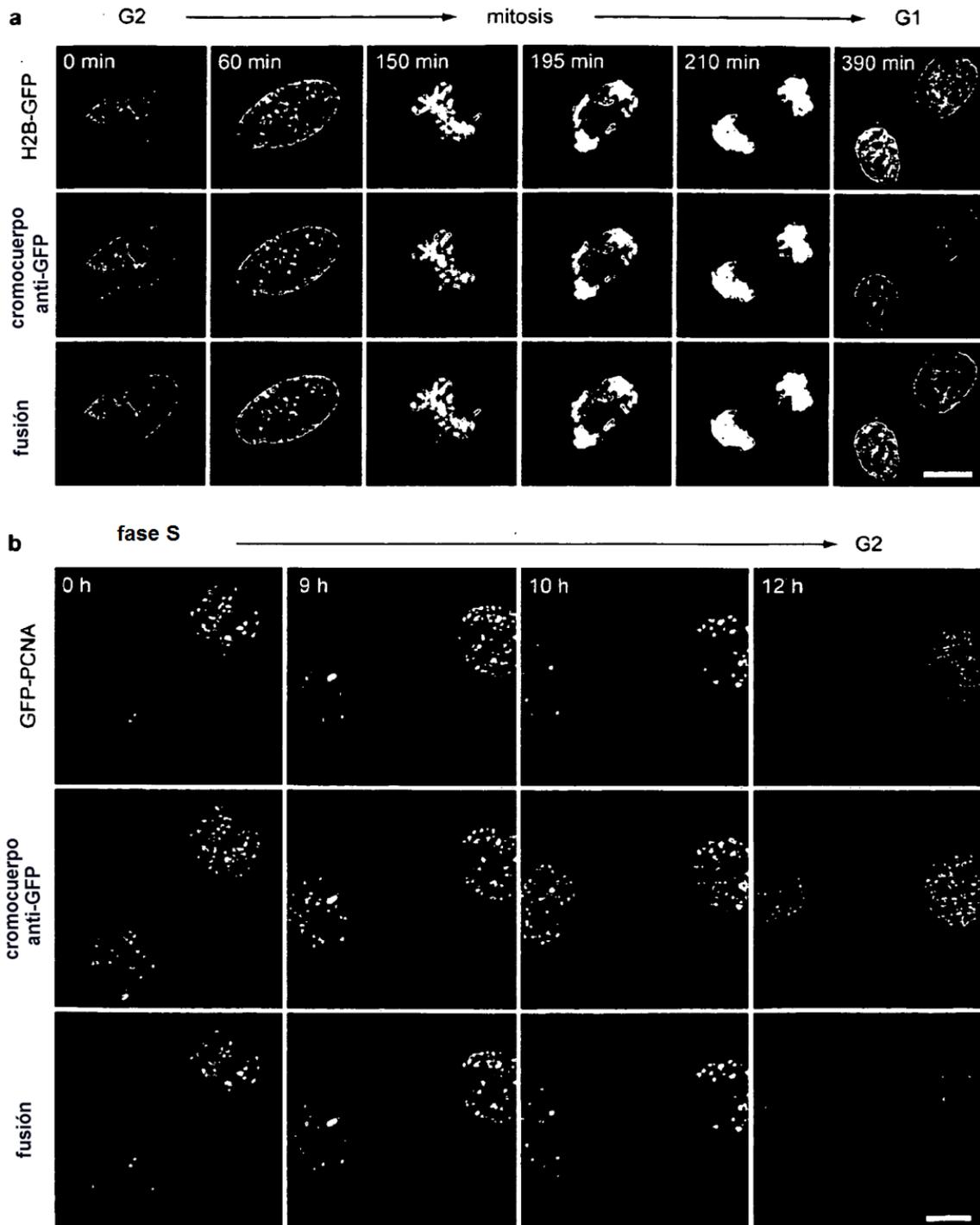
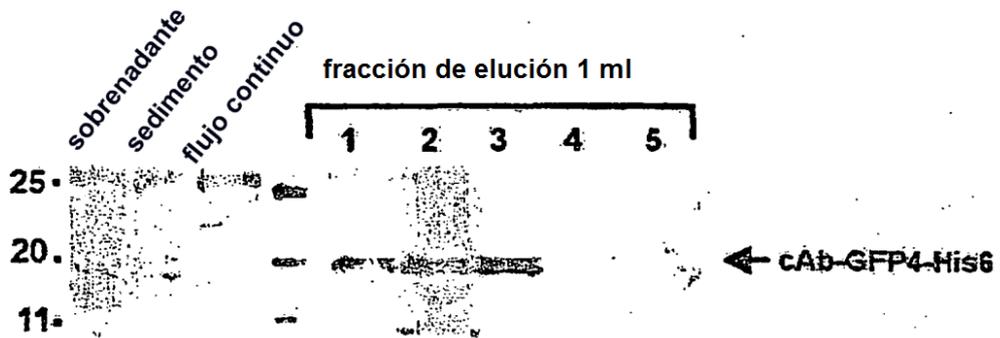


Figura 6

A. Purificación de IMAC de cAb-GFP4-His₆



B. Análisis de filtración en gel

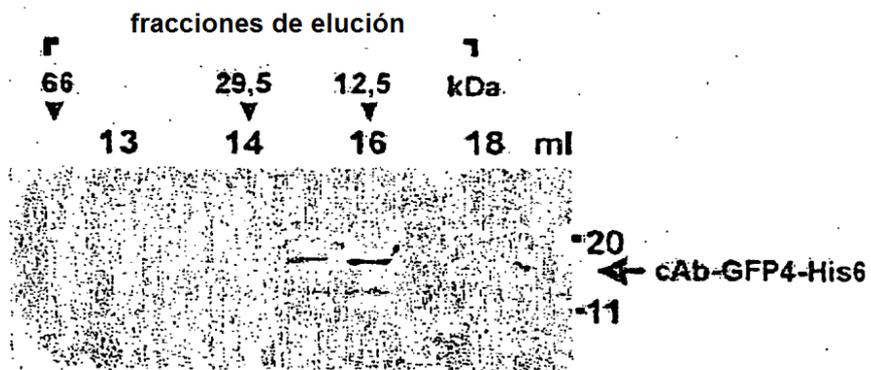


Figura 7

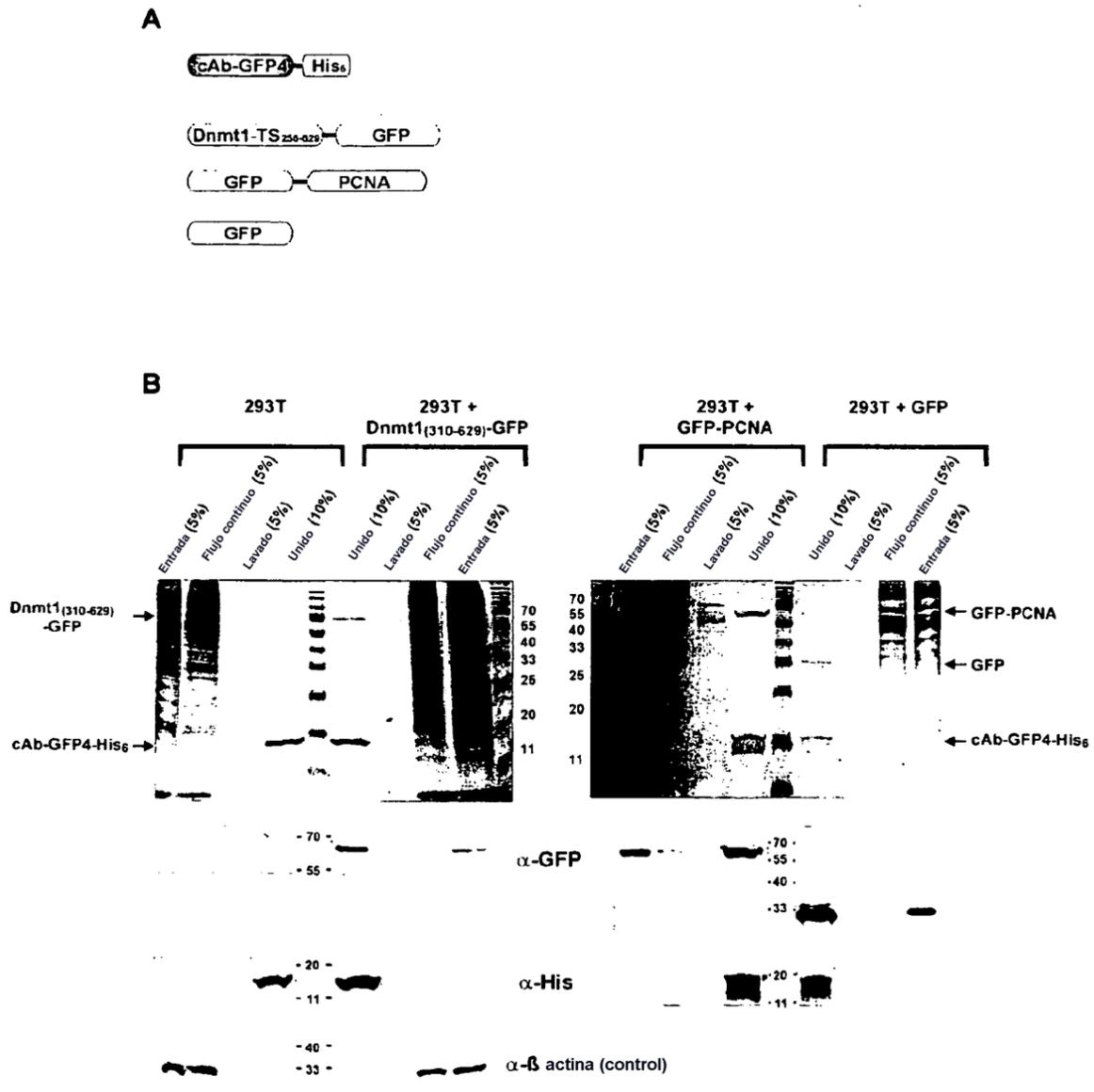


Figura 8

Análisis de inmunoprecipitación conjunta

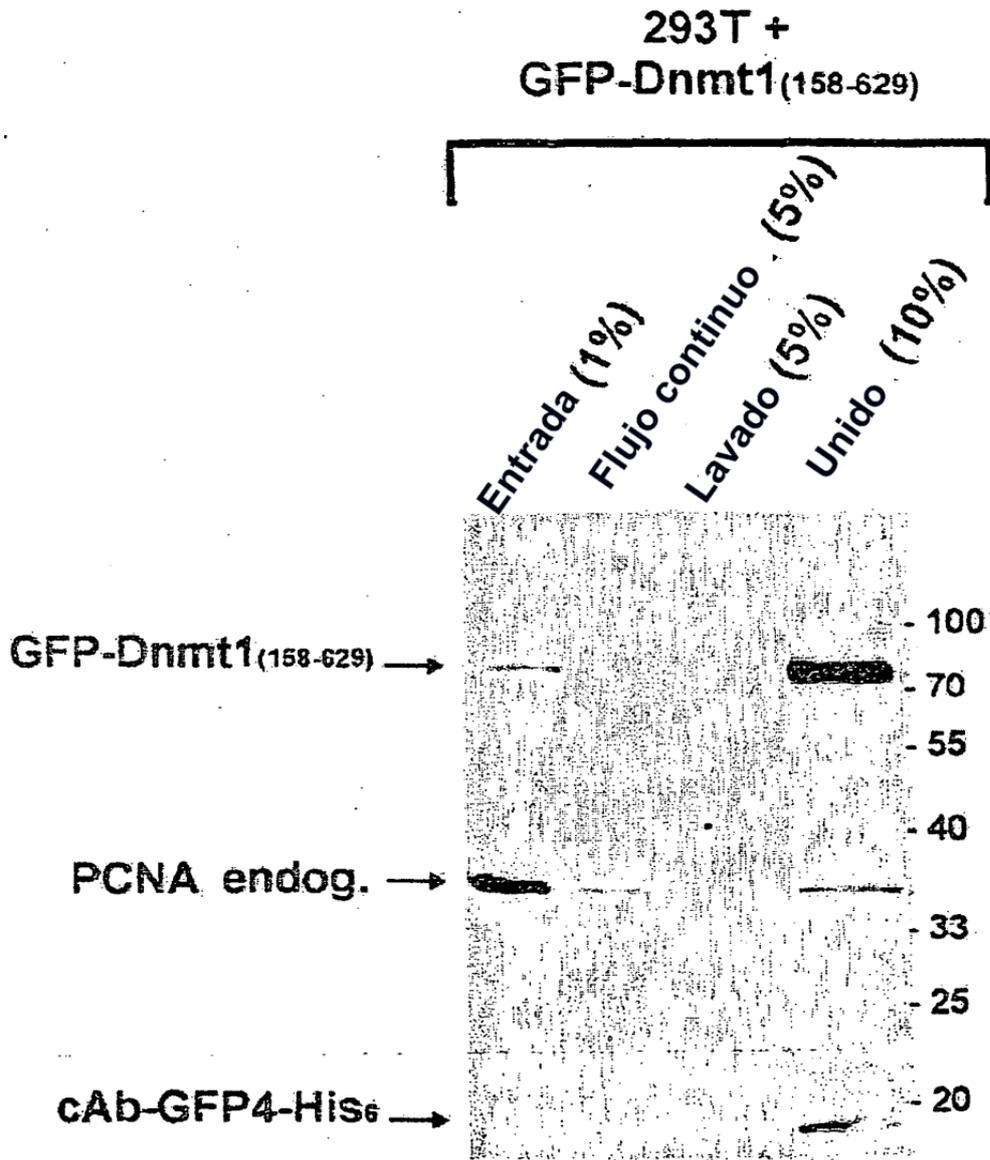
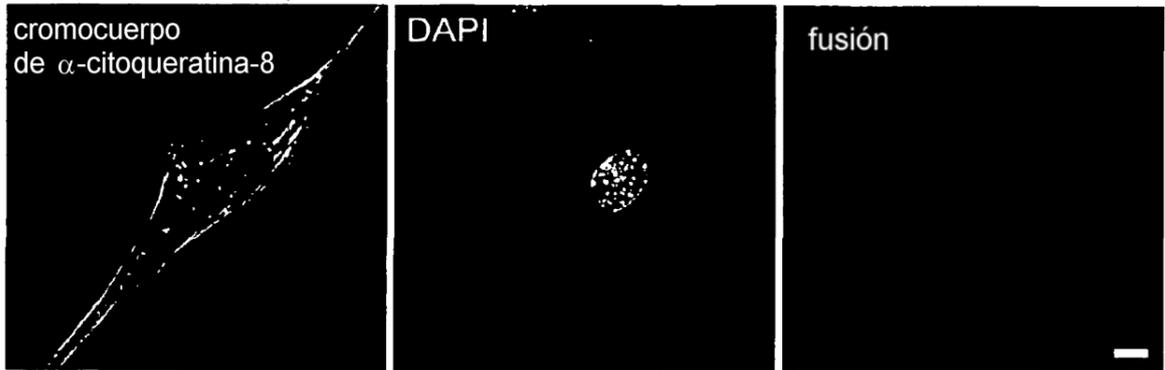


Figura 9

a cromocuerpo de α -citoqueratina-8



b cromocuerpo de α -lamina

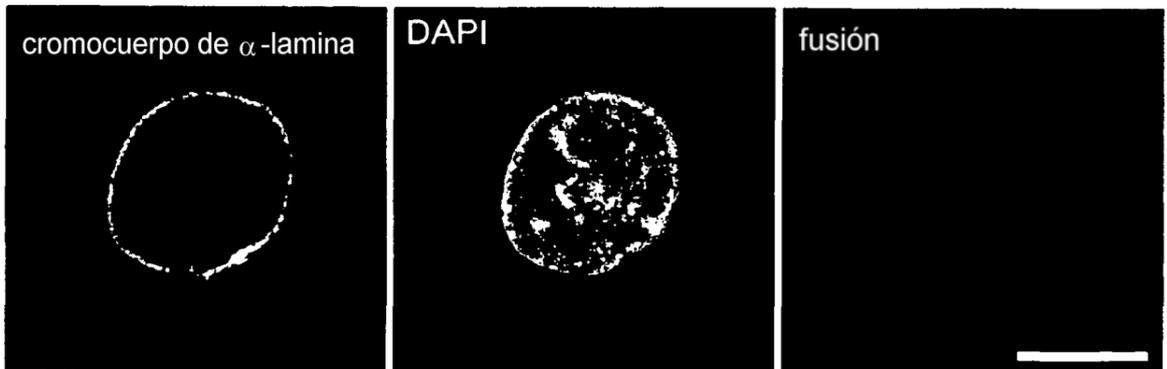


Figura 10

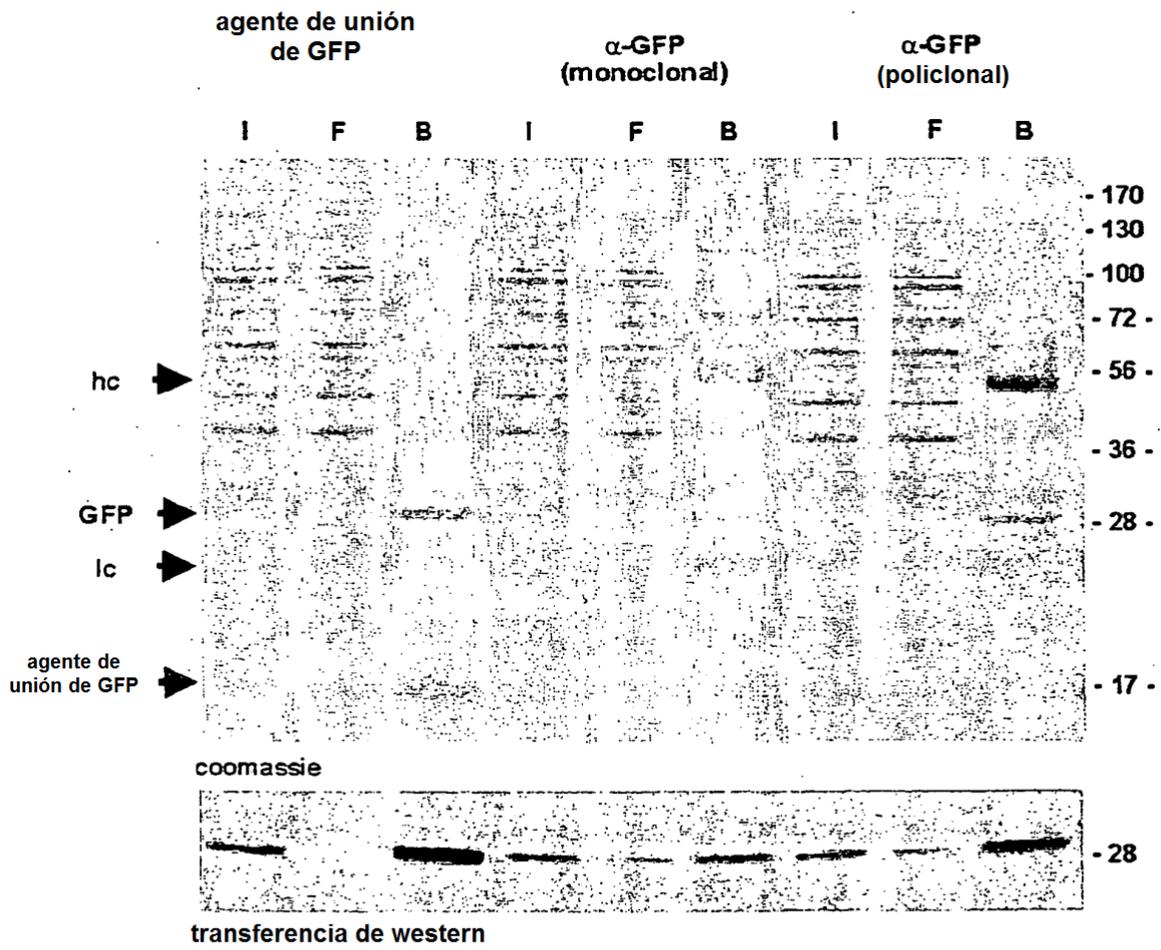


Figura 11

**Purificación en columna
con agente de unión de GFP**

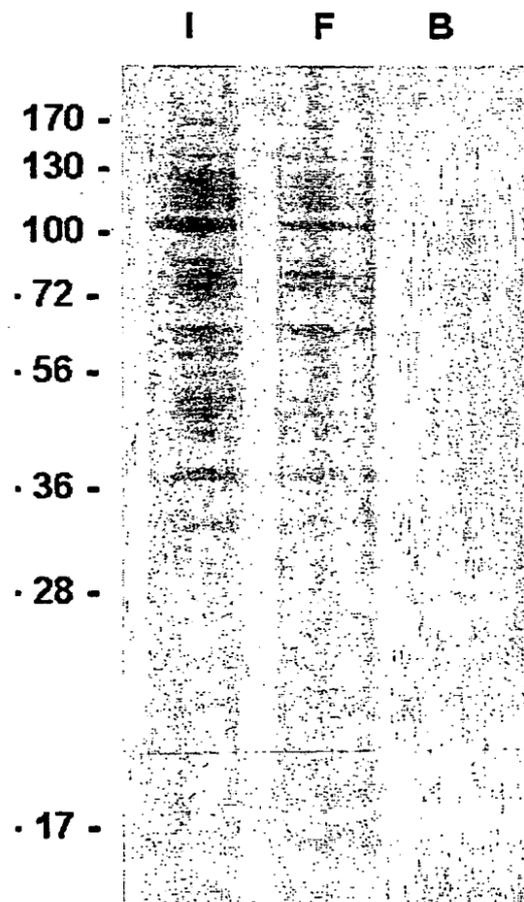


Figura 12

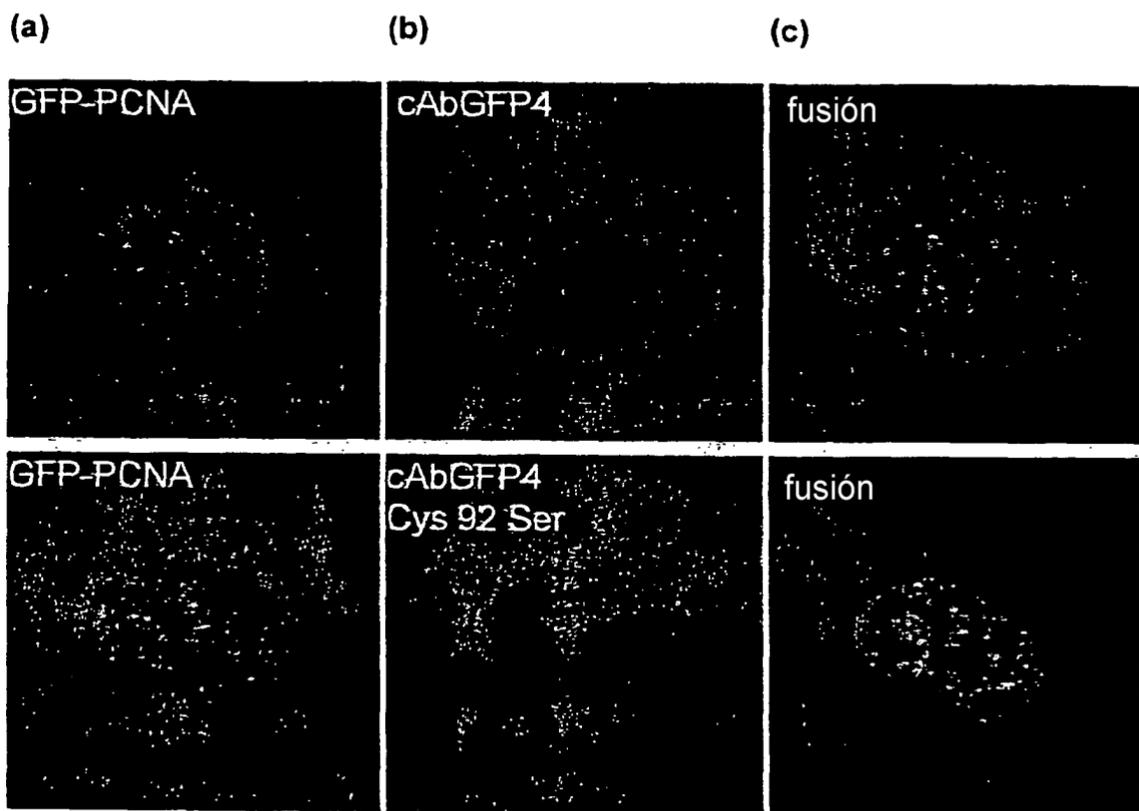
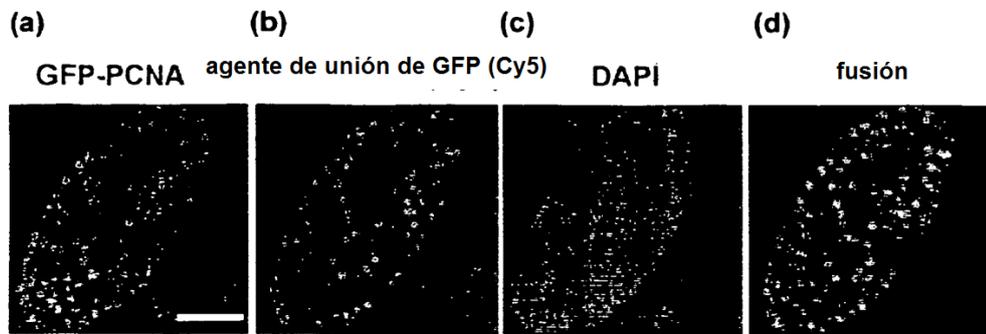


Figura 13



REFERENCIAS CITADAS EN LA DESCRIPCIÓN

La lista de referencias citadas por el solicitante es, únicamente, para conveniencia del lector. No forma parte del documento de patente europea. Si bien se ha tenido gran cuidado al compilar las referencias, no pueden excluirse errores u omisiones y la OEP declina toda responsabilidad a este respecto.

Documentos de patente citados en la descripción

- WO 8810315 A, Gingeras [0019]
- US 74612191 A [0021]
- EP 0464533 A [0031] [0109]
- EP 0232262 A [0031] [0109]
- US 20020123113 A1, Tsien [0033]
- US 20020107362 A1, Thastrup [0033]
- US 20020177189 A1, Bjorn [0033]
- US 6414119 B, Fisher [0033]
- US 20030013849 A, Ward [0033]
- WO 00146233 A, Dove and Hoegh-Guldberg [0038]
- WO 03091415 A2 [0050]
- US 6261804 B [0083] [0084]
- US 6281223 B [0083] [0084]
- US 6207446 B [0083] [0084]
- US 6214553 B [0083] [0084]
- CA 2045869 [0109]

Literatura no patente citada en la descripción

- SDA, WALKER et al. *Nucleic Acid Res.*, 1992, vol. 7, 1691-1696 [0019]
- TAS, KWOH et al. *Proc. Nat. Acad. Sci. USA*, 1989, vol. 86, 1173 [0019]
- Methods in Molecular Biology. vol. 226 [0019]
- PCR Technology: Principles and Applications for DNA Amplification. 1992 [0019]
- PCR Protocols: A guide to methods and applications. Academic Press, 1990 [0019]
- ERLICH, H. A. PCR Technology. Stockton Press, 1989 [0021]
- INNIS, M. A. ; D. H. GELFAND ; J. J. SNINSKY ; T. J. WHITE. PCR Protocols: A guide to methods and applications. Academic Press, 1990 [0021]
- MUYLDERMANS S. *Reviews in Molecular Biology*, 2001, vol. 74, 277-302 [0026]
- HORTON R. et al. *Gene*, 1989, vol. 77, 61-68 [0028] [0074] [0097]
- MURPHY et al. *Biochem J.*, 1991, vol. 227, 277-279 [0030]
- BEBBINGTON et al. *BiolTechnology*, 1992, vol. 10, 169-175 [0030] [0108]
- D. BENNETT et al. *J. Molecular Recognition*, 1995, vol. 8, 52-58 [0031] [0109]
- K. JOHANSON et al. *J. Biol. Chem.*, 1995, vol. 270, 9459-9471 [0031] [0109]
- HOOGENBOOM H.R. et al. Antibody phage display and its applications. *Immunotechnology*, 1998, vol. 4, 1-20 [0032]
- PLUCKTHUN A. Escherichia coli producing recombinant antibodies. *Bioprocess Technol.*, 1994, vol. 19, 233-252 [0032]
- VERMA R. et al. Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems. *J. Immunol. Methods*, 1998, vol. 216, 165-181 [0032]
- ZIMMER. *Chem. Rev.*, 2002, vol. 102, 759-781 [0033] [0036]
- ZHANG et al. *Nature Reviews*, 2002, vol. 3, 906-918 [0033] [0036]
- TSIEN. *Annual Review of Biochemistry*, 1998, vol. 67, 509-544 [0033]
- MATZ et al. *Nat. Biotechnol.*, 1999, vol. 17, 969-973 [0034]
- WIEHLER et al. *FEBS Letters*, 2001, vol. 487, 384-389 [0035]
- TERSKIKH et al. *Science*, 2000, vol. 290, 1585-1588 [0035]
- BAIRD et al. *Proc. Natl. Acad. Sci. USA*, 2000, vol. 97, 11984-11989 [0035]
- ANDO et al. *Proc. Natl. Acad. Sci. USA*, 2002, vol. 99, 12651-12656 [0037]
- WIEDENMANN et al. *Proc. Natl. Acad. Sci. USA*, 2000, vol. 97, 14091-14096 [0037]
- HONGBIN et al. *Biochem. Biophys. Res. Commun.*, 2003, vol. 301, 879-885 [0037]
- LUKYANOV et al. *J. Biol. Chem.*, 2000, vol. 275, 25879-25882 [0037]
- DOVE et al. *Coral Reefs*, 2001, vol. 19, 197-204 [0038]
- BEDDOE et al. *Acta Cryst.*, 2003, vol. D59, 597-599 [0038]
- BULINA et al. *BMC Biochem.*, 2002, vol. 3, 7 [0038]
- LUKYANOV et al. *J. Biol. Chem.*, 2000, vol. 275, 25879-82 [0039]
- GURSKAYA et al. *FEBS Letters*, 2001, vol. 507, 16-20 [0039]
- ZEYTUN et al. *Nature Biotechnology*, 2003, vol. 21 (12), 1473-1479 [0050]
- ZEYTUN et al. *Nature Biotechnology*, 2004, vol. 22 (5), 601 [0050]
- CAMPBELL et al. *PNAS*, 2002, vol. 99, 7877-7882 [0061]

- FISCHER et al. *FEBS Letters*, 2004, vol. 577, 227-232 [0061]
- SHANER, N.C. et al. *Nat. Biotech.*, 2004, vol. 22, 1567-1572 [0061]
- WANG, L. et al. *PNAS*, 2004, vol. 22, 1567-1572 [0061]
- D. KALDERON et al. *Cell*, 1984, vol. 39, 499 [0062] [0102]
- BLOBEL G. ; DOBBERSTEIN B. ER localisation signal of immunoglobulin light chain. *J.Cell.Biol.*, 1975, vol. 67, 835-851 [0062]
- NEUPERT W. *Annu. Rev. Biochem.*, 1997, vol. 66, 863-917 [0062] [0102]
- *Advances in Applied Mathematics*, 1981, vol. 2, 482-489 [0070]
- BRUTLAG. *Comp. App. Biosci.*, 1990, vol. 6, 237-245 [0070]
- THORSBY E. ; BRATILIE A. Histocompatibility Testing. 1970, 655 [0071]
- HARLOW, E. ; LANE, D. Antibodies, A Laboratory Manual. Cold Spring Harbor, 1988 [0072] [0114]
- CLACKSON T. ; LOWMAN H.B. Phage Display: A Practical Approach. Oxford University Press Inc, 2004 [0074] [0115]
- ZOLLER, M.J. *Curr. Opinion in Biotechnology*, 1992, vol. 3, 348-354 [0074]
- VIRNEKAS T. et al. *Nucl. Acids Res.*, 1994, vol. 22, 5600-5607 [0074]
- KIEKE. *Protein Eng.*, 1997, vol. 10, 1303-1310 [0078] [0105] [0114]
- MAIER. *Anal. Biochem.*, 1998, vol. 259, 68-73 [0078] [0114]
- MULLER. *Anal. Biochem.*, 1998, vol. 259, 54-61 [0078] [0114]
- BODER. *Nat. Biotechnol.*, 1997, 15553-15557 [0080]
- MATTHEAKIS et al. *Proc. Natl. Acad. Sci USA*, 1994, vol. 91, 9022-9026 [0082]
- HANES ; PLUCKTHRUN. *Proc. Natl. Acad. Sci USA*, 1997, vol. 94, 4937-4942 [0082]
- GATES et al. *J. Mol. Biol.*, 1996, vol. 255, 373-386 [0086]
- *Methods Enzymol.*, 1996, vol. 267, 171-191 [0086]
- PETERSEN. *Mol. Gen. Genet.*, 1995, vol. 249, 425-31 [0088]
- KAY. *Gene*, 1993, 12859-65 [0088]
- GRAM. *J. Immunol. Methods*, 1993, vol. 161, 169-76 [0088]
- SMITH. *Science*, 1985, vol. 228, 1315-1317 [0089]
- CWIRLA. *Proc. Natl. Acad. Sci. USA*, 1990, vol. 87, 6378-6382 [0089]
- SAMBROOK ; RUSSELL. *Molecule Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 2001 [0089]
- PHAGE. *Display of Peptides and Proteins: A Laboratory Manual*. Academic Press, 1996 [0089]
- CRAMERI. *Eur. J. Biochem.*, 1994, vol. 226, 53-58 [0089]
- DE KRUIF. *Proc. Natl. Acad. Sci. USA*, 1995, vol. 92, 3938-3942 [0089]
- MCGREGOR. *Mol. Biotechnol.*, 1996, vol. 6, 155-162 [0089]
- JACOBSSON. *Biotechniques*, 1996, vol. 20, 1070-1076 [0089]
- JESPERS. *Gene*, 1996, vol. 173, 179-181 [0089]
- JACOBSSON. *Microbiol Res.*, 1997, vol. 152, 121-128 [0089]
- FACK. *J. Immunol. Methods*, 1997, vol. 206, 43-52 [0089]
- ROSSENU. *J. Protein Chem.*, 1997, vol. 16, 499-503 [0089]
- KATZ. *Annu. Rev. Biophys. Biomol. Struct.*, 1997, vol. 26, 27-45 [0089]
- RADER. *Curr. Opin. Biotechnol.*, 1997, vol. 8, 503-508 [0089]
- GRIFFITHS. *Curr. Opin. Biotechnol.*, 1998, vol. 9, 102-108 [0089]
- FELICI. *J. Mol. Biol.*, 1991, vol. 222, 301-310 [0090]
- CRAMERI et al. *Eur. J. Biochem.*, 1994, 22653-58 [0092]
- MCGREGOR ; ROBINS. *Anal. Biochem.*, 2001, vol. 294, 108-117 [0093]
- SPADA. *J. Biol. Chem.*, 1997, vol. 378, 445-456 [0095]
- PEDRAZZI. *FEBS Lett.*, 1997, vol. 41 (5), 289-293 [0095]
- HENNECKE. *Protein Eng.*, 1998, vol. 11, 405-410 [0095]
- BLOBEL G. ; DOBBERSTEIN B. *J.Cell.Biol.*, 1975, vol. 67, 835-851 [0102]
- SKERRA ; SCHMIDT. *Biomolecular Engineering*, 1999, vol. 16, 79-86 [0105]
- GHAHROUDI et al. *FEBS Letters*, 1997, vol. 414, 521-526 [0108]
- MURPHY et al. *Biochem J.*, vol. 227, 277-279 [0108]
- SAMBROOK ; RUSSEL. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 2001 [0110] [0114]
- *Protein Methods*. Wiley-Liss [0111]
- *Methods in Enzymology*, 1990 [0111]
- *Guide to Protein Purification*. Academic Press, vol. 182 [0111]
- SAMBROOK et al. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 2001 [0114]
- *Methods in Enzymology. Guide to Protein Purification*. Academic Press, 1990, vol. 182 [0114]
- BIOCCA, S. ; NEUBERGER, M.S. ; CATTANEO, A. Expression and targeting of intracellular antibodies in mammalian cells. *Embo J*, 1990, vol. 9, 101-108 [0191]
- CATTANEO, A. ; BIOCCA, S. The selection of intracellular antibodies. *Trends Biotechnol.*, 1999, vol. 17, 115-121 [0191]

- **BIOCCA, S. ; PIERANDREI-AMALDI, P. ; CATTANEO, A.** Intracellular expression of anti-p21 ras single chain Fv fragments inhibits meiotic maturation of xenopus oocytes. *Biochem: Biophys. Res. Commun.*, 1993, vol. 197, 422-427 [0191]
- **BIOCCA, S. ; NEUBERGER, M.S. ; CATTANEO, A.** Expression and targeting of intracellular antibodies in mammalian cells. *Embo J.*, 1990, vol. 9, 101-108 [0191]
- **MARASCO, W.A. ; CHEN, S. ; RICHARDSON, J.H. ; RAMSTEDT, U. ; JONES, S.D.** Intracellular antibodies against HIV-1 envelope protein for AIDS gene therapy. *Hum. Gene Ther.*, 1998, vol. 9, 1627-1642 [0191]
- **CARDINALE, A. ; LENER, M. ; MESSINA, S. ; CATTANEO, A. ; BIOCCA, S.** The mode of action of Y13-259 scFv fragment intracellularly expressed in mammalian cells. *FEBS Lett.*, 1998, vol. 439, 197-202 [0191]
- **KONTERMANN, R.E.** Intrabodies as therapeutic agents. *Methods*, 2004, vol. 34, 163-170 [0191]
- **HAMERS-CASTERMAN, C. et al.** Naturally occurring antibodies devoid of light chains. *Nature*, 1993, vol. 363, 446-448 [0191]
- **MUYLDERMANS, S. ; ATARHOUCHE, T. ; SALDANHA, J. ; BARBOSA, J.A. ; HAMERS, R.** Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng.*, 1994, vol. 7, 1129-1135 [0191]
- **SHERIFF, S. ; CONSTANTINE, K.L.** Redefining the minimal antigen-binding fragment. *Nat. Struct. Biol.*, 1996, vol. 3, 733-736 [0191]
- **MUYLDERMANS, S.** Single domain camel antibodies: current status. *J. Biotechnol.*, 2001, vol. 74, 277-302 [0191]
- **ARBABI GHARROUDI, M. ; DESMYTER, A. ; WYNS, L. ; HAMERS, R. ; MUYLDERMANS, S.** Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.*, 1997, vol. 414, 521-526 [0191] [0192]
- **VAN DER LINDEN, R.H. et al.** Comparison of physical chemical properties of llama VHH antibody fragments and mouse monoclonal antibodies. *Biochim. Biophys. Acta*, 1999, vol. 1431, 37-46 [0191]
- **LAUWEREYS, M. et al.** Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *Embo J.*, 1998, vol. 17, 3512-3520 [0191]
- **VAN KONINGSBRUGGEN, S. et al.** Llama-derived phage display antibodies in the dissection of the human disease oculopharyngeal muscular dystrophy. *J. Immunol. Methods*, 2003, vol. 279, 149-161 [0191] [0192]
- **CORTEZ-RETAMOZO, V. et al.** Efficient tumor targeting by single-domain antibody fragments of camels. *Int. J. Cancer*, 2002, vol. 98, 456-462 [0191]
- **JOBLING, S.A. et al.** Immunomodulation of enzyme function in plants by single-domain antibody fragments. *Nat. Biotechnol.*, 2003, vol. 21, 77-80 [0191]
- **KABAT, E.A. ; WU, T.T.** Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. *J. Immunol.*, 1991, vol. 147, 1709-1719 [0191]
- **CONRATH, K.E. ; WERNERY, U. ; MUYLDERMANS, S. ; NGUYEN, V.K.** Emergence and evolution of functional heavy-chain antibodies in Camelidae. *Dev. Comp. Immunol.*, 2003, vol. 27, 87-103 [0191]
- **WESTPHAL, M. et al.** Microfilament dynamics during cell movement and chemotaxis monitored using a GFP-actin fusion protein. *Curr. Biol.*, 1997, vol. 7, 176-183 [0191]
- **DAIGLE, N. et al.** Nuclear pore complexes form immobile networks and have a very low turnover in live mammalian cells. *J. Cell. Biol.*, 2001, vol. 154, 71-84 [0191]
- **WARBRICK, E.** The puzzle of PCNA's many partners. *Bioessays*, 2000, vol. 22, 997-1006 [0191]
- **LEONHARDT, H. et al.** Dynamics of DNA replication factories in living cells. *J. Cell. Biol.*, 2000, vol. 149, 271-280 [0191]
- **SPORBERT, A. ; GAHL, A. ; ANKERHOLD, R. ; LEONHARDT, H. ; CARDOSO, M.C.** DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. *Mol. Cell*, 2002, vol. 10, 1355-1365 [0191]
- **KANDA, T. ; SULLIVAN, K.F. ; WAHL, G.M.** Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr. Biol.*, 1998, vol. 8, 377-385 [0191]
- **CARDINALE, A. ; FILESI, I. ; MATTEI, S. ; BIOCCA, S.** Intracellular targeting and functional analysis of single-chain Fv fragments in mammalian cells. *Methods*, 2004, vol. 34, 171-178 [0191]
- **SAERENS, D. et al.** Identification of a Universal VHH Framework to Graft Non-canonical Antigen-binding Loops of Camel Single-domain Antibodies. *J. Mol. Biol.*, 2005 [0191]
- **DESMYTER, A. et al.** Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat. Struct. Biol.*, 1996, vol. 3, 803-811 [0191]
- **RIECHMANN, L. ; MUYLDERMANS, S.** Single domain antibodies: comparison of camel VH and camelised human VH domains. *J. Immunol. Methods*, 1999, vol. 231, 25-38 [0191]
- **HEIM, R. ; CUBITT, A.B. ; TSIEN, R.Y.** Improved green fluorescence. *Nature*, 1995, vol. 373, 663-664 [0191] [0192]

- **CAMPBELL, R.E. et al.** A monomeric red fluorescent protein. *Proc. Natl Acad. Sci. U S A*, 2002, vol. 99, 7877-7882 [0191]
- **SPINELLI, S. et al.** Camelid heavy-chain variable domains provide efficient combining sites to haptens. *Biochemistry*, 2000, vol. 39, 1217-1222 [0191]
- **SPINELLI, S. et al.** Domain swapping of a llama VHH domain builds a crystal-wide beta-sheet structure. *FEBS Lett.*, 2004, vol. 564, 35-40 [0191]
- **SPINELLI, S. et al.** The crystal structure of a llama heavy chain variable domain. *Nat. Struct. Biol.*, 1996, vol. 3, 752-757 [0191]
- **DESMYTER, A. et al.** Three camelid VHH domains in complex with porcine pancreatic alpha-amylase. Inhibition and versatility of binding topology. *J. Biol. Chem.*, 2002, vol. 277, 23645-23650 [0191]
- **SASSO, E.H. ; SILVERMAN, G.J. ; MANNIK, M.** Human IgA and IgG F(ab')₂ that bind to staphylococcal protein A belong to the VHIII subgroup. *J Immunol*, 1991, vol. 147, 1877-1883 [0192]
- **RIECHMANN, L. ; DAVIES, J.** Backbone assignment, secondary structure and protein A binding of an isolated, human antibody VH domain. *J Biomol NMR*, 1995, vol. 6, 141-152 [0192]
- **RIECHMANN, L. ; MUYLDERMANS, S.** Single domain antibodies: comparison of camel VH and camelised human VH domains. *J. Immunol. Methods*, 1999, vol. 231, 25-38 [0192]