

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 617 783**

51 Int. Cl.:

C12N 15/62 (2006.01)

C12N 15/10 (2006.01)

C12N 9/12 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.02.2012** **PCT/EP2012/053463**

87 Fecha y número de publicación internacional: **07.09.2012** **WO2012117031**

96 Fecha de presentación y número de la solicitud europea: **29.02.2012** **E 12706065 (5)**

97 Fecha y número de publicación de la concesión europea: **14.12.2016** **EP 2681319**

54 Título: **Sensor del sustrato de quinasa**

30 Prioridad:

01.03.2011 GB 201103453

01.03.2011 US 201161464285 P

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

19.06.2017

73 Titular/es:

VIB VZW (50.0%)

Rijvisschestraat 120

9052 Gent, BE y

UNIVERSITEIT GENT (50.0%)

72 Inventor/es:

TAVERNIER, JAN y

LIEVENS, SAMUEL

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 617 783 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

Sensor del sustrato de quinasa

La presente invención se refiere a un complejo de proteínas citoplasmáticas, que comprende (a) una primera proteína de fusión recombinante que comprende una quinasa, condensada a un primer polipéptido de interacción y (b) una segunda proteína de fusión recombinante que comprende un dominio que comprende un sitio de fosforilación informador, en el que dicho dominio está condensado a un segundo polipéptido de interacción. La presente invención se refiere, además, a un método para detectar la interacción compuesto-compuesto utilizando dicho complejo de proteína citoplasmática, y a células que comprenden dicho complejo de proteínas citoplasmáticas.

Las interacciones proteína-proteína son una clave esencial en todos los procesos biológicos, desde la replicación y expresión de los genes a la morfogénesis de los organismos. Las interacciones proteína-proteína gobiernan, entre otros, la interacción ligando-receptor y la subsiguiente vía de señalización; son importantes en el ensamblaje de subunidades de la enzima, en la formación de estructuras supramoleculares biológicas tales como ribosomas, filamentos y partículas de virus y en las interacciones antígeno-anticuerpo.

Los investigadores han desarrollado varias estrategias en un intento de identificar las interacciones proteína-proteína. La co-purificación de proteínas y la co-inmunoprecipitación estaban entre las primeras técnicas utilizadas. Sin embargo, estos métodos son tediosos y no permiten la selección de alto rendimiento. Además de ello, requieren una lisis que deteriora el contexto celular normal. Un gran avance se obtuvo mediante la introducción de los enfoques genéticos, de los cuales el doble híbrido en levadura (Fields y Song, 1989) es el más importante. Aunque esta técnica se convirtió en ampliamente utilizada, tiene varios inconvenientes. Las proteínas de fusión deben ser translocadas al núcleo, lo que no siempre es evidente. Las proteínas con propiedades de activación de la transcripción intrínsecas pueden provocar falsos positivos. Además de ello, las interacciones que dependen de modificaciones secundarias de la proteína tales como fosforilación no se pueden detectar fácilmente.

Se han desarrollado varios sistemas alternativos para resolver uno o más de estos problemas. Enfoques basados en la presentación de fagos evitan la translocación nuclear. El documento WO9002809 describe cómo una proteína de unión se puede presentar en la superficie de un paquete genético, tal como un fago filamentoso, en el que el gen que codifica la proteína de unión está empaquetado dentro del fago. Los fagos, que portan la proteína de unión que reconoce la molécula diana se aíslan y amplifican. Se han propuesto varias mejoras de la estrategia de presentación de fagos, tal como se describe, p. ej., en los documentos WO9220791, WO9710330 y WO9732017.

Sin embargo, todos estos métodos adolecen de las dificultades que son inherentes a la metodología de presentación de fagos: las proteínas necesitan ser expuestas en la superficie del fago y están tan expuestas a un entorno que no es fisiológica relevante para la interacción *in vivo*. Además de ello, cuando se rastrea una fagoteca, habrá una competencia entre los fagos, lo que resulta en una selección de los aglutinantes de alta afinidad.

El documento US5637463 describe una mejora del sistema de doble híbrido en levadura, mediante el cual se pueden rastrear interacciones proteína-proteína dependientes de la modificación. Sin embargo, este método se basa en la co-expresión de la enzima modificadora, que ejercerá su actividad en el citoplasma y puede modificar otras enzimas que la implicada en la interacción proteína-proteína, que puede a su vez afectar la viabilidad del organismo huésped.

Una evolución interesante se describe en el documento US5776689, por el denominado sistema de reclutamiento de proteínas. Las interacciones proteína-proteína son detectadas por el reclutamiento de un factor de intercambio de nucleótidos de guanina (Sos) en la membrana plasmática, en donde Sos activa una molécula informadora Ras. Esto resulta en la supervivencia de la célula que de otro modo no sobreviviría en las condiciones del cultivo utilizadas. Aunque este método tiene ciertamente la ventaja de que la interacción proteína-proteína se lleva a cabo en condiciones fisiológicas en el espacio sub-membranal, tiene varios inconvenientes. No pueden ser detectadas interacciones que dependen de la modificación. Además, el método utiliza la vía Ras pleiotrópica, lo cual puede provocar complicaciones técnicas.

Una mejora importante en la detección de interacciones proteína-proteína se describió en el documento WO0190188, que describe el denominado sistema Mappit. El método, basado en un receptor de citoquinas, permite no sólo una detección fiable de las interacciones proteína-proteína en células de mamífero, sino también se pueden detectar interacciones entre proteínas dependientes de la modificación, así como tres interacciones proteína-proteína híbridas complejas mediadas por un compuesto pequeño (Caligiuri et al., 2006). Sin embargo, aunque muy útil, el sistema está limitado en su sensibilidad y algunas interacciones débiles no se pueden detectar. Además de ello, como este es un sistema basado en membrana, las interacciones nucleares normalmente no se detectan.

Sigue existiendo la necesidad de un sistema de identificación sensible para las interacciones compuesto-compuesto que puede estudiar estas interacciones en condiciones fisiológicas, con un fondo bajo y mediante el cual se pueden aislar las interacciones proteína-proteína dependiente de la modificación.

La presente invención, basada en una quinasa citoplasmática mutante, preferiblemente una quinasa citoplasmática mutante constitutiva satisface esta necesidad y proporciona ventajas adicionales. Mientras que MAPPIT utiliza una quinasa inducible controlada por el receptor del ligando, para limitar y corregir el fondo, la quinasa citoplasmática, especialmente la quinasa citoplasmática constitutiva sorprendentemente proporciona una relación más alta de señal/ruido, permitiendo la detección de interacciones que no se pueden detectar en el sistema MAPPIT. La alta relación señal/ruido es inesperada para un sistema citoplasmático, dado que la persona experta en la técnica esperaría que la quinasa citoplasmática se puede mover libremente en el citoplasma y llegar a su sustrato sin otras señales de reclutamiento que la quinasa y el dominio de fosforilación. Mediante el uso de un sistema citoplasmático se ha resuelto el problema de la detección de interacciones proteína-proteína nuclear. La sensibilidad inesperadamente alta del presente método permite la detección de interacciones de proteínas débiles tales como la interacción p51/p51, que no puede ser detectada por Mappit. Alternativamente, se puede utilizar una quinasa mutante inducible por un compuesto pequeño exógeno. Esta estrategia combina las ventajas de la quinasa constitutiva, con la capacidad de inducción, sin interferir en las vías endógenas.

Una representación esquemática de la invención se da en la Figura 1.

Un primer aspecto de la invención es un complejo de proteínas citoplasmáticas que comprende (a) una primera proteína de fusión recombinante que comprende una quinasa condensada a un primer polipéptido de interacción y (b) una segunda proteína de fusión recombinante que comprende un dominio que comprende un sitio de fosforilación del informador, con lo que dicho dominio está condensado a un segundo polipéptido de interacción.

Preferentemente, dicha primera proteína de fusión recombinante no comprende un sitio de fosforilación del informador. Preferiblemente, dicha quinasa es una quinasa mutante, más preferiblemente una tirosina quinasa mutante. En una realización preferida, dicha tirosina quinasa mutante es una quinasa constitutiva. En otra realización preferida, dicha tirosina quinasa mutante es un mutante inactivo que es activado por la adición de una pequeña molécula exógena. Una quinasa mutante de este tipo es conocida por la persona experta en la técnica, y se ha descrito, como un ejemplo no limitativo, por Qiao et al (2006) como un mutante Src 388R/A o una mutación 391R/A en la correspondiente proteína Src humana (número de acceso NP_938033, versión NP_938033.1). Alternativamente, puede ser una mutación similar en la familia de quinasas Jak tal como, pero no limitado a Tyk2 1027R/A.

En una realización preferida, dicha quinasa es un mutante quinasa constitutivo derivado de Tyk2 tal como, pero no limitado a un mutante de delección constitutiva Tyk2 o y/o un mutante de Tyk2 V678F. Derivado de Tyk2, tal como se utiliza en esta memoria, significa que la quinasa es una parte de la tirosina-proteína quinasa no receptora Tyk2 humana (GenBank número de acceso NP_003322; versión NP_003322.3; SEQ ID N° 26) o un mutante o una variante de la misma, en donde dicha parte muestra actividad de la quinasa constitutiva. Una variante, como un ejemplo no limitativo, es un homólogo, parálogo u ortólogo. "Homólogos" de una proteína abarcan péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que tienen sustituciones de aminoácidos, delecciones y/o inserciones con relación a la proteína no modificada en cuestión y que tienen actividad biológica y funcional similar a la proteína no modificada de la que se derivan. "Ortólogos" y "parálogos" abarcan conceptos evolutivos utilizados para describir las relaciones ancestrales de los genes. "Parálogos" son los genes dentro de la misma especie que se han originado a través de la duplicación de un gen ancestral; "ortólogos" son genes de diferentes organismos que se han originado a través de la especiación, y también se derivan de un gen ancestral común. Preferiblemente, dicho homólogo, "ortólogo" o "parálogo" tiene una identidad de secuencia a nivel de proteína de al menos 50%, 51%, 52%, 53%, 54% o 55%, 56%, 57%, 58%, 59 %, preferiblemente 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, más preferiblemente 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, incluso más preferiblemente 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, lo más preferiblemente 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más, según se mide en un BLASTp (Altschul et al, 1997.; Altschul et al, 2005). Variantes y partes de las mismas de acuerdo con la invención sí muestran actividad de quinasa. Preferiblemente, dicha parte es una parte con actividad de quinasa constitutiva, preferiblemente el fragmento 589-1187 de SEQ ID N° 26. Alternativamente, dicha parte es la parte que corresponde al fragmento 589-1187 de SEC ID N° 26 en un homólogo, parálogo u ortólogo tal como se definen anteriormente, en el que dicha parte tiene actividad de quinasa constitutiva. En una realización alternativa dicha quinasa constitutiva es una quinasa constitutiva derivada de una quinasa Jak, preferiblemente de una quinasa Jak seleccionada del grupo que consiste en Jak1 (Número de acceso P23458, versión P23458.2), Jak2 (Número de acceso O60674, versión O60674.2) y Jak3 (Número de acceso P52333, versión P52333.2) o un mutante o variante del mismo tal como se define anteriormente. Preferiblemente, dicha quinasa constitutiva es un mutante de delección constitutiva Jak2. Todavía en otra realización alternativa dicha quinasa constitutiva es una quinasa constitutiva derivada de una quinasa Src (Número de acceso NP_005408, versión NP_005408.1) o un mutante o variante de la misma tal como se define anteriormente. Preferiblemente, dicha quinasa constitutiva es un mutante de delección de Src tal como se representa en SEQ ID N° 27. Derivado, tal como se utiliza en esta memoria, significa que la quinasa es una parte de la tirosina-proteína no receptora citada, o de un mutante o variante de la misma, en el que dicha parte muestra la actividad de quinasa constitutiva.

En una realización preferida, dicha primera proteína de fusión recombinante comprende SEQ ID N° 1 (pseudo-quinasa + quinasa). En otra realización preferida dicha primera proteína recombinante comprende SEQ ID N°2 (quinasa). Todavía en otra realización preferida, dicha primera proteína recombinante comprende una secuencia seleccionada del grupo que consiste en SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7 y SEQ ID N° 27. En una realización preferida, dicha segunda proteína de fusión recombinante comprende SEQ ID N° 3 (sitio STAT). En otra realización preferida, dicha segunda proteína de fusión recombinante comprende SEQ ID N° 4. Lo más preferiblemente, dicho complejo de proteínas citoplasmáticas comprende una primera proteína de fusión recombinante que comprende SEQ ID N° 1 o SEQ ID N° 2 y una segunda proteína de fusión recombinante que comprende SEQ ID N° 3 o SEQ ID N° 4.

En los casos en los que se utiliza un mutante inactivo que puede ser activado por un compuesto pequeño, una realización preferida es una quinasa mutante en donde dicha quinasa mutante es una mutación Src 388R/A o una mutación 391R/A en la correspondiente proteína Src humana (Número de acceso NP_938033, versión NP_938033.1). Otra realización preferida es un mutante en el que dicha quinasa mutante es un mutante Tyk2 1027R/A (numeración de SEQ ID N° 26). Preferiblemente, dicho compuesto pequeño es imidazol.

La quinasa de dicha primera proteína de fusión recombinante está condensada a un primer polipéptido de interacción tal como se define más adelante. Dicha fusión se puede producir en el extremo amino-terminal, en el extremo carboxi-terminal, o internamente en el polipéptido recombinante. En los casos de quinasas que consisten en dos dominios, el primer polipéptido de interacción puede incluso producirse como una fusión interna entre los dos dominios. El dominio que comprende el sitio de fosforilación del informador de dicha segunda proteína de fusión recombinante está condensado a un segundo polipéptido de interacción tal como se define más adelante. Dicha fusión se puede producir en el extremo amino-terminal, en el extremo carboxi-terminal, o internamente en el polipéptido recombinante. La primera proteína de fusión recombinante puede no comprender un sitio de fosforilación del informador. La primera proteína de fusión recombinante, la segunda proteína de fusión recombinante o ambas pueden comprender, además, otras secuencias tales como, pero no limitadas a una señal de localización para dirigir el complejo citoplasmático a una organela específica en el citoplasma de la célula, o al núcleo. Preferiblemente, dicho complejo de proteínas citoplasmáticas es fosforilado en dicho sitio de fosforilación del informador, incluso más preferiblemente dicha fosforilación es llevada a cabo por dicha quinasa. En una realización preferida, el complejo de proteínas de acuerdo con la invención es ensamblado por la interacción directa del primer polipéptido de interacción con el segundo polipéptido de interacción. La primera y la segunda proteína de interacción pueden ser idénticas en los casos en que se estudia la homodimerización. Sin embargo, el complejo de proteínas no se limita a dos compuestos, y la interacción puede ser mediada por uno o más de otros compuestos. Como un ejemplo no limitante, el primer polipéptido de interacción puede unirse a una proteína que se une, a su vez, por el segundo polipéptido de interacción. En lugar de una proteína, dicho compuesto de puente puede ser una o más moléculas de naturaleza no proteica, o mezclas de compuestos pequeños y proteínas. Otro ejemplo no limitativo es un complejo de proteínas, en el que un compuesto pequeño se une al primero, así como al segundo polipéptido de interacción.

La formación del complejo puede depender de la modificación del primer y/o segundo polipéptido de interacción. Dicha modificación puede ser, pero no se limita a la presencia o ausencia de fosforilación, acetilación, acilación, metilación, ubiquitinilación o glucosilación o a la aparición de una escisión proteolítica o no, o una combinación de las mismas. Preferiblemente, dicha modificación se lleva a cabo por una enzima modificadora (como se define a continuación). Preferentemente, dicha enzima modificadora se condensa con la primera y/o segunda proteína de fusión recombinante. En el caso de la fosforilación, la modificación de dicha enzima puede ser la quinasa constitutiva condensada al primer polipéptido de interacción tal como se describe anteriormente.

Otro aspecto de la invención es un método para detectar la interacción compuesto-compuesto que comprende (a) condensar una quinasa citoplasmática a un primer polipéptido de interacción, resultando en una primera proteína de fusión recombinante (sin sitio de fosforilación del informador) de acuerdo con la invención (b) condensar un dominio que comprende un sitio de fosforilación del informador a un segundo polipéptido de interacción, resultando en una segunda proteína de fusión recombinante de acuerdo con la invención, (c) expresar ambas proteínas de fusión recombinantes en una célula y (d) identificar y/o seleccionar aquellas células en las que dicho sitio de fosforilación del informador es fosforilado.

Preferiblemente, dicha quinasa es una quinasa mutante, incluso más preferiblemente una tirosina quinasa mutante. En una realización preferida, dicha quinasa mutante es una quinasa constitutiva, incluso más preferiblemente dicha quinasa es un mutante constitutivo derivado de una quinasa Tyk2, preferiblemente un mutante de delección o un mutante V678F constitutivo. Alternativamente, se puede utilizar un mutante de delección Jak2 constitutivo, o un mutante de delección de quinasa Src constitutivo. En otra realización preferida, dicha quinasa es una quinasa mutante inactiva que es activada por la adición de un compuesto pequeño exógeno. Una quinasa mutante de este tipo es conocida por la persona experta en la técnica, y se ha descrito, como un ejemplo no limitativo, por Qiao et al (2006).

En una realización preferida, dicha primera proteína de fusión recombinante comprende la SEQ ID N°1. En otra realización preferida, dicha primera proteína de fusión recombinante comprende la SEQ ID N° 2. En aún otra realización más preferida, dicha primera proteína recombinante comprende una secuencia seleccionada del grupo que consiste en SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7 y SEQ ID N° 27. En una realización preferida, dicha

segunda proteína de fusión recombinante comprende la SEQ ID N° 3. En otra realización preferida, dicha segunda proteína de fusión recombinante comprende la SEQ ID N° 4. Lo más preferiblemente, dicho complejo de proteínas citoplasmáticas comprende una primera proteína de fusión recombinante que comprende la SEQ ID N° 1 o la SEQ ID N° 2 y una segunda proteína de fusión recombinante que comprende la SEQ ID N° 3 o la SEQ ID N° 4.

El gen que codifica la primera y/o la segunda proteína de fusión recombinante se puede disponer aguas debajo de un promotor constitutivo o un promotor inducible. Esta última construcción puede tener algunas ventajas en los casos en los que existe una competencia por el sitio de unión entre el polipéptido de interacción y polipéptidos endógenos. La inducción de la proteína de fusión recombinante que comprende el primer polipéptido de interacción en presencia de los segundos polipéptidos de interacción puede facilitar la unión y evitar la saturación de los sitios de unión con los polipéptidos endógenos. Dicha célula puede ser cualquier célula, incluyendo, pero no limitada a células bacterianas, células fúngicas, células de levadura, células de insectos y células de mamíferos. Preferiblemente, dicha célula es una célula eucariota, incluso más preferiblemente dicha célula es una célula de mamíferos. En los casos en los que se estudia la homodimerización, la primera y la segunda proteína de interacción pueden ser idénticas. En este caso, la misma proteína se condensa con la quinasa en un lado y a un dominio que comprende el sitio de fosforilación del informador en el otro lado.

Dicha primera proteína de fusión recombinante y dicha segunda proteína de fusión recombinante pueden estar situadas en uno o en vectores separados. El vector utilizado para la transformación y la expresión del primero y/o segundo polipéptido recombinante puede ser cualquier vector, conocido por la persona experta en la técnica, incluyendo pero no limitado a vectores episomales, vectores de integración y vectores virales. En una realización preferida, una célula eucariota que porta un gen que codifica una primera proteína de fusión recombinante de acuerdo con la invención se transforma o transfecta con un banco de vectores que codifica segundas proteínas de fusión recombinantes de acuerdo con la invención. La interacción del primer polipéptido de interacción con el segundo polipéptido de interacción dará lugar a la fosforilación del sitio de fosforilación del informador y puede ser detectada por el uso de un sistema informador. Una realización específica del método para detectar la unión compuesto-compuesto es un método por el que dichos compuestos son proteínas y la unión es una interacción proteína-proteína. Otra realización específica es un método para detectar la interacción proteína-proteína, por lo que dicha interacción depende del estado de modificación, y dicha modificación es llevada a cabo por una enzima de modificación. Todavía otra realización específica es un método para detectar la unión compuesto-compuesto, en el que dicha unión está mediada por tres o más participantes. En este caso, uno o más participantes pueden no ser o no ser completamente de naturaleza proteica. Es evidente para una persona experta en la técnica que el primer polipéptido de interacción, de acuerdo con la invención, puede, como un ejemplo no limitativo, unirse a una molécula pequeña. Por otro lado, el segundo polipéptido de interacción, de acuerdo con la invención, puede también unirse a la molécula pequeña, de modo que la primera y la segunda están enlazadas entre sí por dicha molécula pequeña. Dicha molécula pequeña puede estar presente en la célula huésped tal como un compuesto producido por la propia célula, o como un compuesto que ha sido recogido del medio. Preferiblemente, dicho método para detectar la unión compuesto-compuesto comprende la construcción de una célula eucariota que comprende una primera proteína de fusión recombinante de acuerdo con la invención, seguido de la transformación o transfección de dicha célula por un banco de vectores que codifican segundas proteínas de fusión recombinantes de acuerdo con la invención. La unión compuesto-compuesto es detectada por la fosforilación del sitio de fosforilación del informador. La detección de la fosforilación del complejo se puede realizar mediante el aislamiento del complejo, y el análisis de la fosforilación (p. ej., utilizando fósforo radiactivo marcado). Preferiblemente, dicha fosforilación del sitio de fosforilación del informador resulta en la inducción de un sistema informador. Un sistema informador puede ser cualquier sistema que permita la detección y/o la selección de las células que portan un complejo de proteínas citoplasmáticas fosforilado de acuerdo con la invención. Es evidente para la persona experta en la técnica que se pueden utilizar varios sistemas indicadores. Como ejemplo no limitativo, dicha fosforilación puede conducir a la inducción de una ruta de señalización, y un gen de luciferasa, un gen de resistencia a antibióticos o un gen marcador de la superficie celular puede ser colocado bajo el control de un promotor que es inducido por dicha ruta de señalización. Alternativamente, se pueden utilizar sistemas informadores que se basan en el cambio en las características de los compuestos de dicha ruta de señalización, cuando dicha ruta es activa tal como la fosforilación y/o dimerización de tales compuestos. Todavía otra posibilidad es el uso de anticuerpos y/o de otros polipéptidos de unión que reconocen específicamente el sitio de fosforilación del informador fosforilado.

Una realización preferida es un sitio de fosforilación del informador que es parte de un sitio de unión de transductor de señales y activador de la transcripción (STAT), más preferentemente parte de un sitio de unión STAT1 y/o STAT3. En este caso, el gen informador se puede colocar bajo el control de un promotor inducible STAT tal como, pero no limitado al promotor de la Proteína Asociada a la Pancreatitis 1 (rPAP). Alternativamente, dado que la fosforilación del sitio de fosforilación del informador resultará en la unión de un polipéptido de STAT al sitio de fosforilación del informador fosforilado, seguido de la fosforilación de dicho polipéptido STAT y la subsiguiente dimerización de dos moléculas STAT fosforiladas, la dimerización en sí puede ser utilizada como señal del informador. Todavía otro sistema indicador alternativo consiste en un ensayo de complementación de proteínas, en el que una parte de la proteína se incorpora en o se asocia con el complejo de proteínas citoplasmáticas de acuerdo

con la invención, y la segunda parte de la proteína es reclutada al sitio de fosforilación del informador fosforilado, lo que conduce a una actividad detectable de la proteína reconstituida.

Todavía otro aspecto de la invención es una célula, que comprende un complejo de proteínas citoplasmáticas de acuerdo con la invención. Dicha célula puede ser cualquier célula, incluyendo, pero no limitada a células bacterianas, células fúngicas, células de levadura, células de insectos y células de mamíferos. Preferiblemente, dicha célula es una célula eucariota, incluso más preferiblemente dicha célula es una célula de mamífero. Preferiblemente, dicho sitio de fosforilación del informador en dicho complejo de proteínas citoplasmáticas está fosforilado. Preferiblemente, dicha célula también comprende un sistema informador, lo que permite la detección de la fosforilación del sitio de fosforilación del informador de dicho complejo de proteínas citoplasmáticas.

Todavía otro aspecto de la invención es un método para detectar compuestos que interrumpen una interacción polipéptido-polipéptido, que comprende (a) cultivar una célula que comprende un complejo de proteínas de acuerdo con la invención, en donde dicho complejo de proteínas citoplasmáticas comprende la interacción polipéptido-polipéptido que se quiere interrumpir, en ausencia y en presencia de al menos un compuesto, (b) comparar la fosforilación del sitio de fosforilación del informador en la segunda proteína de fusión recombinante de dicho complejo de proteínas citoplasmáticas de las células, que se cultiva en presencia y en ausencia de dicho compuesto y (c) identificar y/o seleccionar aquellas células en las que dicho sitio de fosforilación del informador no está fosforilado. Es evidente para la persona experta en la técnica que la interacción polipéptido-polipéptido, tal como se ha mencionado anteriormente, puede ser una interacción indirecta por tres de más participantes, con lo que incluso uno o más de los participantes puede no ser de naturaleza proteica.

DEFINICIONES

Se recogen las siguientes definiciones para ilustrar y definir el significado y alcance de diversos términos utilizados para describir la invención en esta memoria.

Proteína, tal como se utiliza en esta memoria, significa una cadena compuesta de aminoácidos, independiente de la longitud. Los términos 'proteína' y 'polipéptido' son intercambiables. La proteína puede ser modificada por modificaciones tales como, pero no limitadas a fosforilación, glucosilación, ubiquitinilación y acetilación.

Dominio, tal como se utiliza en esta memoria, es una parte de un polipéptido, en donde dicha parte puede portar una función específica tal como, pero no limitada a y el centro enzimático o un sitio de fosforilación.

Complejo de proteínas, tal como se utiliza en esta memoria, significa una estructura que comprende al menos dos moléculas de proteína, no enlazadas de modo covalente. Los complejos de proteínas pueden consistir en más de dos proteínas, e incluyen otras moléculas que no son proteínas. Algunos ejemplos no limitantes de tales moléculas son iones de metales, ATP, o moléculas de hidratos de carbono.

Complejo de proteínas citoplasmáticas significa que el complejo de proteínas tal como se describe anteriormente se mueve libremente en el citoplasma de la célula, y no está enlazado a la membrana celular. Sin embargo, dicho complejo de proteínas citoplasmáticas puede dirigirse al núcleo o a organelas citoplasmáticas utilizando secuencias de señales de localización.

Una quinasa, tal como se utiliza en esta memoria, es un polipéptido que puede transferir un grupo fosfato a un aminoácido del mismo o de otro polipéptido. Preferiblemente, dicho aminoácido es una serina, una treonina o una tirosina. Incluso más preferiblemente dicho aminoácido está incrustado en un sitio de fosforilación. Un *sitio de fosforilación*, tal como se

utiliza en esta memoria, es un patrón de varios aminoácidos que comprende preferiblemente una serina, treonina o una tirosina, y que determina el aminoácido que será fosforilado por la quinasa. La mayoría de las quinasas pueden aparecer en un estado inactivo y en un estado activo, en el que el sitio de fosforilación del informador sólo es fosforilado en el estado activo de la quinasa. Las quinasas se pueden activar desde la forma inactiva a la forma activa por fosforilación, o por otras modificaciones tales como proteólisis, o por mutación. La fosforilación puede ser una autofosforilación, fosforilación cruzada (por un complejo de proteínas de quinasas idénticas) o por la acción de otra quinasa. Una quinasa citoplasmática es una quinasa que se mueve libremente en el citoplasma, y no ligada ni es reclutada a la membrana celular. La quinasa citoplasmática de acuerdo con la invención se puede derivar de una quinasa enlazada a la membrana, mediante la supresión del dominio de anclaje a la membrana o de reclutamiento de membrana, sin pérdida de la actividad de quinasa enzimática.

Una *quinasa mutante* es una quinasa de la que la secuencia difiere de la secuencia que se produce normalmente en la naturaleza, por la sustitución, delección o inserción de uno o más aminoácidos. Homólogos, ortólogos y parálogos no se consideran mutantes, ya que están presentes como una forma normal en la naturaleza.

Constitutiva, tal como se utiliza aquí, significa que la quinasa está continuamente en el estado activo, normalmente como consecuencia de una mutación, o por escisión proteolítica que separa un inhibidor. Quinasas constitutivas son conocidas por la persona experta en la técnica y comprenden, pero no se limitan a formas truncadas de Tyk2, formas truncadas de quinasa Src y mutaciones puntuales tales como Tyk2 (V678F), Jak1 (V658F) y Jak2 (V617F).

Un mutante de quinasa *inactivo* significa que la forma mutante muestra una actividad de quinasa que es significativamente menor que la forma original no mutada. Preferiblemente, la actividad restante es inferior a 50% de la actividad original, incluso más preferible inferior a 20%, más preferiblemente menor que 10%, lo más preferiblemente menor que 5% de la actividad original.

Activado por la adición de un pequeño compuesto exógeno, tal como se utiliza en esta memoria, significa que la actividad de la quinasa inactiva es parcial o totalmente restaurada por adición de un pequeño compuesto a las células, con lo que dicho pequeño compuesto, exógeno a la célula, es absorbido por la célula y activa la quinasa como un compuesto exógeno intracelular. "Activado por la adición de un pequeño compuesto exógeno" se utiliza para hacer una distinción con la activación inducida por el receptor del ligando, en donde un ligando se une a la parte extracelular de un receptor, e induce la activación de la quinasa. "Exógeno", tal como se utiliza aquí, significa que el compuesto no está presente normalmente en la célula.

El *sitio de fosforilación del informador* es el sitio que está fosforilado en el complejo de proteínas tras la interacción del primer y del segundo polipéptido de interacción; es distinto de un posible sitio de fosforilación en el dominio quinasa que se autofosforila en la quinasa constitutiva.

Primer polipéptido de interacción, tal como se utiliza aquí, es un polipéptido del que se quiere estudiar la interacción con uno o más compuestos. Al primer polipéptido de interacción se le alude normalmente como "cebo" en la terminología de doble híbrido.

Segundo polipéptido de interacción, tal como se utiliza aquí, es un polipéptido que se presenta para estudiar su interacción con el primer polipéptido de interacción. Al segundo polipéptido de interacción se le alude normalmente como "presa" en la terminología de doble híbrido. Está claro para la persona experta en la técnica que el primer y el segundo polipéptido de interacción son intercambiables en la invención, a este respecto que, un "cebo" o una "presa" pueden condensarse con quinasa constitutiva de acuerdo con la invención. De hecho, el complejo de proteínas resultante tendrá una composición global idéntica, compuesto de los cuatro elementos esenciales (primer polipéptido de interacción, segundo polipéptido de interacción, quinasa constitutiva y sitio de fosforilación del informador), e independiente de si el primer polipéptido de interacción se condensa a la quinasa constitutiva o al sitio de fosforilación del informador (en donde el segundo polipéptido de interacción se condensa a continuación al sitio de fosforilación del informador, y la quinasa constitutiva, respectivamente), la interacción del primer con el segundo polipéptido de interacción dará lugar a la formación de un complejo de proteínas citoplasmáticas de acuerdo con la invención, y dará lugar a la fosforilación del sitio de fosforilación del informador. En una realización preferida, la primera y la segunda proteína de interacción son idénticas para estudiar la homodimerización u homomultimerización de una proteína. En otra realización preferida, la primera y la segunda proteínas son diferentes, lo que permite estudiar las interacciones proteína-proteína de heterodímeros o heteromultímeros.

Modificación de la enzima, tal como se utiliza aquí, significa cualquier actividad enzimática que resulta en una modificación del primer y/o segundo polipéptido de interacción. Tal modificación puede ser, como un ejemplo no limitativo, fosforilación, acetilación, acilación, metilación, ubiquitinilación o glucosilación, la aparición de una escisión proteolítica o una combinación de los mismos. Preferiblemente, dicha modificación es la fosforilación. La modificación de la enzima puede estar asociada a o incorporada en la primera y/o la segunda proteína de fusión recombinante del complejo de proteínas citoplasmáticas. Se puede co-expresar con el primer y/o el segundo polipéptido de interacción, o se puede condensar a uno o a ambos de dichos polipéptidos. En el caso de la fosforilación, la quinasa constitutiva del complejo de proteínas citoplasmáticas de acuerdo con la invención puede llevar a cabo tanto la fosforilación del sitio de activación como la fosforilación de uno o ambos de los polipéptidos de interacción.

Compuesto significa cualquier compuesto químico o biológico, incluyendo moléculas simples o complejas, orgánicas o inorgánicas, péptidos, péptido-miméticos, proteínas, anticuerpos, hidratos de carbono, ácidos nucleicos o derivados de los mismos.

Interacción significa cualquier interacción, ya sea directa o indirecta. Una interacción directa implica un contacto entre los participantes en la interacción. Una interacción indirecta significa cualquier interacción en la que los participantes en la interacción interactúan en un complejo de más de dos compuestos. Esta interacción puede ser completamente indirecta, con la ayuda de uno o más compuestos de puenteo, o parcialmente indirecta, en que todavía existe un contacto directo que se estabiliza por la interacción de uno o más compuestos.

BREVE DESCRIPCIÓN DE LAS FIGURAS

Figura 1: Representación esquemática del complejo de proteínas citoplasmáticas de acuerdo con la invención. Un primer polipéptido interactuante (X) se condensa a una quinasa constitutiva (K) y un segundo polipéptido interactuante (Y) se condensa a un sitio de fosforilación del informador (R). La interacción entre los polipéptidos que interactúan X e Y resulta en el sitio de fosforilación del informador que está fosforilado (P) por la quinasa constitutiva, lo que conduce a una actividad detectable.

Figura 2: Detección de la interacción entre subunidades de la transcriptasa inversa (RT) de VIH1 en una variante de ensayo que comprende proteínas de fusión de quinasa Tyk2 mutantes.

A. Vista esquemática del ensayo. El primer polipéptido interactuante (X) se condensa a la región C-terminal de Tyk2 que comprende el dominio quinasa, mientras que el segundo polipéptido interactuante (Y) se condensa a un fragmento de gp130 que contiene sitios de fosforilación. Cuando los polipéptidos X e Y interactúan, el dominio quinasa Tyk2 fosforila los sitios de fosforilación de gp130. Factores de transcripción de STAT3 se reclutan a estos sitios fosforilados y son a su vez fosforilados por el dominio de quinasa Tyk2, lo que resulta en su activación. Los dímeros de factores de transcripción STAT3 activados son capaces de unirse al promotor rPAPI específico que impulsa la expresión de un gen informador de luciferasa de luciérnaga. La actividad de este gen informador se mide como la producción de luz en un ensayo de detección de luciferasa utilizando un luminómetro.

B. Aplicación al análisis de las subunidades de RT de VIH1. Las células fueron transfectadas con los siguientes plásmidos:

- a) pMET7-HA-Tyk2(C)-RTP66 + pMG2-RTP51 + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Tyk2(C)-RTP66 + pMG2-RTP66 + pXP2d2-rPAPI-luciferasa
- c) pMET7-HA-Tyk2(C)-RTP51 + pMG2-RTP51 + pXP2d2-rPAPI-luciferasa

La actividad de luciferasa se muestra como la inducción múltiple con respecto a la actividad de luciferasa medida en células transfectadas con la misma fusión Tyk2(C), un fragmento de gp130 no condensado y el plásmido informador de luciferasa (pMET7-HA-Tyk2(C)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa en a) y b); pMET7-HA-Tyk2(C)-RTP51 + pMG1 + pXP2d2-rPAPI-luciferasa en c)). Las barras de error indican la desviación estándar.

Figura 3: Detección de la interacción entre la integrasa (IN) de VIH1 y LEDGF humano y entre p53 humano y MDM2. Las células fueron transfectadas con los siguientes plásmidos:

- a) pMET7-HA-Tyk2(C)-LEDGF + pMG2-IN + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Tyk2(C)-MDM2 + pMG1-p53 + pXP2d2-rPAPI-luciferasa

La actividad de luciferasa se muestra como la inducción múltiple con respecto a la actividad de luciferasa medida en células transfectadas con la misma fusión Tyk2(C), un fragmento de gp130 no condensado y el plásmido informador de luciferasa (pMET7-HA-Tyk2(C)-LEDGF + pMG1 + pXP2d2-rPAPI-luciferasa en a); pMet7-HA-Tyk2(C)-MDM2 + pMG1 + pXP2d2-rPAPI-luciferasa en b)). Las barras de error indican la desviación estándar.

Figura 4: Interrupción dependiente de la dosis de la interacción entre p53 humano y MDM2 por Nutlina-3. Las células fueron transfectadas con los siguientes plásmidos:

- a) pMET7-HA-Tyk2(C)-p53(N) + pMG1 + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Tyk2(C)-p53(N) + pMG1-EFHA1 + pXP2d2-rPAPI-luciferasa
- c) pMET7-HA-Tyk2(C)-p53(N) + pMG2-MDM2 + pXP2d2-rPAPI-luciferasa

Después de la transfección, las células se trataron con 0-14-41-123-370-1111-3333nM Nutlina-3, concentración final. Las barras de error indican la desviación estándar.

Figura 5: Detección de la interacción entre subunidades p66 de transcriptasa inversa (RT) de VIH1 en las variantes de ensayo que comprenden proteínas de fusión de quinasa mutante Jak2 o c-Src. Las células fueron transfectadas con los siguientes plásmidos:

- a) pMET7-HA-Tyk2(C)-RTP66 + pMG2-RTP66 + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Jak2(C)-RTP66 + pMG2-RTP66 + pXP2d2-rPAPI-luciferasa
- c) pMET7-HA-c-Src(K)-RTP66 + pMG2-RTP66 + pXP2d2-rPAPI-luciferasa

La actividad de luciferasa se muestra como la inducción múltiple con respecto a la actividad de luciferasa medida en células transfectadas con la misma fusión RTP66-quinasa, un fragmento de gp130 no condensado y el plásmido informador de luciferasa (pMET7-HA-Tyk2(C)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa en a); pMet7-HA-Jak2(C)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa en b); pMET7-HA-c-Src(K)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa en c)). Las barras de error indican la desviación estándar.

Figura 6: Detección de la interacción entre las subunidades de transcriptasa inversa (RT) de VIH1 en una variante de ensayo inducible. Las células fueron transfectadas con los siguientes plásmidos:

- a) pMET7-HA-Tyk2(C)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Tyk2(C)-RTP66 + pMG2-RTP51 + pXP2d2-rPAPI-luciferasa
- c) pMET7-HA-Tyk2(R1027A)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa
- d) pMET7-HA-Tyk2(R1027A)-RTP66 + pMG2-RTP51 + pXP2d2-rPAPI-luciferasa

Después de la transfección, las células se trataron con imidazol 0-5-50 mM, concentración final. Las barras de error indican la desviación estándar.

EJEMPLOS

Materiales y métodos de la invención

Plásmidos utilizados en los ejemplos

Un primer tipo de plásmidos expresan las proteínas quiméricas que consisten en una porción C-terminal marcada con HA de Tyk2 humana condensada en su extremo C al primer polipéptido interactuante y se generan en el vector pMET7, que contiene un promotor SR α híbrido constitutivo fuerte (Takebe et al., 1988). Para generar el plásmido pMET7-HA-Tyk2(C)-RTP66, la secuencia que codifica el extremo C-terminal de Tyk2 humana que comprende el dominio quinasa (a partir de los aminoácidos 589 y omitiendo el codón de terminación) se amplificó por PCR en ADNc a partir de células HEK293 con cebadores MBU-O-6486-O y MBU-O-6487. Además de una secuencia codificadora de HA el primer cebador contenía un sitio de reconocimiento de la enzima de restricción Apal, mientras que el último cebador contenía un sitio de reconocimiento de la enzima de restricción EcoRI. El amplicón de PCR fue digerido con Apal y EcoRI y fue ligado en el plásmido pMG2-RTP66 cortado con Apal-EcoRI (un plásmido derivado de pMET7 que codifica una proteína quimérica Flag-gp130-RTP66) que contenía los sitios Apal y EcoRI que flanquean las secuencias que codifican la proteína de fusión Flag-gp130 en los extremos 5' y 3', respectivamente (Pattyn et al., 2008). pMET7-HA-Tyk2(C)-RTP51 fue generado mediante el intercambio de la inserción RTP51 de pMG2-RTP51 (Pattyn et al., 2008) con RTP66 de pMET7-HA-Tyk2(C)-RTP66 utilizando los sitios EcoRI y NotI que flanquean estas inserciones en el extremo 5' y 3', respectivamente. LEDGF humana de longitud completa se amplificó por PCR en ADNc de células HeLa utilizando los cebadores MBU-O-3879 y MBU-O-3880 y se intercambió con la inserción RTP66 de pMET7-HA-Tyk2(C)-RTP66 utilizando sitios de restricción EcoRI y NotI para generar pMET7-HA-Tyk2(C)-LEDGF. De manera similar, la secuencia MDM2 humana de longitud completa se amplificó por PCR en un clon de entrada a MDM2 de la colección hORFeome (Lamesch et al., 2007) con los cebadores MBU-O-3912 y MBU-O-3913 y se intercambió con EcoRI-NotI para producir pMET7-HA-Tyk2(C)-MDM2. La región N-terminal de p53 humano (aminoácidos 2-72) se amplificó por PCR en un clon de entrada a p53 de la colección hORFeome (Lamesch et al., 2007) con los cebadores MBU-O-2277 y MBU-O-2273 y se intercambió utilizando enzimas de restricción EcoRI y NotI, proporcionando pMET7-HA-Tyk2(C)-p53(N). El plásmido pMET7-HA-Jak2(C)-RTP66 se generó por PCR amplificando la secuencia que codifica el extremo C-terminal de Jak2 de ratón (desde el aminoácido 535 hasta el extremo de la proteína, dejando fuera el codón de parada) en pRK5-mJak2 (Silvennoinen et al., 1993) con los cebadores MBU-O-6653 y MBU-O-6655. Este fragmento se intercambió con la secuencia que codifica Tyk2(C) a partir de pMET7-HA-Tyk2(C)-RTP66 a través de una digestión de restricción PacI-EcoRI. Del mismo modo, la secuencia codificadora para el dominio quinasa de c-Src (aminoácidos 266-523) se amplificó por PCR en un clon de entrada c-Src de la colección hORFeome (Lamesch et al., 2007) con los cebadores MBU-O-6656 y MBU-O-6657 y se intercambió utilizando enzimas de restricción EcoRI y PacI, generando pMET7-HA-c-Src(K)-RTP66. El plásmido pMET7-HA-Tyk2(R1027A)-RTP66 se generó amplificando primero por PCR la secuencia de Tyk2 humana en ADNc de HEK293 con cebadores MBU-O-7811 y MBU-O-7812 e intercambiando el amplicón resultante por la secuencia Tyk2(C) de pMET7-HA-Tyk2(C)-RTP66 utilizando las enzimas de restricción PacI y EcoRI para producir pMET7-HA-Tyk2-RTP66. A continuación, la mutación R1027A se introdujo por mutagénesis dirigida al sitio utilizando los cebadores MBU-O-7341 y MBU-O-7342.

Los plásmidos que codifican las fusiones con el segundo polipéptido interactuante eran del tipo también utilizado en MAPPIT, designado pMG1 y pMG2 (documento WO0190188, Eyckerman et al., 2001; Lemmens et al., 2003). Estos plásmidos codifican proteínas de fusión del segundo polipéptido interactuante acoplado a un fragmento de la cadena del receptor de citoquina gp130 humana, que contiene múltiples residuos tirosina que, después de la fosforilación, constituyen sitios de reclutamiento para STAT3. RTP66 y RTP51 que contienen plásmidos pMG2-RTP66 y pMG2-RTP51 se han descrito en otra parte (Pattyn et al., 2008). El plásmido pMG1 que codifica un fragmento de receptor gp130 no condensado se obtuvo cortando la inserción de presa MAPPIT de un vector pMG1 utilizando EcoRI y XhoI, haciendo roma la cadena principal del vector a través de Pfu ADN polimerasa y auto-ligamiento. El plásmido pMG2-IN fue construido por PCR amplificando la secuencia que codifica la integrasa de VIH1 en el molde del plásmido pNL4-3 (Adachi et al., 1986) utilizando los cebadores MBU-O-3813 y MBU-O-3814 e intercambiando esta secuencia con la inserción de un vector pMG2 MAPPIT utilizando enzimas de restricción EcoRI y NotI. De manera similar, la región codificadora de MDM2 humano se amplificó por PCR con los cebadores MBU-O-3912 y MBU-O-3913 y se intercambió con EcoRI-NotI para producir pMG2-MDM2. pMG1-p53 y pMG1-EFHA1 fueron generados por transferencia mediada por recombinación Gateway de las secuencias de longitud completa de p53 y EFHA1 humanas de vectores de entrada de la colección hORFeome (Lamesch et al., 2007) en una versión compatible con Gateway del vector pMG1 tal como se describió anteriormente (Lievens et al., 2009).

El plásmido informador pXP2d2-rPAPI-luciferasa utilizado en los ejemplos contiene el promotor rPAPI (Proteína I Asociada a la Pancreatitis de rata) dependiente STAT3 que dirige un gen informador de luciferasa de luciérnaga como se ha descrito previamente (documento WO0190188, Eyckerman et al., 2001).

Procedimiento de transfección

- 5 Las transfecciones se llevaron a cabo utilizando un método de fosfato de calcio estándar. Células HEK293-T fueron sembradas en placas negras de 96 pocillos tratadas con cultivo de tejido a razón de 10.000 células/pocillo en un medio de cultivo de 100 µl (DMEM suplementado con FCS al 10%). Veinticuatro horas más tarde, se prepararon mezclas de ADN de plásmido que contenía los plásmidos que codifican las proteínas de fusión con la primera y segunda proteínas interactuantes y plásmidos informadores. El ADN fue suplementado con 10 µl de CaCl₂ 2,5 M y
- 10 agua doblemente destilada a un volumen final de 100 µl. Esta mezcla se añadió gota a gota a 100 µl de tampón 2xHeBS (NaCl 280 mM, Na₂HPO₄ 1,5 mM, Hepes 50 mM, pH 7,05), al tiempo que se sometían vigorosamente a un vórtex. Después de incubación a temperatura ambiente durante 15 min para permitir que se formaran precipitados de ADN, la disolución se añadió a las células a razón de 10 µl/pocillo. Las células se incubaron a 37°C, 8% de CO₂.
- 15 Cuarenta y ocho horas después de la transfección, se midió la actividad luciferasa utilizando el kit de sistema de ensayo de luciferasa (Promega) en un luminómetro TopCount (Perkin-Elmer). Cada una de las transfecciones se realizó por triplicado y se utilizó la media de las lecturas de actividad de luciferasa en los cálculos. Una disolución madre de Nutlina-3 (Sigma) de 20 mM en DMSO se diluyó en medio de cultivo y se añadió a las células 24 h después de la transfección. Se añadió imidazol (Sigma) diluido en medio de cultivo a las células 24 h después de la transfección.

Código del cebador de oligonucleótidos	SEQ ID N°	Secuencia (5' > 3')
MBU-O-6486	8	CCCGGGCCCCACCATGTATCCATATGATGTTCCAGATTA TGCTTTAATTAAAATCACCCAGCTGTCCCACTTGG
MBU-O-6487	9	GGGGAATTCGCACACGCTGAACACTGAAGG
MBU-O-3879	10	CGTACGAATTCGGGAGCTCGATGACTCGCGATTTCAAA CCTGGAG
MBU-O-3880	11	GGTCATCTAGACCGCGGCCGCTCAGTTATCTAGTGTAG AATCCTTCAG
MBU-O-3912	12	GCGGAATTCATGTGCAATACCAACATGTCTG
MBU-O-3913	13	CGCGCGGCCGCCTAGGGGAAATAAGTTAGCAC
MBU-O-2277	14	GCGAGAATTCGAGGAGCCGAGTCAGATCC
MBU-O-2273	15	CGCTGCGGCCGCTTAGCGGGGAGCAGCCTCTGGC
MBU-O-6653	16	CCCGCGGCCGCTTTAATTAAAATGGTGTTCACAAAATC AG
MBU-O-6655	17	GGGCTCGAGGAATTCGCAGCTATACTGTCCCGG
MBU-O-6656	18	CCCGCGGCCGCTTTAATTAAACCTCGGGAGTCGCTGC GGC

Código del cebador de oligonucleótidos	SEQ ID N°	Secuencia (5' > 3')
MBU-O-6657	19	GGGCTCGAGGAATTCGAAGTAGTCCTCCAGGAAGG
MBU-O-3813	20	CGTACGAATTCGGGAGCTCGTTTTTAGATGGAATAG
MBU-O-3814	21	GGTCATCTAGACCGCGGCCGCTCAATCCTCATCCTGCTC TAC
MBU-O-7811	22	CATTTAATTAAACCTCTGCGCCACTGGGGG
MBU-O-7812	23	CATGAATTCGCACACGCTGAACACTGAAGGG
MBU-O-7341	24	CCGAGACCTAGCCGCGGCCAACGTGCTGC
MBU-O-7342	25	GCAGCACGTTGGCCGCGGCTAGGTCTCGG

Ejemplo 1: Detección de la interacción entre subunidades de transcriptasa inversa (RT) de VIH1

Con el fin de determinar la funcionalidad del ensayo, se sometió a ensayo la interacción entre las subunidades de VIH1 que forman homodímeros y heterodímeros mediante la transfección de las siguientes combinaciones de plásmidos (100 ng de la construcción de fusión Tyk2(C), 1 µg de la construcción de fusión gp130 y 50 ng de la construcción de informador de luciferasa) de acuerdo con los métodos descritos anteriormente:

- a) pMET7-HA-Tyk2(C)-RTP66 + pMG2-RTP51 + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Tyk2(C)-RTP66 + pMG2-RTP66 + pXP2d2-rPAPI-luciferasa
- c) pMET7-HA-Tyk2(C)-RTP51 + pMG2-RTP51 + pXP2d2-rPAPI-luciferasa

Se determinó la señal de fondo para cada uno de los polipéptidos de fusión Tyk2(C) mediante la transfección del plásmido que lo codifica junto con un plásmido que codifica un fragmento de gp130 no condensado (pMG1) y el plásmido informador de luciferasa (pMET7-HA-Tyk2(C)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa en a) y b); pMET7-HA-Tyk2(C)-RTP51 + pMG1 + pXP2d2-rPAPI-luciferasa en c)). La inducción múltiple para cada una de las interacciones testadas se calculó como la relación de la actividad de luciferasa medida en relación con la actividad de luciferasa medida para la señal de fondo correspondiente. Los resultados (Figura 2A) muestran señales claras tanto para los homodímeros como para los heterodímeros de RT. La señal más fuerte era la del heterodímero p66-p51, se obtuvieron señales inferiores, pero aún robustas para los homodímeros p66 y p51. Esta tendencia se corresponde con las afinidades medidas para las interacciones p66-p51, p66-p66 y p51-p51 en ensayos de interacción *in vitro*, que se informó que eran 0,3 µM, 4µM y 230 µM, respectivamente (Venezia et al., 2006). Estos datos ilustran la alta sensibilidad del método.

Ejemplo 2: Interacción entre proteínas nucleares

Para determinar si el método puede detectar interacciones entre proteínas nucleares, los autores de la invención han sometido a ensayo la interacción entre la integrasa (IN) de VIH1 y LEDGF humana y entre p53 y MDM2 humanas, proteínas con una localización nuclear. Las células se transfectaron con las siguientes combinaciones de plásmidos (250 ng de la construcción de fusión Tyk2(C), 500 ng de la construcción de fusión gp130 y 50 ng de la construcción de informador de luciferasa) de acuerdo con los métodos descritos anteriormente:

- a) pMET7-HA-Tyk2(C)-LEDGF + pMG2-IN + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Tyk2(C)-MDM2 + pMG1-p53 + pXP2d2-rPAPI-luciferasa

Se determinó la señal de fondo para cada uno de los polipéptidos de fusión Tyk2(C) mediante la transfección del plásmido que lo codifica junto con un plásmido que codifica un fragmento de gp130 no condensado (pMG1) y el plásmido informador de luciferasa (pMET7-HA-Tyk2(C)-LEDGF + pMG1 + pXP2d2-rPAPI-luciferasa en a); pMET7-HA-Tyk2(C)-MDM2 + pMG1 + pXP2d2-rPAPI-luciferasa en b)). La inducción múltiple para cada una de las interacciones testadas se calculó como la relación de la actividad de luciferasa medida en relación con la actividad de luciferasa medida para la señal de fondo correspondiente. El resultado (Figura 3) demuestra que estas

interacciones dan lugar a señales fuertes, lo que indica que el método es capaz de detectar interacciones en el núcleo.

Ejemplo 3: Interrupción dependiente de la dosis de la interacción entre p53 y MDM2 humana por Nutlina-3

Para demostrar que el método puede detectar la modulación de las interacciones proteína-proteína de moléculas pequeñas los autores de la invención analizaron la interacción entre p53 y MDM2, que se ha informado que es interrumpida por Nutlina-3, un miembro de la familia de nutlinas de potenciales nuevos compuestos contra el cáncer (Vassilev et al., 2004). Las células fueron transfectadas con las siguientes combinaciones de plásmidos (100 ng de la construcción de fusión Tyk2(C), 1000 ng de la construcción de fusión gp130 y 50 ng de la construcción de informador de luciferasa) de acuerdo con los métodos descritos anteriormente:

- a) pMET7-HA-Tyk2(C)-p53(N) + pMG1 + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Tyk2(C)-p53(N) + pMG1-EFHA1 + pXP2d2-rPAPI-luciferasa
- c) pMET7-HA-Tyk2(C)-p53(N) + pMG2-MDM2 + pXP2d2-rPAPI-luciferasa

Después de la transfección, las células fueron tratadas con 0-14-41-123-370-1111-3333nM Nutlina-3, concentración final. Los resultados mostrados en la Figura 4 demuestran una disminución dependiente de la dosis de la señal para la interacción entre p53 (región N-terminal que contiene el dominio de unión de MDM2) y MDM2 tras el tratamiento con Nutlina-3, mientras que la señal correspondiente a la interacción con EFHA1 (que se une a Tyk2(C) en sí) no se ve afectada por este tratamiento. La señal procedente de una transfección de control de combinar pMET7-HA-Tyk2(C)-p53(N) y el plásmido pMG1 que codifica un fragmento de receptor gp130 no condensado es baja y tampoco se ve afectada por el tratamiento de Nutlina-3. Estos datos indican que el método puede detectar cambios dinámicos en los complejos de proteínas que dependen de la adición de componentes exógenos.

Ejemplo 4: Incorporación de otros dominios de quinasa en el método

Para apoyar el hecho de que el método no se limita al uso del dominio quinasa derivado de Tyk2, los autores de la invención han sometido a ensayo proteínas de fusión que comprenden los dominios de quinasa de Jak2 o c-Src. De manera similar a Tyk2, Jak2 pertenece a la familia Jak de tirosina quinasas asociadas al receptor. c-Src pertenece a la familia de tirosina quinasas no receptora de Src. Las células fueron transfectadas con las siguientes combinaciones de plásmidos (250 ng de la construcción de fusión de quinasa, 1000 ng de la construcción de fusión gp130 y 50 ng de la construcción de informador de luciferasa) de acuerdo con los métodos descritos anteriormente:

- a) pMET7-HA-Tyk2(C)-RTP66 + pMG2-RTP66 + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Jak2(C)-RTP66 + pMG2-RTP66 + pXP2d2-rPAPI-luciferasa
- c) pMET7-HA-c-Src(K)-RTP66 + pMG2-RTP66 + pXP2d2-rPAPI-luciferasa

La actividad de luciferasa se muestra como la inducción múltiple con respecto a la actividad de luciferasa medida en células transfectadas con la misma fusión del dominio RTP66-quinasa, un fragmento de gp130 no condensado y el plásmido informador de luciferasa (pMET7-HA-Tyk2(C)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa en a); pMET7-HA-Jak2(C)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa en b); pMET7-HA-c-Src(K)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa en c)). La inducción múltiple para cada una de las interacciones testadas se calculó como la relación de la actividad de luciferasa medida en relación con la actividad de luciferasa medida para la señal de fondo correspondiente. El resultado (Figura 5) indica que la interacción entre las subunidades RTP66 también puede detectarse utilizando proteínas de fusión con dominios de quinasa de otras quinasas distintas de Tyk2.

Ejemplo 5: Versión inducible del método

Una versión inducible del método fue ideada para añadir un nivel adicional de control y para permitir la separación temporal de los eventos de interacción de proteínas y de generación de la señal. Esto último podría ser importante en los casos en que la modificación de los polipéptidos de interacción por parte de la actividad de quinasa prohíbe la interacción. Se ha descrito un mutante de tirosina quinasa Src cuya actividad se hace inducible por la pequeña molécula imidazol por una mutación de arginina a alanina en el centro catalítico de la enzima (Qiao et al., 2006). La mutación puntual correspondiente de esta arginina catalítica conservada (R1027) fue generada en Tyk2, condensada a RTP66 y testada en una versión inducible por imidazol del ensayo. Las células fueron transfectadas con las siguientes combinaciones de plásmidos (250 ng de la construcción de fusión de quinasa, 250 ng de la construcción de fusión gp130 y 50 ng de la construcción informador de luciferasa) de acuerdo con los métodos descritos anteriormente:

- a) pMET7-HA-Tyk2(C)-RTP66 + pMG1 + pXP2d2-rPAPI-luciferasa
- b) pMET7-HA-Tyk2(C)-RTP66 + pMG2-RTP51 + pXP2d2-rPAPI-luciferasa

c) pMET7-HA-Tyk2(R1027A)-RTp66 + pMG1 + pXP2d2-rPAPI-luciferasa

d) pMET7-HA-Tyk2(R1027A)-RTp66 + pMG2-RTp51 + pXP2d2-rPAPI-luciferasa

Después de la transfección, las células fueron tratadas con imidazol 0-5-50 mM, concentración final. Los resultados se representan en la Figura 5, que muestra una señal inducible por imidazol específicamente en las células que expresan una fusión que contiene la arginina Tyk2 a la mutación de alanina. La inducción máxima dependiente de imidazol de la señal indicativa de la interacción entre el VIH y RTp66 RTp51 fue de 36 veces en comparación con células no tratadas. Las células transfectadas con la fusión Tyk2(C) descritas anteriormente no mostraron regulación alguna de imidazol. La señal de fondo observada en las células transfectadas con combinaciones de la construcción de fusión Tyk2 y un plásmido que codifica un fragmento de gp130 no condensado (pMG1) de manera similar no se vio afectada por el tratamiento con imidazol. Estos resultados indican que es posible controlar el ensayo de interacción mediante la aplicación de una quinasas químicamente inducible.

REFERENCIAS

- 15 - Adachi, A., Gendelman, H.E., Koenig, S., Folks, T., Willey, R., Rabson, A. y Martin, M.A. (1986). Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J Virol* 59, 284-291.
- Caligiuri, M., Molz, L., Liu, Q., Kaplan, F., Xu, J.P., Majeti, J.Z., Ramos-Kelsey, R., Murthi, K., Lievens, S., Tavernier, J. y Kley, N (2006). MASPIT: three-hybrid trap for quantitative fingerprinting of small molecule-protein interactions in mammalian cells. *Chem. Biol.* 13, 711-722.
- 20 - Eyckerman, S., Verhee, A., Van der Heyden, J., Lemmens, I., Van Ostade, X., Vandekerckhove, J. y Tavernier, J. (2001). Design and application of a cytokine-receptor-based interaction trap. *Nature Cell Biology* 3, 1114-1119.
- Fields, S. y Song, O.K. (1989). A novel genetic system to detect protein-protein interactions. *Nature*, 340, 245 - 246.
- Lamesch, P., Li, N., Milstein, S., Fan, C., Hao, T., Szabo, G., Hu, Z., Venkatesan, K., Bethel, G., Martin, P., et al. (2007). hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. *Genomics* 89, 307-315.
- 25 - Lemmens, I., Eyckerman, S., Zabeau, L., Catteeuw, D., Vertenten, E., Verschueren, K., Huylebroeck, D., Vandekerckhove, J. y Tavernier, J. (2003). Heteromeric MAPPIT: a novel strategy to study modification-dependent proteinprotein interactions in mammalian cells. *Nucleic Acids Research* 31.
- 30 - Lievens, S., Vanderroost, N., Van der Heyden, J., Gesellchen, V., Vidal, M. y Tavernier, J. (2009). Array MAPPIT: high-throughput interactome analysis in mammalian cells. *J Proteome Res* 8, 877-886.
- Pattyn, E., Lavens, D., Van der Heyden, J., Verhee, A., Lievens, S., Lemmens, I., Hallenberger, S., Jochmans, D. y Tavernier, J. (2008). MAPPIT (MAMmalian Protein-Protein Interaction Trap) as a tool to study HIV reverse transcriptase dimerization in intact human cells. *Journal of Virological Methods* 153, 7-15.
- 35 - Qiao, Y., Molina, H., Pandey, A., Zhang, J. y Cole, P.A. (2006). Chemical rescue of a mutant enzyme in living cells. *Science* 311, 1293-1297.
- Silvennoinen, O., Witthuhn, B.A., Quelle, F.W., Cleveland, J.L., Yi, T. e Ihle, J.N. (1993). Structure of the murine Jak2 protein-tyrosine kinase and its role in interleukin 3 signal transduction. *Proc Natl Acad Sci U S A* 90, 8429-8433.
- 40 - Takebe, Y., Seiki, M., Fujisawa, J., Hoy, P., Yokota, K., Arai, K., Yoshida, M. y Arai, N. (1988). SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol Cell Biol* 8, 466-472.
- Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., et al. (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* 303, 844-848.
- 45 - Venezia, C.F., Howard, K.J., Ignatov, M.E., Holladay, L.A. y Barkley, M.D. (2006). Effects of efavirenz binding on the subunit equilibria of HIV-1 reverse transcriptase. *Biochemistry* 45, 2779-2789.

LISTADO DE SECUENCIAS

<110> VIB VZW UNIVERSITEIT GENT

<120> SENSOR DEL SUSTRATO DE QUINASA

5 <130> JT/TYK2/361

<150> GB 1103453.5

<151> 01-03-2011

<150> US 61/464,285

<151> 01-03-2011

10 <160> 27

<170> PatentIn version 3.5

<210> 1

15 <211> 599

<212> PRT

<213> Homo sapiens

<400> 1

20

```

Ile Thr Gln Leu Ser His Leu Gly Gln Gly Thr Arg Thr Asn Val Tyr
1          5          10          15

Glu Gly Arg Leu Arg Val Glu Gly Ser Gly Asp Pro Glu Glu Gly Lys
20          25          30

Met Asp Asp Glu Asp Pro Leu Val Pro Gly Arg Asp Arg Gly Gln Glu
35          40          45

Leu Arg Val Val Leu Lys Val Leu Asp Pro Ser His His Asp Ile Ala
50          55          60

Leu Ala Phe Tyr Glu Thr Ala Ser Leu Met Ser Gln Val Ser His Thr
65          70          75          80

His Leu Ala Phe Val His Gly Val Cys Val Arg Gly Pro Glu Asn Ile
85          90          95

Met Val Thr Glu Tyr Val Glu His Gly Pro Leu Asp Val Trp Leu Arg
100         105         110

Arg Glu Arg Gly His Val Pro Met Ala Trp Lys Met Val Val Ala Gln
115         120         125

Gln Leu Ala Ser Ala Leu Ser Tyr Leu Glu Asn Lys Asn Leu Val His
130         135         140

Gly Asn Val Cys Gly Arg Asn Ile Leu Leu Ala Arg Leu Gly Leu Ala
145         150         155         160

```

ES 2 617 783 T3

Glu Gly Thr Ser Pro Phe Ile Lys Leu Ser Asp Pro Gly Val Gly Leu
 165 170 175
 Gly Ala Leu Ser Arg Glu Glu Arg Val Glu Arg Ile Pro Trp Leu Ala
 180 185 190
 Pro Glu Cys Leu Pro Gly Gly Ala Asn Ser Leu Ser Thr Ala Met Asp
 195 200 205
 Lys Trp Gly Phe Gly Ala Thr Leu Leu Glu Ile Cys Phe Asp Gly Glu
 210 215 220
 Ala Pro Leu Gln Ser Arg Ser Pro Ser Glu Lys Glu His Phe Tyr Gln
 225 230 235 240
 Arg Gln His Arg Leu Pro Glu Pro Ser Cys Pro Gln Leu Ala Thr Leu
 245 250 255
 Thr Ser Gln Cys Leu Thr Tyr Glu Pro Thr Gln Arg Pro Ser Phe Arg
 260 265 270
 Thr Ile Leu Arg Asp Leu Thr Arg Leu Gln Pro His Asn Leu Ala Asp
 275 280 285
 Val Leu Thr Val Asn Pro Asp Ser Pro Ala Ser Asp Pro Thr Val Phe
 290 295 300
 His Lys Arg Tyr Leu Lys Lys Ile Arg Asp Leu Gly Glu Gly His Phe
 305 310 315 320
 Gly Lys Val Ser Leu Tyr Cys Tyr Asp Pro Thr Asn Asp Gly Thr Gly
 325 330 335
 Glu Met Val Ala Val Lys Ala Leu Lys Ala Asp Cys Gly Pro Gln His
 340 345 350
 Arg Ser Gly Trp Lys Gln Glu Ile Asp Ile Leu Arg Thr Leu Tyr His
 355 360 365
 Glu His Ile Ile Lys Tyr Lys Gly Cys Cys Glu Asp Gln Gly Glu Lys
 370 375 380
 Ser Leu Gln Leu Val Met Glu Tyr Val Pro Leu Gly Ser Leu Arg Asp
 385 390 395 400
 Tyr Leu Pro Arg His Ser Ile Gly Leu Ala Gln Leu Leu Leu Phe Ala
 405 410 415

ES 2 617 783 T3

Gln Gln Ile Cys Glu Gly Met Ala Tyr Leu His Ala Gln His Tyr Ile
420 425 430

His Arg Asp Leu Ala Ala Arg Asn Val Leu Leu Asp Asn Asp Arg Leu
435 440 445

Val Lys Ile Gly Asp Phe Gly Leu Ala Lys Ala Val Pro Glu Gly His
450 455 460

Glu Tyr Tyr Arg Val Arg Glu Asp Gly Asp Ser Pro Val Phe Trp Tyr
465 470 475 480

Ala Pro Glu Cys Leu Lys Glu Tyr Lys Phe Tyr Tyr Ala Ser Asp Val
485 490 495

Trp Ser Phe Gly Val Thr Leu Tyr Glu Leu Leu Thr His Cys Asp Ser
500 505 510

Ser Gln Ser Pro Pro Thr Lys Phe Leu Glu Leu Ile Gly Ile Ala Gln
515 520 525

Gly Gln Met Thr Val Leu Arg Leu Thr Glu Leu Leu Glu Arg Gly Glu
530 535 540

Arg Leu Pro Arg Pro Asp Lys Cys Pro Cys Glu Val Tyr His Leu Met
545 550 555 560

Lys Asn Cys Trp Glu Thr Glu Ala Ser Phe Arg Pro Thr Phe Glu Asn
565 570 575

Leu Ile Pro Ile Leu Lys Thr Val His Glu Lys Tyr Gln Gly Gln Ala
580 585 590

Pro Ser Val Phe Ser Val Cys
595

<210> 2
<211> 304
5 <212> PRT
<213> Homo sapiens

<400> 2

Ser Pro Ala Ser Asp Pro Thr Val Phe His Lys Arg Tyr Leu Lys Lys
1 5 10 15

Ile Arg Asp Leu Gly Glu Gly His Phe Gly Lys Val Ser Leu Tyr Cys
20 25 30

Tyr Asp Pro Thr Asn Asp Gly Thr Gly Glu Met Val Ala Val Lys Ala
35 40 45

ES 2 617 783 T3

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

ES 2 617 783 T3

<210> 3

<211> 159

<212> PRT

5 <213> Homo sapiens

<400> 3

Ser Thr Val Val His Ser Gly Tyr Arg His Gln Val Pro Ser Val Gln
1 5 10 15

Val Phe Ser Arg Ser Glu Ser Thr Gln Pro Leu Leu Asp Ser Glu Glu
20 25 30

Arg Pro Glu Asp Leu Gln Leu Val Asp His Val Asp Gly Gly Asp Gly
35 40 45

Ile Leu Pro Arg Gln Gln Tyr Phe Lys Gln Asn Cys Ser Gln His Glu
50 55 60

Ser Ser Pro Asp Ile Ser His Phe Glu Arg Ser Lys Gln Val Ser Ser
65 70 75 80

Val Asn Glu Glu Asp Phe Val Arg Leu Lys Gln Gln Ile Ser Asp His
85 90 95

Ile Ser Gln Ser Cys Gly Ser Gly Gln Met Lys Met Phe Gln Glu Val
100 105 110

Ser Ala Ala Asp Ala Phe Gly Pro Gly Thr Glu Gly Gln Val Glu Arg
115 120 125

Phe Glu Thr Val Gly Met Glu Ala Ala Thr Asp Glu Gly Met Pro Lys
130 135 140

Ser Tyr Leu Pro Gln Thr Val Arg Gln Gly Gly Tyr Met Pro Gln
145 150 155

10

<210> 4

<211> 187

<212> PRT

15 <213> Homo sapiens

<400> 4

ES 2 617 783 T3

Ser Thr Val Val His Ser Gly Tyr Arg His Gln Val Pro Ser Val Gln
1 5 10 15

Val Phe Ser Arg Ser Glu Ser Thr Gln Pro Leu Leu Asp Ser Glu Glu
20 25 30

Arg Pro Glu Asp Leu Gln Leu Val Asp His Val Asp Gly Gly Asp Gly
35 40 45
Ile Leu Pro Arg Gln Gln Tyr Phe Lys Gln Asn Cys Ser Gln His Glu
50 55 60

Ser Ser Pro Asp Ile Ser His Phe Glu Arg Ser Lys Gln Val Ser Ser
65 70 75 80

Val Asn Glu Glu Asp Phe Val Arg Leu Lys Gln Gln Ile Ser Asp His
85 90 95

Ile Ser Gln Ser Cys Gly Ser Gly Gln Met Lys Met Phe Gln Glu Val
100 105 110

Ser Ala Ala Asp Ala Phe Gly Pro Gly Thr Glu Gly Gln Val Glu Arg
115 120 125

Phe Glu Thr Val Gly Met Glu Ala Ala Thr Asp Glu Gly Met Pro Lys
130 135 140

Ser Tyr Leu Pro Gln Thr Val Arg Gln Gly Gly Tyr Met Pro Gln Gly
145 150 155 160

Gly Ser Glu Leu Ser Thr Ser Leu Tyr Lys Lys Ala Gly Tyr Leu Pro
165 170 175

Gln Thr Val Arg Gln Gly Gly Tyr Met Pro Gln
180 185

5 <210> 5
<211> 598
<212> PRT
<213> Mus musculus

10 <400> 5

ES 2 617 783 T3

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

ES 2 617 783 T3

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

ES 2 617 783 T3

340	345	350
Ser Thr Glu Glu His Leu Arg Asp Phe Glu Arg Glu Ile Glu Ile Leu 355 360 365		
Lys Ser Leu Gln His Asp Asn Ile Val Lys Tyr Lys Gly Val Cys Tyr 370 375 380		
Ser Ala Gly Arg Arg Asn Leu Arg Leu Ile Met Glu Tyr Leu Pro Tyr 385 390 395 400		
Gly Ser Leu Arg Asp Tyr Leu Gln Lys His Lys Glu Arg Ile Asp His 405 410 415		
Lys Lys Leu Leu Gln Tyr Thr Ser Gln Ile Cys Lys Gly Met Glu Tyr 420 425 430		
Leu Gly Thr Lys Arg Tyr Ile His Arg Asp Leu Ala Thr Arg Asn Ile 435 440 445		
Leu Val Glu Asn Glu Asn Arg Val Lys Ile Gly Asp Phe Gly Leu Thr 450 455 460		
Lys Val Leu Pro Gln Asp Lys Glu Tyr Tyr Lys Val Lys Glu Pro Gly 465 470 475 480		
Glu Ser Pro Ile Phe Trp Tyr Ala Pro Glu Ser Leu Thr Glu Ser Lys 485 490 495		
Phe Ser Val Ala Ser Asp Val Trp Ser Phe Gly Val Val Leu Tyr Glu 500 505 510		
Leu Phe Thr Tyr Ile Glu Lys Ser Lys Ser Pro Pro Val Glu Phe Met 515 520 525		
Arg Met Ile Gly Asn Asp Lys Gln Gly Gln Met Ile Val Phe His Leu 530 535 540		
Ile Glu Leu Leu Lys Ser Asn Gly Arg Leu Pro Arg Pro Glu Gly Cys 545 550 555 560		
Pro Asp Glu Ile Tyr Val Ile Met Thr Glu Cys Trp Asn Asn Asn Val 565 570 575		
Ser Gln Arg Pro Ser Phe Arg Asp Leu Ser Leu Arg Val Asp Gln Ile 580 585 590		
Arg Asp Ser Ile Ala Ala 595		

<210> 6

<211> 308

5 <212> PRT

<213> Mus musculus

<400> 6

```

Met Arg Ile Gly Ala Leu Gly Phe Ser Gly Ala Phe Glu Asp Arg Asp
 1           5           10           15

Pro Thr Gln Phe Glu Glu Arg His Leu Lys Phe Leu Gln Gln Leu Gly
      20           25           30

Lys Gly Asn Phe Gly Ser Val Glu Met Cys Arg Tyr Asp Pro Leu Gln
      35           40           45

Asp Asn Thr Gly Glu Val Val Ala Val Lys Lys Leu Gln His Ser Thr
      50           55           60

Glu Glu His Leu Arg Asp Phe Glu Arg Glu Ile Glu Ile Leu Lys Ser
65           70           75           80

Leu Gln His Asp Asn Ile Val Lys Tyr Lys Gly Val Cys Tyr Ser Ala
      85           90           95

Gly Arg Arg Asn Leu Arg Leu Ile Met Glu Tyr Leu Pro Tyr Gly Ser
      100          105          110

Leu Arg Asp Tyr Leu Gln Lys His Lys Glu Arg Ile Asp His Lys Lys
      115          120          125

Leu Leu Gln Tyr Thr Ser Gln Ile Cys Lys Gly Met Glu Tyr Leu Gly
      130          135          140

Thr Lys Arg Tyr Ile His Arg Asp Leu Ala Thr Arg Asn Ile Leu Val
145          150          155          160

Glu Asn Glu Asn Arg Val Lys Ile Gly Asp Phe Gly Leu Thr Lys Val
      165          170          175

Leu Pro Gln Asp Lys Glu Tyr Tyr Lys Val Lys Glu Pro Gly Glu Ser
      180          185          190

Pro Ile Phe Trp Tyr Ala Pro Glu Ser Leu Thr Glu Ser Lys Phe Ser
      195          200          205

Val Ala Ser Asp Val Trp Ser Phe Gly Val Val Leu Tyr Glu Leu Phe
210          215          220

```

ES 2 617 783 T3

Thr Tyr Ile Glu Lys Ser Lys Ser Pro Pro Val Glu Phe Met Arg Met
225 230 235 240

Ile Gly Asn Asp Lys Gln Gly Gln Met Ile Val Phe His Leu Ile Glu
245 250 255

Leu Leu Lys Ser Asn Gly Arg Leu Pro Arg Pro Glu Gly Cys Pro Asp
260 265 270

Glu Ile Tyr Val Ile Met Thr Glu Cys Trp Asn Asn Asn Val Ser Gln
275 280 285

Arg Pro Ser Phe Arg Asp Leu Ser Leu Arg Val Asp Gln Ile Arg Asp
290 295 300

Ser Ile Ala Ala
305

<210> 7
<211> 258
<212> PRT
<213> Homo sapiens

<400> 7

Pro Arg Glu Ser Leu Arg Leu Glu Val Lys Leu Gly Gln Gly Cys Phe
1 5 10 15

Gly Glu Val Trp Met Gly Thr Trp Asn Gly Thr Thr Arg Val Ala Ile
20 25 30

Lys Thr Leu Lys Pro Gly Thr Met Ser Pro Glu Ala Phe Leu Gln Glu
35 40 45

Ala Gln Val Met Lys Lys Leu Arg His Glu Lys Leu Val Gln Leu Tyr
50 55 60

Ala Val Val Ser Glu Glu Pro Ile Tyr Ile Val Thr Glu Tyr Met Ser
65 70 75 80

Lys Gly Ser Leu Leu Asp Phe Leu Lys Gly Glu Thr Gly Lys Tyr Leu
85 90 95

Arg Leu Pro Gln Leu Val Asp Met Ala Ala Gln Ile Ala Ser Gly Met
100 105 110

Ala Tyr Val Glu Arg Met Asn Tyr Val His Arg Asp Leu Arg Ala Ala
115 120 125

Asn Ile Leu Val Gly Glu Asn Leu Val Cys Lys Val Ala Asp Phe Gly
130 135 140

ES 2 617 783 T3

Leu Ala Arg Leu Ile Glu Asp Asn Glu Tyr Thr Ala Arg Gln Gly Ala
 145 150 155 160
 Lys Phe Pro Ile Lys Trp Thr Ala Pro Glu Ala Ala Leu Tyr Gly Arg
 165 170 175
 Phe Thr Ile Lys Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Thr Glu
 180 185 190
 Leu Thr Thr Lys Gly Arg Val Pro Tyr Pro Gly Met Val Asn Arg Glu
 195 200 205
 Val Leu Asp Gln Val Glu Arg Gly Tyr Arg Met Pro Cys Pro Pro Glu
 210 215 220
 Cys Pro Glu Ser Leu His Asp Leu Met Cys Gln Cys Trp Arg Lys Glu
 225 230 235 240
 Pro Glu Glu Arg Pro Thr Phe Glu Tyr Leu Gln Ala Phe Leu Glu Asp
 245 250 255

5 Tyr Phe

<210> 8
 <211> 73
 <212> ADN

10 <213> Secuencia Artificial

<220>
 <223> Cebador

15 <400> 8

cccgggccca ccatgtatcc atatgatgtt ccagattatg ctttaattaa aatcacccag 60

ctgtcccact tgg 73

<210> 9
 <211> 30
 <212> ADN
 <213> Secuencia Artificial

25 <220>
 <223> Cebador

<400> 9

ggggaattcg cacacgtga aactgaagg 30

30 <210> 10
 <211> 45
 <212> ADN
 <213> Secuencia Artificial

35 <220>

<223> Cebador

<400> 10
 5 cgtacgaatt cgggagctcg atgactcgcg atttcaaacc tggag 45

<210> 11
 <211> 48
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <223> Cebador

<400> 11
 15 ggatcatctag accgcggccg ctgagttatc tagttagaa tcttcag 48

<210> 12
 <211> 31
 20 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Cebador

25 <400> 12

gcggaattca tgtgcaatac caacatgtct g 31

30 <210> 13
 <211> 32
 <212> ADN
 <213> Secuencia Artificial

35 <220>
 <223> Cebador

<400> 13

40 cgcgcggccg cctaggggaa ataagttagc ac 32

<210> 14
 <211> 30
 <212> ADN
 45 <213> Secuencia Artificial

<220>
 <223> Cebador

50 <400> 14

gcgagaattc gaggagccgc agtcagatcc 30

<210> 15
 55 <211> 34
 <212> ADN
 <213> Secuencia Artificial

<220>
 60 <223> Cebador

	<400> 15	
	cgctgcggcc gcttagcggg gagcagcctc tggc	34
5	<210> 16 <211> 41 <212> ADN <213> Secuencia Artificial	
10	<220> <223> Cebador	
	<400> 16 cccgcggccg cttaattaa aatggtgtt caaaaatca g	41
15	<210> 17 <211> 34 <212> ADN <213> Secuencia Artificial	
20	<220> <223> Cebador	
	<400> 17 gggctcgagg aattccgcag ctatactgtc ccgg	34
25	<210> 18 <211> 40 <212> ADN <213> Secuencia Artificial	
30	<220> <223> Cebador	
	<400> 18 cccgcggccg cttaattaa acctcgggag tcgctgcggc	40
35	<210> 19 <211> 35 <212> ADN <213> Secuencia Artificial	
40	<220> <223> Cebador	
	<400> 19 gggctcgagg aattcgaagt agtcctccag gaagg	35
45	<210> 20 <211> 36 <212> ADN <213> Secuencia Artificial	
50	<220> <223> Cebador	
	<400> 20 cgtacgaatt cgggagctcg ttttagatg gaatag	36
55	<210> 21 <211> 41	
60		

<212> ADN
 <213> Secuencia Artificial

 <220>
 5 <223> Cebador
 <400> 21
 ggtcatctag accgcggccg ctcaatcctc atcctgtcta c 41

 <210> 22
 10 <211> 30
 <212> ADN
 <213> Secuencia Artificial

 <220>
 15 <223> Cebador

 <400> 22
 catttaatta aacctctgcg ccactggggg 30

 <210> 23
 20 <211> 31
 <212> ADN
 <213> Secuencia Artificial

 <220>
 25 <223> Cebador

 <400> 23
 catgaattcg cacacgctga acactgaagg g 31
 30
 <210> 24
 <211> 29
 <212> ADN
 <213> Secuencia Artificial
 35
 <220>
 <223> Cebador

 <400> 24
 40 ccgagacctg gccgcggcca acgtgctgc 29

 <210> 25
 <211> 29
 <212> ADN
 45 <213> Secuencia Artificial

 <220>
 <223> Cebador

 <400> 25
 50 gcagcacgtt ggccgcggct aggtctcgg 29

 <210> 26
 <211> 1187
 55 <212> PRT
 <213> Homo sapiens

 <400> 26

 60 Met Pro Leu Arg His Trp Gly Met Ala Arg Gly Ser Lys Pro Val Gly

ES 2 617 783 T3

1	5	10	15
Asp Gly Ala Gln Pro Met Ala Ala Met Gly Gly Leu Lys Val Leu Leu	20	25	30
His Trp Ala Gly Pro Gly Gly Gly Glu Pro Trp Val Thr Phe Ser Glu	35	40	45
Ser Ser Leu Thr Ala Glu Glu Val Cys Ile His Ile Ala His Lys Val	50	55	60
Gly Ile Thr Pro Pro Cys Phe Asn Leu Phe Ala Leu Phe Asp Ala Gln	65	70	75
Ala Gln Val Trp Leu Pro Pro Asn His Ile Leu Glu Ile Pro Arg Asp	85	90	95
Ala Ser Leu Met Leu Tyr Phe Arg Ile Arg Phe Tyr Phe Arg Asn Trp	100	105	110
His Gly Met Asn Pro Arg Glu Pro Ala Val Tyr Arg Cys Gly Pro Pro	115	120	125
Gly Thr Glu Ala Ser Ser Asp Gln Thr Ala Gln Gly Met Gln Leu Leu	130	135	140
Asp Pro Ala Ser Phe Glu Tyr Leu Phe Glu Gln Gly Lys His Glu Phe	145	150	155
Val Asn Asp Val Ala Ser Leu Trp Glu Leu Ser Thr Glu Glu Glu Ile	165	170	175
His His Phe Lys Asn Glu Ser Leu Gly Met Ala Phe Leu His Leu Cys	180	185	190
His Leu Ala Leu Arg His Gly Ile Pro Leu Glu Glu Val Ala Lys Lys	195	200	205
Thr Ser Phe Lys Asp Cys Ile Pro Arg Ser Phe Arg Arg His Ile Arg	210	215	220
Gln His Ser Ala Leu Thr Arg Leu Arg Leu Arg Asn Val Phe Arg Arg	225	230	235
Phe Leu Arg Asp Phe Gln Pro Gly Arg Leu Ser Gln Gln Met Val Met	245	250	255
Val Lys Tyr Leu Ala Thr Leu Glu Arg Leu Ala Pro Arg Phe Gly Thr	260	265	270

ES 2 617 783 T3

Glu Arg Val Pro Val Cys His Leu Arg Leu Leu Ala Gln Ala Glu Gly
 275 280 285
 Glu Pro Cys Tyr Ile Arg Asp Ser Gly Val Ala Pro Thr Asp Pro Gly
 290 295 300
 Pro Glu Ser Ala Ala Gly Pro Pro Thr His Glu Val Leu Val Thr Gly
 305 310 315 320
 Thr Gly Gly Ile Gln Trp Trp Pro Val Glu Glu Glu Val Asn Lys Glu
 325 330 335
 Glu Gly Ser Ser Gly Ser Ser Gly Arg Asn Pro Gln Ala Ser Leu Phe
 340 345 350
 Gly Lys Lys Ala Lys Ala His Lys Ala Val Gly Gln Pro Ala Asp Arg
 355 360 365
 Pro Arg Glu Pro Leu Trp Ala Tyr Phe Cys Asp Phe Arg Asp Ile Thr
 370 375 380
 His Val Val Leu Lys Glu His Cys Val Ser Ile His Arg Gln Asp Asn
 385 390 395 400
 Lys Cys Leu Glu Leu Ser Leu Pro Ser Arg Ala Ala Ala Leu Ser Phe
 405 410 415
 Val Ser Leu Val Asp Gly Tyr Phe Arg Leu Thr Ala Asp Ser Ser His
 420 425 430
 Tyr Leu Cys His Glu Val Ala Pro Pro Arg Leu Val Met Ser Ile Arg
 435 440 445
 Asp Gly Ile His Gly Pro Leu Leu Glu Pro Phe Val Gln Ala Lys Leu
 450 455 460
 Arg Pro Glu Asp Gly Leu Tyr Leu Ile His Trp Ser Thr Ser His Pro
 465 470 475 480
 Tyr Arg Leu Ile Leu Thr Val Ala Gln Arg Ser Gln Ala Pro Asp Gly
 485 490 495
 Met Gln Ser Leu Arg Leu Arg Lys Phe Pro Ile Glu Gln Gln Asp Gly
 500 505 510
 Ala Phe Val Leu Glu Gly Trp Gly Arg Ser Phe Pro Ser Val Arg Glu
 515 520 525

ES 2 617 783 T3

Leu Gly Ala Ala Leu Gln Gly Cys Leu Leu Arg Ala Gly Asp Asp Cys
 530 535 540

Phe Ser Leu Arg Arg Cys Cys Leu Pro Gln Pro Gly Glu Thr Ser Asn
 545 550 555 560

Leu Ile Ile Met Arg Gly Ala Arg Ala Ser Pro Arg Thr Leu Asn Leu
 565 570 575

Ser Gln Leu Ser Phe His Arg Val Asp Gln Lys Glu Ile Thr Gln Leu
 580 585 590

Ser His Leu Gly Gln Gly Thr Arg Thr Asn Val Tyr Glu Gly Arg Leu
 595 600 605

Arg Val Glu Gly Ser Gly Asp Pro Glu Glu Gly Lys Met Asp Asp Glu
 610 615 620

Asp Pro Leu Val Pro Gly Arg Asp Arg Gly Gln Glu Leu Arg Val Val
 625 630 635 640

Leu Lys Val Leu Asp Pro Ser His His Asp Ile Ala Leu Ala Phe Tyr
 645 650 655

Glu Thr Ala Ser Leu Met Ser Gln Val Ser His Thr His Leu Ala Phe
 660 665 670

Val His Gly Val Cys Val Arg Gly Pro Glu Asn Ile Met Val Thr Glu
 675 680 685

Tyr Val Glu His Gly Pro Leu Asp Val Trp Leu Arg Arg Glu Arg Gly
 690 695 700

His Val Pro Met Ala Trp Lys Met Val Val Ala Gln Gln Leu Ala Ser
 705 710 715 720

Ala Leu Ser Tyr Leu Glu Asn Lys Asn Leu Val His Gly Asn Val Cys
 725 730 735

Gly Arg Asn Ile Leu Leu Ala Arg Leu Gly Leu Ala Glu Gly Thr Ser
 740 745 750

Pro Phe Ile Lys Leu Ser Asp Pro Gly Val Gly Leu Gly Ala Leu Ser
 755 760 765

Arg Glu Glu Arg Val Glu Arg Ile Pro Trp Leu Ala Pro Glu Cys Leu
 770 775 780

ES 2 617 783 T3

Pro Gly Gly Ala Asn Ser Leu Ser Thr Ala Met Asp Lys Trp Gly Phe
 785 790 795 800
 Gly Ala Thr Leu Leu Glu Ile Cys Phe Asp Gly Glu Ala Pro Leu Gln
 805 810 815
 Ser Arg Ser Pro Ser Glu Lys Glu His Phe Tyr Gln Arg Gln His Arg
 820 825 830
 Leu Pro Glu Pro Ser Cys Pro Gln Leu Ala Thr Leu Thr Ser Gln Cys
 835 840 845
 Leu Thr Tyr Glu Pro Thr Gln Arg Pro Ser Phe Arg Thr Ile Leu Arg
 850 855 860
 Asp Leu Thr Arg Leu Gln Pro His Asn Leu Ala Asp Val Leu Thr Val
 865 870 875 880
 Asn Pro Asp Ser Pro Ala Ser Asp Pro Thr Val Phe His Lys Arg Tyr
 885 890 895
 Leu Lys Lys Ile Arg Asp Leu Gly Glu Gly His Phe Gly Lys Val Ser
 900 905 910
 Leu Tyr Cys Tyr Asp Pro Thr Asn Asp Gly Thr Gly Glu Met Val Ala
 915 920 925
 Val Lys Ala Leu Lys Ala Asp Cys Gly Pro Gln His Arg Ser Gly Trp
 930 935 940
 Lys Gln Glu Ile Asp Ile Leu Arg Thr Leu Tyr His Glu His Ile Ile
 945 950 955 960
 Lys Tyr Lys Gly Cys Cys Glu Asp Gln Gly Glu Lys Ser Leu Gln Leu
 965 970 975
 Val Met Glu Tyr Val Pro Leu Gly Ser Leu Arg Asp Tyr Leu Pro Arg
 980 985 990
 His Ser Ile Gly Leu Ala Gln Leu Leu Leu Phe Ala Gln Gln Ile Cys
 995 1000 1005
 Glu Gly Met Ala Tyr Leu His Ala Gln His Tyr Ile His Arg Asp
 1010 1015 1020
 Leu Ala Ala Arg Asn Val Leu Leu Asp Asn Asp Arg Leu Val Lys
 1025 1030 1035
 Ile Gly Asp Phe Gly Leu Ala Lys Ala Val Pro Glu Gly His Glu

ES 2 617 783 T3

1040	1045	1050	
Tyr Tyr Arg Val Arg Glu Asp Gly Asp Ser Pro Val Phe Trp Tyr	1055	1060	1065
Ala Pro Glu Cys Leu Lys Glu Tyr Lys Phe Tyr Tyr Ala Ser Asp	1070	1075	1080
Val Trp Ser Phe Gly Val Thr Leu Tyr Glu Leu Leu Thr His Cys	1085	1090	1095
Asp Ser Ser Gln Ser Pro Pro Thr Lys Phe Leu Glu Leu Ile Gly	1100	1105	1110
Ile Ala Gln Gly Gln Met Thr Val Leu Arg Leu Thr Glu Leu Leu	1115	1120	1125
Glu Arg Gly Glu Arg Leu Pro Arg Pro Asp Lys Cys Pro Cys Glu	1130	1135	1140
Val Tyr His Leu Met Lys Asn Cys Trp Glu Thr Glu Ala Ser Phe	1145	1150	1155
Arg Pro Thr Phe Glu Asn Leu Ile Pro Ile Leu Lys Thr Val His	1160	1165	1170
Glu Lys Tyr Gln Gly Gln Ala Pro Ser Val Phe Ser Val Cys	1175	1180	1185

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

<210> 27
 <211> 522
 <212> PRT
 <213> Homo sapiens

<400> 27

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

5

Tyr Ile Val Thr Glu Tyr Met Ser Lys Gly Ser Leu Leu Asp Phe Leu
 340 345 350

Lys Gly Glu Thr Gly Lys Tyr Leu Arg Leu Pro Gln Leu Val Asp Met
 355 360 365

Ala Ala Gln Ile Ala Ser Gly Met Ala Tyr Val Glu Arg Met Asn Tyr
 370 375 380

Val His Arg Asp Leu Arg Ala Ala Asn Ile Leu Val Gly Glu Asn Leu
 385 390 395 400

Val Cys Lys Val Ala Asp Phe Gly Leu Ala Arg Leu Ile Glu Asp Asn
 405 410 415

Glu Tyr Thr Ala Arg Gln Gly Ala Lys Phe Pro Ile Lys Trp Thr Ala
 420 425 430

Pro Glu Ala Ala Leu Tyr Gly Arg Phe Thr Ile Lys Ser Asp Val Trp
 435 440 445

Ser Phe Gly Ile Leu Leu Thr Glu Leu Thr Thr Lys Gly Arg Val Pro
 450 455 460

Tyr Pro Gly Met Val Asn Arg Glu Val Leu Asp Gln Val Glu Arg Gly
 465 470 475 480

Tyr Arg Met Pro Cys Pro Pro Glu Cys Pro Glu Ser Leu His Asp Leu
 485 490 495

Met Cys Gln Cys Trp Arg Lys Glu Pro Glu Glu Arg Pro Thr Phe Glu
 500 505 510

Tyr Leu Gln Ala Phe Leu Glu Asp Tyr Phe
 515 520

REIVINDICACIONES

1. Un complejo de proteínas citoplasmáticas que comprende (a) una primera proteína de fusión recombinante que comprende una quinasa mutante condensada a un primer polipéptido de interacción, en donde dicha quinasa mutante se selecciona del grupo que consiste en una tirosina quinasa constitutiva mutante y una tirosina quinasa inactiva que es activada por la adición de un pequeño compuesto exógeno y (b) una segunda proteína de fusión recombinante que comprende un dominio que comprende un sitio de fosforilación del informador, con lo que dicho dominio está condensado a un segundo polipéptido de interacción.
2. El complejo de proteínas citoplasmáticas de acuerdo con la reivindicación 1, en el que dicho sitio de fosforilación del informador está fosforilado.
3. El complejo de proteínas citoplasmáticas de acuerdo con la reivindicación 1 ó 2, en donde dicha tirosina quinasa se selecciona del grupo que consiste en un mutante Tyk2, un mutante JAK o un mutante Src.
4. El complejo de proteínas citoplasmáticas de acuerdo con la reivindicación 1 ó 2, en donde dicha primera proteína de fusión recombinante comprende SEQ ID N° 1 o SEQ ID N° 2.
5. El complejo de proteínas citoplasmáticas de acuerdo con la reivindicación 1 ó 2, en donde dicho dominio de fosforilación del informador está comprendido en SEQ ID N° 3 o SEQ ID N° 4.
6. Un método para detectar la interacción compuesto-compuesto que comprende (a) condensar una quinasa mutante a un primer polipéptido de interacción, en donde dicha quinasa mutante se selecciona del grupo que consiste en una tirosina quinasa constitutiva mutante y una tirosina quinasa inactiva que es activada por la adición de un pequeño compuesto exógeno resultando en una primera proteína de fusión recombinante sin sitio de fosforilación del informador, (b) condensar un dominio que comprende un sitio de fosforilación del informador a un segundo polipéptido de interacción, resultando en una segunda proteína de fusión recombinante, (c) expresar ambas proteínas de fusión en una célula y (d) identificar y/o seleccionar aquellas células en las que dicho sitio de fosforilación del informador es fosforilado.
7. El método de la reivindicación 6, en el que dicha célula es una célula eucariota.
8. El método de la reivindicación 7, en el que dicha célula es una célula de mamíferos.
9. El método de acuerdo con cualquiera de las reivindicaciones 6-8, en el que dicha tirosina quinasa constitutiva se selecciona del grupo que consiste en un mutante Tyk2, un mutante JAK o un mutante Src.
10. El método de acuerdo con cualquiera de las reivindicaciones 6-8, en el que dicha quinasa comprende SEQ ID N° 1 o SEQ ID N° 2.
11. El método de acuerdo con cualquiera de las reivindicaciones 6-8, en el que dicho dominio de fosforilación del informador está comprendido en SEQ ID N° 3 o SEQ ID N° 4.
12. Una célula, que comprende el complejo de proteínas citoplasmáticas de acuerdo con cualquiera de las reivindicaciones 1-5.
13. La célula de acuerdo con la reivindicación 12, en donde dicha célula es una célula eucariota.
14. La célula de acuerdo con la reivindicación 13, en donde dicha célula eucariota es una célula de mamíferos.
15. Un método para detectar compuestos que interrumpen una interacción polipéptido-polipéptido, que comprende (a) cultivar una célula que comprende un complejo de proteínas de acuerdo con cualquiera de las reivindicaciones 12-14, en donde dicho complejo de proteínas citoplasmáticas comprende la interacción polipéptido-polipéptido que se quiere interrumpir, en ausencia y en presencia de al menos un compuesto, (b) comparar la fosforilación del sitio de fosforilación del informador en la segunda proteína de fusión recombinante de dicho complejo de proteínas citoplasmáticas de las células, que se cultiva en presencia y en ausencia de dicho compuesto y (c) identificar y/o seleccionar aquellas células en las que dicho sitio de fosforilación del informador no está fosforilado.

Figura 1

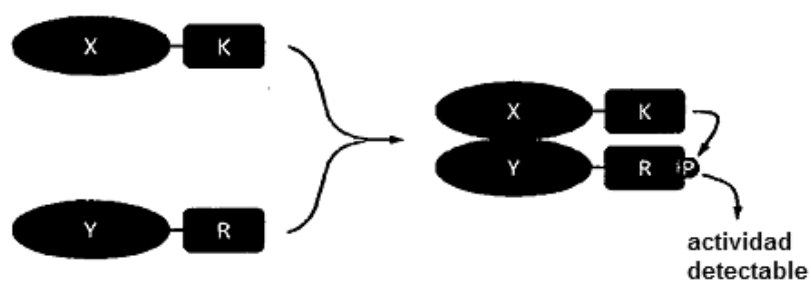
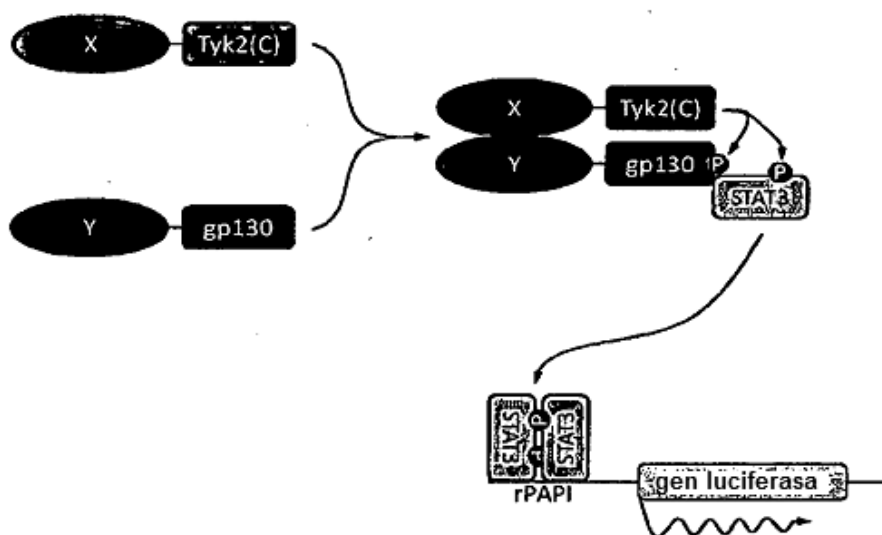


Figura 2

A



B

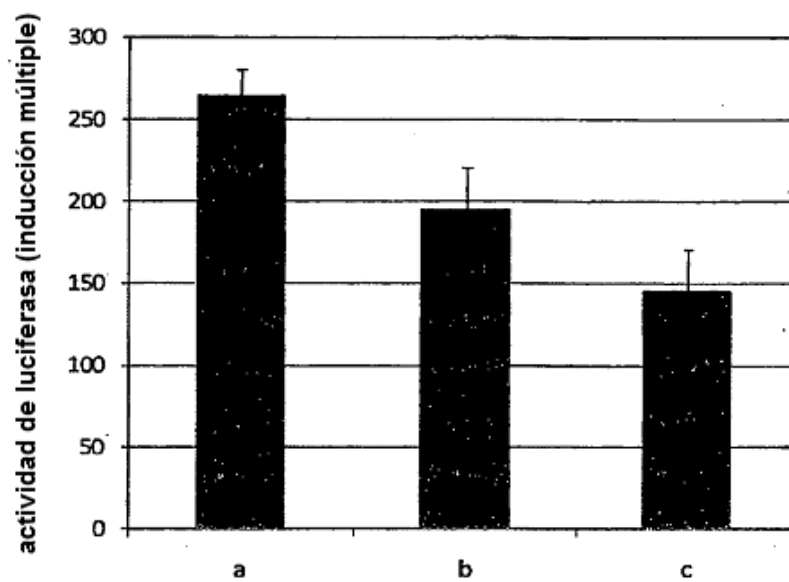


Figura 3

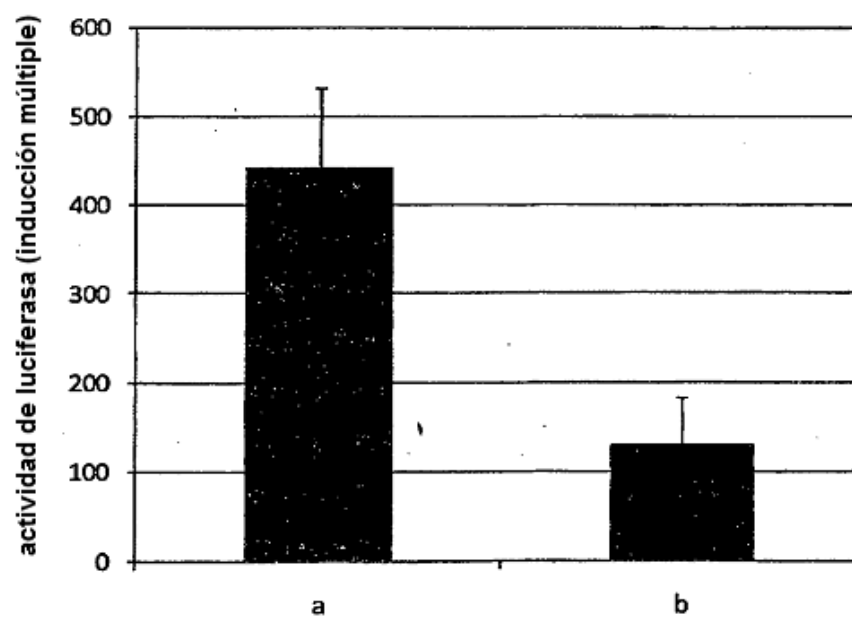


Figura 4

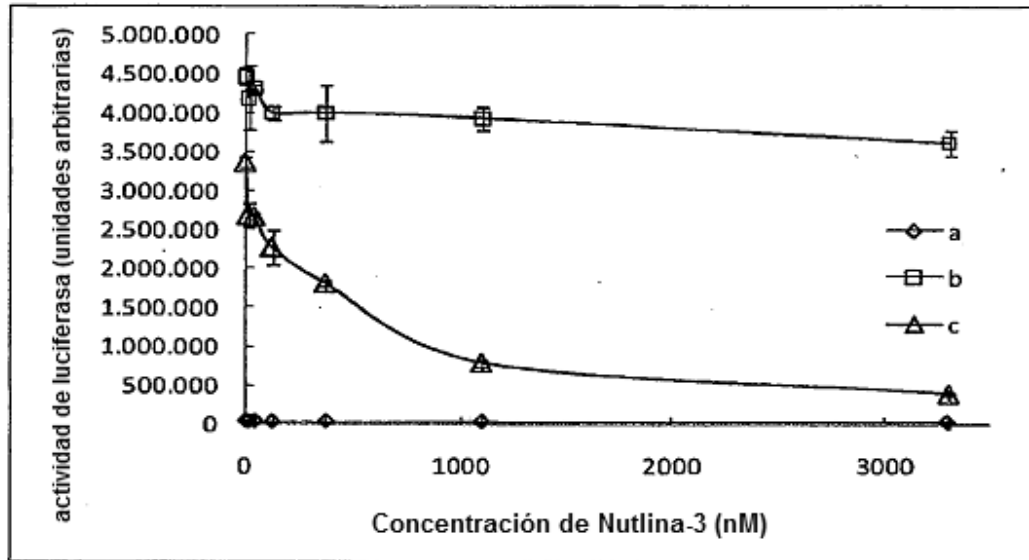


Figura 5

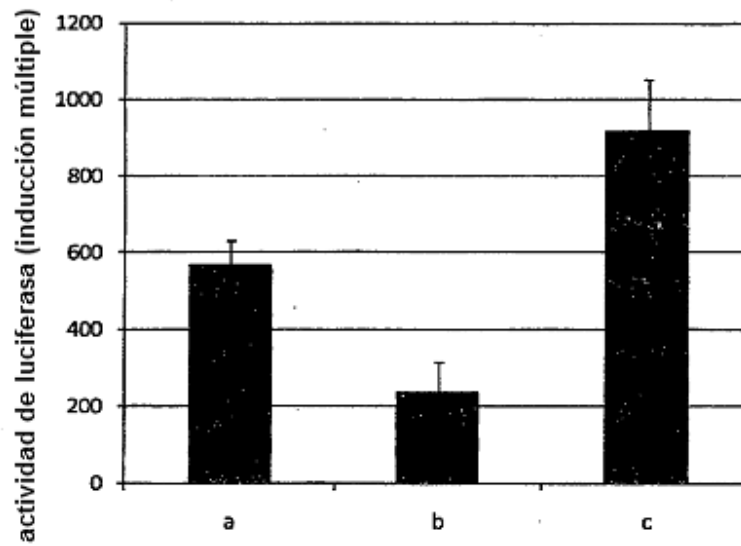


Figura 6