



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 360 205**

(51) Int. Cl.:

G01N 33/566 (2006.01)

A61K 31/01 (2006.01)

C07K 14/00 (2006.01)

C12Q 1/00 (2006.01)

C12N 15/62 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **02723332 .9**

(96) Fecha de presentación : **04.03.2002**

(97) Número de publicación de la solicitud: **1364212**

(97) Fecha de publicación de la solicitud: **26.11.2003**

(54) Título: **Sistema de ensayo de tres híbridos.**

(30) Prioridad: **02.03.2001 US 272932 P**
23.03.2001 US 278233 P
15.10.2001 US 329437 P

(45) Fecha de publicación de la mención BOPI:
01.06.2011

(45) Fecha de la publicación del folleto de la patente:
01.06.2011

(73) Titular/es: **AGENNIX AG.**
Im Neuenheimer Feld 515
69120 Heidelberg, DE
AGENNIX USA Inc.

(72) Inventor/es: **Becker, Frank;**
Come, Jon, H.;
Kley, Nikolai A. y
Reichel, Christoph

(74) Agente: **Zea Checa, Bernabé**

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

Referencia a solicitudes relacionadas

[0001] Esta solicitud reivindica prioridad respecto a la solicitud provisional de Estados Unidos 60/272.932 presentada el 2 de marzo de 2001; la solicitud provisional de Estados Unidos 60/278.233 presentada el 23 de marzo de 2001; y la solicitud provisional de Estados Unidos 60/329.437 presentada el 15 de octubre de 2001.

Antecedentes de la invención

[0002] Las interacciones de proteínas facilitan la mayoría de los procesos biológicos, incluyendo la transducción de señales y la homeostasia. Se ha avanzado en la elucidación de compañeros proteicos de interacción particulares que faciliten estos procesos biológicos mediante el desarrollo de métodos de “doble híbrido” o “trampa de interacción” para detectar y seleccionar compañeros proteicos de interacción (véase Fields y Songs (1989) Nature 340: 245-6; Gyuris et al. (1993) Cell 75: 791-803; Patente de Estados Unidos N° 5.468.614; y Yang et al. (1995) Nucleic Acid Research 23, 1152-1156). Estos métodos dependen de la reconstitución de un activador de la transcripción nuclear a través de la interacción de dos polipéptidos compañeros de unión —es decir, un primer polipéptido fusionado a un dominio de unión a ADN (BD) y un segundo polipéptido fusionado a un dominio de activación de la transcripción (AD). Cuando el primer y el segundo polipéptidos interaccionan, la interacción puede detectarse por activación de un gen indicador que contiene sitios de unión para el dominio de unión a ADN. Para que funcione este método, ambas proteínas tienen que ser solubles y deben ser capaces de localizarse en el núcleo. Por consiguiente, la interacción de polipéptidos que normalmente se localizan en otros compartimentos puede no detectarse debido a la ausencia de otros componentes polipeptídicos no nucleares que faciliten la interacción o a modificaciones postraduccionales no nucleares particulares que no suceden en el núcleo o porque las proteínas de interacción no se pliegan apropiadamente cuando se localizan en el compartimento nuclear. En particular, el ensayo de doble híbrido nuclear es incompatible con la detección de interacciones proteicas que suceden dentro de o en la superficie de membranas celulares. Además, este ensayo no sirve para explorar interacciones de moléculas pequeñas-proteínas debido a que depende únicamente de proteínas de fusión codificadas genéticamente.

[0003] Un área fundamental de investigación en farmacología y medicina es la determinación de interacciones de ligando-receptor. La base farmacológica de la acción de fármacos a nivel celular es bastante frecuentemente la consecuencia de interacciones no covalentes entre moléculas orgánicas pequeñas terapéuticamente relevantes y proteínas de unión de alta afinidad dentro de un tipo celular específico. Estos ligandos orgánicos pequeños pueden funcionar como agonistas o antagonistas de acontecimientos reguladores clave que organizan funciones celulares tanto normales como anormales. Durante años la estrategia de la industria farmacéutica para descubrir dichos ligandos ha sido la de explorar aleatoriamente miles de moléculas pequeñas en ensayos *in vitro* e *in vivo* específicos para determinar un compuesto candidato potente para sus esfuerzos de descubrimiento de fármacos. Usando estas herramientas, puede encontrarse un compuesto candidato que ejerza efectos muy bien definidos con respecto a una función en un tipo celular particular (por ejemplo, inhibición de la producción de citocinas o de la replicación del ADN en una línea celular de cáncer particular). Sin embargo, dichos resultados pueden proporcionar escasos indicios en lo que se refiere al mecanismo de acción a nivel molecular (interacción de ligando-proteína). Además, la exploración de una acción potente sobre una función celular puede dejar pasar reactividades cruzadas de un compuesto candidato que dan origen a efectos secundarios no deseados. Dichos efectos secundarios son con frecuencia la consecuencia de proteínas con estructuras muy similares que tienen funciones diferentes, o de una proteína que cumple diferentes funciones cuando se expresa en diferentes tipos celulares, o incluso cuando se localiza en diferentes compartimentos subcelulares. Por lo tanto, la identificación de las posiblemente diversas dianas proteicas para un agente farmacológico que muestra una actividad dada es desafiante pero altamente deseable. Existe una necesidad no cumplida de un método general y eficaz para identificar las dianas celulares para estos agentes farmacológicos de modo que se acelere la búsqueda de nuevos fármacos tanto a nivel de la investigación básica como aplicada.

[0004] De forma similar, existe la necesidad de una estrategia general para identificar una molécula pequeña capaz de unirse a cualquier diana celular seleccionada independientemente de su función biológica. Fowlkes et al. (WO 94/23025) y Broach et al. (WO/ 95/30012) describieron un ensayo de exploración para identificar moléculas capaces de unirse a receptores de la superficie celular de modo que se active una ruta de transducción de señales seleccionada. Estas referencias describen la modificación de rutas de señalización de levadura seleccionadas para mimetizar las etapas de la ruta de señalización de mamíferos. Esta última estrategia es específica para ciertas rutas de señalización y tiene una utilidad limitada para descubrir ampliamente moléculas pequeñas que interaccionen con cualquier diana celular. Por lo tanto, existe también una necesidad no cumplida de un método de exploración general para determinar la interacción entre moléculas pequeñas y proteínas diana para identificar nuevos fármacos que sean capaces de efectos terapéuticos específicos en una diversidad de patologías, así como para identificar agonistas y antagonistas que puedan interferir o competir con la unión de las moléculas pequeñas a estas dianas.

[0005] En este momento, existen pocas (si alguna) metodologías eficaces para identificar rápidamente una diana biológica, tal como una proteína para un ligando molecular pequeño particular. Las estrategias existentes incluyen el uso de cromatografía de afinidad, unión de ligando radiomarcado y marcaje por fotoafinidad en combinación con métodos de purificación de proteínas para detectar y aislar supuestas proteínas diana. Esto viene seguido de clonación del gen que codifica la proteína diana basándose en la secuencia peptídica de la diana aislada. Estas estrategias dependen de la abundancia de la supuesta proteína diana en la muestra y son laboriosas y meticulosas.

- [0006]** Crabtree et al. (documento WO/94/18317) describieron un método para activar un gen diana en células que comprenden (a) el suministro de células que contienen y son capaces de expresar (i) al menos una construcción de ADN que comprende al menos un dominio de receptor, capaz de unirse a un ligando seleccionado, fusionarse a una proteína adicional heteróloga capaz de iniciar un proceso biológico tras la exposición de la construcción de fusión al ligando, en el que el proceso biológico comprende la expresión del gen diana, en el que el ligando es capaz de unirse a dos o más proteínas de fusión, y en el que el proceso biológico sólo se inicia tras la unión del ligando a dos o más proteínas de fusión, siendo las dos proteínas de fusión iguales o diferentes, y (ii) el gen diana bajo el control de la expresión de un elemento de control que responde transcripcionalmente al inicio de dicho proceso biológico; y (b) exponer dichas células a dicho ligando en una cantidad eficaz para dar como resultado la expresión del gen indicador. Se describen además construcciones de ADN, ligandos y kits útiles para realizar dicho método. Los documentos relacionados US 5.830.462, US 5.869.337, US 6.165.787 muestran estas y otras realizaciones; en concreto, Holt et al. (documento WO 96/06097) describe la síntesis de ligandos híbridos para su uso con los presentes métodos.
- [0007]** El documento describe compuestos con restos de ligando de N-oxalil-pipecolilo y N-oxalil-prolilo para multimerización de inmunofilinas.
- [0008]** El fin previsto para estos métodos y composiciones se limita a la investigación de procesos celulares, la regulación de la síntesis de proteínas de importancia terapéutica o agrícola y la regulación de procesos celulares en terapia génica. Nada en los mismos sugiere el uso de estos métodos y composiciones para estudiar la interacción de proteínas con moléculas pequeñas, particularmente en su aplicación a la investigación farmacéutica y al desarrollo de fármacos.
- [0009]** Schreiber et al. (documento WO 96/13613) describen métodos para identificar un agente capaz de lograr un acontecimiento biológico mediado por la asociación de dos o más mediadores de proteínas endógenos, que implica la unión covalente de compuestos que se unen a estos mediadores para formar un agente de dimerización capaz de unirse a ambos mediadores. La unión de los compuestos puede realizarse por medio de un enlazador químico.
- [0010]** Haralambidis et al. (Patente de Estados Unidos N° 5.846.728) describen conjugados de oligonucleótido-poliámida en los que el oligonucleótido y la poliámida se unen por medio de un enlazador L para formar un heterodímero. La porción oligonucleotídica del conjugado comprende una secuencia antisentido que es complementaria a e hibrida con el polinucleótido diana.
- [0011]** Licitra y Liu (Proc. Natl. Acad. Sci. 1996; 93: 12817-21) describen un “ensayo de exploración de triple híbrido” que implica ligandos híbridos para detectar interacciones de ligando pequeño-receptor de proteína.
- [0012]** Licitra y Liu (documento WO 97/41255) también describieron un “ensayo de exploración de triple híbrido” en el que se pone en práctica el sistema de ensayo de doble híbrido de levadura básico. La diferencia significativa es: en lugar de depender de la interacción entre un denominado “cebo” y una denominada proteína “presa”, la transcripción del gen indicador depende de la proximidad de las dos proteínas, pudiendo cada una de ellas unirse específicamente a uno de los dos restos de un ligando híbrido pequeño. El ligando híbrido pequeño constituye el “tercer” componente del sistema de ensayo híbrido. En ese sistema, un resto conocido del ligando híbrido se unirá a la proteína “cebo”, mientras que la interacción entre el otro resto y la proteína “presa” puede aprovecharse para explorar para una proteína que pueda unirse a un resto conocido, o un pequeño resto (compuesto farmacéutico o fármaco) que pueda unirse a una diana proteica conocida.
- [0013]** Sin embargo, el sistema de triple híbrido de Liu presenta varias limitaciones: 1) el uso de un ensayo indicador de activación de transcripcional es incompatible con proteínas no nucleares, por ejemplo, proteínas unidas a membrana y proteínas citosólicas; 2) el ligando híbrido debe localizarse en el núcleo y permanece estable; y 3) la interacción entre la proteína “cebo” y su resto de unión en el ligando híbrido debe tener gran afinidad, preferiblemente a nivel nanomolar. Por ejemplo, se usó la interacción FK506-FKBP que proporciona una afinidad micromolar. Se desea una mayor afinidad entre proteína cebo y su compañero de unión para mejorar el rendimiento del sistema.
- [0014]** Lin et al. (J. Am. Chem. Soc. 2000, 122: 4247-8) realizaron mejoras sobre el sistema de triple híbrido existente por sustitución de la pareja FK506-FKBP con un ligando híbrido que consistía en dihidrofolato reductasa (DHFR) unida a metotrexato (Mtx) (DHFR-Mtx), que proporciona una afinidad picomolar, mejorando por lo tanto significativamente el rendimiento del sistema.
- [0015]** Las Patentes de Estados Unidos N° 5.585.245 y 5.503.977 describen los métodos de “ubiquitina dividida”, que pueden detectar interacciones de proteína-proteína mediante el uso de una proteasa específica de ubiquitina para escindir un polipéptido indicador de una proteína de fusión. Se construyen dos proteínas de fusión, consistiendo una en la mitad N-terminal de la ubiquitina y una proteína presa (Nub-presa o presa-Nub) y consistiendo la otra en la mitad C-terminal de la ubiquitina, una proteína cebo y el indicador (cebo-Cub-indicador). La asociación de presa y cebo reconstituye una estructura de ubiquitina reconocida por la proteasa específica de ubiquitina, por lo que el indicador se escinde de la proteína de fusión. La escisión del indicador de la proteína de fusión puede detectarse por varias técnicas, por ejemplo, escisión o desestabilización del indicador o permitiendo su translocación.

Descripción resumida de la invención

[0016] Un aspecto de la presente invención proporciona un ligando híbrido representado por la fórmula general: R1-Y-R2, en la que:

R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, 2,4-diaminopteridina o ciclosporina, o un derivado de los mismos con modificaciones estructurales minoritarias;

Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es número entero de 2 a 25; y

R2 representa un segundo ligando especificado por el usuario diferente de R1 seleccionado de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o un nucleótido.

[0017] En una realización, el primer ligando se une a un polipéptido. En una realización preferida, la afinidad de unión corresponde a una constante de disociación de ligando/polipéptido K_D de menos de 1 μM . En otra realización preferida, el primer ligando es capaz de formar un enlace covalente con el polipéptido.

[0018] En otra realización X es O. En otra realización, Y es $(CH_2-O-CH_2)_n$, donde n = 2 a 5. En otra realización, R1 es dexametasona. En otra realización, R1 es metotrexato, un derivado de metotrexato, FK506, un derivado de FK506 o un derivado de 2,4-diaminopteridina. En una realización preferida, R1 es dexametasona, Y es $(CH_2OCH_2)_3$ y R2 es metotrexato o un derivado de 2,4-diaminopteridina. En una realización más preferida, R1 es metotrexato e Y es $(CH_2-O-CH_2)_n$, donde n = 2 a 5.

[0019] En otra realización, R2 es un ligando seleccionado de: un compuesto con un efecto biológico conocido, un compuesto con un mecanismo de acción desconocido, un compuesto que se une a más de un polipéptido, un compuesto candidato a fármaco, o un compuesto que se une a una proteína desconocida.

[0020] En otra realización, R2 se une a o inhibe una quinasa.

[0021] El número entero n puede ser de 2 a 20, o de 2 a 15, o de 2 a 10, o de 2 a 5.

[0022] Un aspecto relacionado de la invención proporciona un ligando híbrido representado por la fórmula general: R1-Y-R2, en la que:

R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, derivado de 2,4-diaminopteridina o ciclosporina, o un derivado con modificaciones estructurales minoritarias;

Y representa un enlazador; y

R2 representa un segundo ligando especificado por el usuario diferente de R1 seleccionado de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o un nucleótido;

en la que R2 se une a o inhibe una quinasa.

[0023] En una realización, la quinasa es una quinasa dependiente de ciclina. En otra realización, R2 es un compuesto seleccionado de la Tabla 2 que contiene aproximadamente 600 compuestos que se sabe que son capaces de unirse a o inhibir una quinasa, o un derivado del mismo con modificaciones estructurales minoritarias. En otra realización, Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es número entero de 2 a 25.

[0024] Otro aspecto de la invención proporciona un polipéptido de fusión que comprende los segmentos P1, Cub-Z y RM, en un orden en el que Cub-Z está más próximo al extremo N-terminal del polipéptido de fusión que RM, en el que 1) P1 es un polipéptido de unión a ligando que se une a un ligando no peptídico de un ligando híbrido, que tiene la fórmula general R1-Y-R1, en la que R1 y R2 son ligandos e Y es un enlazador, 2) Cub es un subdominio carboxi-

terminal de ubiquitina, 3) Z es un resto aminoacídico, 4) RM es un resto indicador.

[0025] Otro aspecto de la invención proporciona un polipéptido de fusión que comprende los segmentos P1 y Nux, en el que 1) Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida y 2) P1 es un polipéptido de unión a ligando que se une a un ligando no peptídico de un ligando híbrido, que tiene la fórmula general R1-Y-R1, en la que R1 y R2 son ligandos e Y es un enlazador.

[0026] En una realización preferida, los ligandos no peptídicos de las proteínas de fusión son: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, 2,4-diaminopteridina, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, ciclosporina o un derivado de los mismos con modificaciones estructurales minoritarias; o

[0027] un carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido, o un nucleótido.

[0028] En otra realización, Z es un aminoácido distinto de metionina. En otra realización, RM es: un polipéptido capaz de emitir luz con la excitación, un polipéptido con una actividad enzimática, un marcador detectable o un factor de transcripción. En otra realización, RM es: proteína verde fluorescente, URA3 o PLV.

[0029] Otro aspecto de la invención proporciona un ácido nucleico que codifica el polipéptido de fusión de cualquiera de la presente invención.

[0030] En otra realización, X es O. En otra realización, Y es $(CH_2OCH_2)_3$. En otra realización, R1 es dexametasona, Y es $(CH_2OCH_2)_3$ y R2 es metotrexato o 2,4-diaminopteridina.

[0031] Otro aspecto de la invención proporciona una composición que comprende: 1) un ligando híbrido de fórmula general R1-Y-R2, en la que R1 y R2 son ligandos, R1 es diferente de R2 y al menos uno de R1 y R2 no es un péptido, Y es un enlazador; y 2) al menos uno de los dos polipéptidos de fusión, que comprenden: a) un primer polipéptido de fusión que comprende los segmentos P2, Cub-Z y RM en un orden en el que Cub-Z está más próximo al extremo N-terminal del primer polipéptido de fusión que RM, en el que P2 es un polipéptido de unión a ligando que puede unirse al ligando R1 o R2 del ligando híbrido, Cub es un subdominio carboxi-terminal de ubiquitina y RM es un resto indicador, y Z es un resto aminoacídico; b) un segundo polipéptido de fusión que comprende los segmentos Nux y P1, en el que Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida, y P1 es un polipéptido de unión a ligando que puede unirse al ligando R1 o R2 del ligando híbrido.

[0032] Un aspecto relacionado la invención proporciona una composición que comprende: 1) un ligando híbrido representado por la fórmula general: R1-Y-R2, en la que: a) R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, derivado de 2,4-diaminopteridina, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel o ciclosporina, o un derivado de los mismos con modificaciones estructurales minoritarias; b) Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es un número entero de 2 a 25; c) R2 representa un segundo ligando especificado por el usuario diferente de R1 seleccionado de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o un nucleótido; 2) al menos un polipéptido de fusión seleccionado de: a) un primer polipéptido de fusión que comprende: un dominio de unión a ligando P1 y un dominio seleccionado del grupo que consiste en: un dominio de unión a ADN y un dominio de activación de la transcripción, en el que el dominio de unión a ligando puede unirse al primer ligando R1; y b) un segundo polipéptido de fusión que comprende: un dominio unión a ligando candidato P2 que puede unirse al ligando especificado por el usuario R2 y un dominio seleccionado del grupo que consiste en: un dominio de unión a ADN y un dominio de activación de la transcripción, en el que uno del primer y segundo polipéptidos de fusión contiene un dominio de unión a ADN y el otro polipéptido de fusión contiene un dominio de activación de la transcripción.

[0033] Otro aspecto relacionado de la invención proporciona una composición que comprende: 1) un ligando híbrido representado por la fórmula general: R1-Y-R2, en la que: a) R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, derivado de 2,4-diaminopteridina, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel o ciclosporina, o un derivado de los mismos con modificaciones estructurales minoritarias; b) Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es un número entero de 2 a 25; c) R2 representa un segundo ligando especificado por el usuario diferente de R1 seleccionado de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de

- alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o un nucleótido; y 2) un polipéptido de fusión que incluye: a) al menos un dominio de unión a ligando; y b) un dominio funcional heterólogo respecto al dominio de unión a ligando que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional.
- [0034]** En una realización, la composición es un complejo. En otra realización, la composición está proporcionada en un entorno seleccionado de: una célula, un recipiente, un kit, una solución o un medio de cultivo.
- [0035]** Otro aspecto de la invención proporciona un método de identificación de una secuencia polipeptídica que se une a un ligando especificado por el usuario que comprende: 1) proporcionar un ligando híbrido que tiene la fórmula general $R1-Y-R2$, en la que $R1$ es un primer ligando, $R2$ es un ligando especificado por el usuario e Y es un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O , S , SO o SO_2 y n es un número entero de 2 a 25; 2) introducir el ligando híbrido en una población de células, conteniendo cada célula un sistema de exploración de ligandos híbridos que incluye: a) un gen indicador unido operativamente a una secuencia reguladora de la transcripción, incluyendo dicha secuencia reguladora una secuencia de ADN que se une a un dominio de unión a ADN; b) un primer gen quimérico que codifica un primer polipéptido de fusión que comprende: un dominio de unión a ligando $P1$ y un dominio seleccionado de un dominio de unión a ADN o un dominio de activación de la transcripción, en el que el dominio de unión a ligando se une al primer ligando $R1$; y c) un segundo gen quimérico que codifica un segundo polipéptido de fusión que comprende: un dominio de unión a ligando candidato $P2$ para el ligando especificado por el usuario $R2$ y un dominio seleccionado de un dominio de unión a ADN o un dominio de activación de la transcripción; en el que uno de los dos polipéptidos de fusión contiene un dominio de unión a ADN y el otro polipéptido de fusión contiene un dominio de activación de la transcripción; 3) permitir que el ligando híbrido se una al dominio de unión a ligando del primer polipéptido de fusión a través del primer ligando $R1$ y contacte con el dominio de unión a ligando candidato del segundo polipéptido de fusión a través del ligando especificado por el usuario $R2$ de modo que, si $R2$ se une al dominio de unión a ligando candidato, se produce un aumento en el nivel de transcripción del gen indicador; 4) identificar una célula de unión a ligando positiva en la que se ha producido un aumento en el nivel de transcripción del gen indicador; y 5) identificar la secuencia de ácido nucleico del segundo gen quimérico que codifica el dominio de unión a ligando candidato que se une al ligando especificado por el usuario $R2$, identificando de este modo una secuencia polipeptídica que se une a un ligando especificado por el usuario.
- [0036]** En una realización, la secuencia de ácido nucleico que codifica el polipéptido de dominio de unión a ligando candidato del segundo polipéptido de fusión es de una biblioteca seleccionada de: una biblioteca de oligonucleótidos sintéticos, una genoteca de ADNc, una genoteca de fragmentos de ADN genómico bacteriano o una genoteca de fragmentos de ADN genómico eucariota.
- [0037]** En otra realización, la biblioteca tiene aproximadamente 2-10 miembros, o aproximadamente 10-500 miembros, o aproximadamente 500-10.000 miembros o al menos 10.000 miembros.
- [0038]** En otra realización, la secuencia de ácido nucleico que codifica la secuencia polipeptídica del dominio de unión a ligando candidato representa una sola diana farmacológica seleccionada por el usuario.
- [0039]** En otra realización, el primer ligando $R1$ del ligando híbrido se une al dominio de unión a ligando $P1$ con una gran afinidad. En una realización preferida, la afinidad de unión corresponde a una constante de disociación de ligando/proteína de unión a ligando K_D inferior a $1 \mu M$.
- [0040]** En otra realización, el primer ligando es capaz de formar un enlace covalente con el dominio de unión a ligando $P1$.
- [0041]** En otra realización, X es O . En otra realización, Y es $(CH_2-O-CH_2)_n$, donde $n = 2$ a 5 . En otra realización, $R1$ es metotrexato e Y es $(CH_2-O-CH_2)_n$, donde $n = 2$ a 5 . En otra realización, el gen indicador se selecciona de: $HIS3$, $LEU2$, $TRP2$, $TRP1$, $ADE2$, $LYS2$, $URA3$, $CYH1$, $CAN1$, $lacZ$, gfp o CAT . En otra realización, $R2$ se une a o inhibe una quinasa.
- [0042]** Otro aspecto de la invención proporciona un método de identificación de una secuencia polipeptídica que se une a un ligando especificado por el usuario que comprende: 1) proporcionar un ligando híbrido que tiene la fórmula general $R1-Y-R2$, en la que $R1$ es un primer ligando, $R2$ es un ligando especificado por el usuario diferente de $R1$ que se une a o inhibe una quinasa, al menos uno de $R1$ y $R2$ no es un péptido e Y es un enlazador; 2) introducir el ligando híbrido en una población de células, conteniendo cada célula un sistema de exploración de ligandos híbridos que incluye: a) un gen indicador unido operativamente a una secuencia reguladora de la transcripción, incluyendo dicha secuencia reguladora una secuencia de ADN que se une a un dominio de unión a ADN; b) un primer gen quimérico que codifica un primer polipéptido de fusión que comprende: un dominio de unión a ligando y un dominio seleccionado del dominio de unión a ADN o un dominio de activación de la transcripción, en el que el dominio de unión a ligando se une al primer ligando $R1$; y c) un segundo gen quimérico que codifica un segundo polipéptido de fusión que comprende: un dominio de unión a ligando candidato para el ligando especificado por el usuario $R2$ y un dominio seleccionado del dominio de unión a ADN o del dominio de activación de la transcripción; en el que uno de los dos polipéptidos de fusión contiene un

- dominio de unión a ADN y el otro polipéptido de fusión contiene un dominio de activación de la transcripción; 3) permitir que el ligando híbrido se una al dominio de unión a ligando del primer polipéptido de fusión a través del primer ligando R1 y contacte con el dominio de unión a ligando candidato del segundo polipéptido de fusión a través del ligando especificado por el usuario R2 de modo que, si R2 se une al dominio de unión a ligando candidato, se produce un aumento en el nivel de transcripción del gen indicador; 4) identificar una célula de unión a ligando positiva en la que se ha producido un aumento en el nivel de transcripción del gen indicador; y 5) identificar la secuencia de ácido nucleico del segundo gen quimérico que codifica el dominio de unión a ligando candidato que se une al ligando especificado por el usuario R2, identificando de este modo una secuencia polipeptídica que se une a un ligando especificado por el usuario.
- [0043]** En una realización, la quinasa es una quinasa dependiente de ciclina. En una realización, R2 es un compuesto seleccionado de la Tabla 2. En una realización, Y es $(CH_2-X-CH_2)_n$, $n = 2$ a 25. En una realización, R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, derivado de 2,4-diaminopteridina o ciclosporina, o un derivado de los mismos con modificaciones estructurales minoritarias.
- [0044]** En otra realización, el método comprende además determinar la afinidad de unión del ligando híbrido a los dominios de unión a ligando P1 y/o P2. En una realización preferida, la determinación de la afinidad de unión se realiza por resonancia de plasmón superficial.
- [0045]** En otra realización, el método comprende además determinar los efectos del ligando híbrido que son independientes de la formación de un complejo trimérico que comprende el ligando híbrido, P1 y P2.
- [0046]** En otra realización, el método comprende además la etapa de: realizar al menos un método separado adicional para confirmar que la transcripción del gen indicador depende de la presencia del ligando híbrido y los dominios de unión a ligando P1 y P2. En una realización preferida, dicho método separado adicional se selecciona de: un método de ensayo de crecimiento en halo o ensayo de crecimiento por detección de fluorescencia. En una realización más preferida, dicho método separado adicional se realiza individualmente en más de aproximadamente 10, 100, 1000 ó 10000 tipos celulares de unión a ligando positiva diferentes identificados en la etapa 4).
- [0047]** Un aspecto relacionado de la invención proporciona un método de identificación de una secuencia polipeptídica que se une a un ligando especificado por el usuario que comprende: proporcionar un ligando híbrido que tiene la fórmula general R1-Y-R2, en la que R1 es un primer ligando, R2 es un ligando especificado por el usuario e Y es un enlazador; poner en contacto el ligando híbrido con una célula cultivada que comprende: un primer gen quimérico que codifica un primer polipéptido de fusión que comprende: los segmentos P1, Cub-Z y RM en un orden en el que Cub-Z está más próximo al extremo N-terminal del primer polipéptido de fusión que RM, en el que P1 es un polipéptido de unión a ligando que se une al primer ligando R1, Cub es un subdominio carboxi-terminal de ubiquitina, Z es un resto aminoácido distinto de metionina y RM es un resto indicador, un segundo gen quimérico que codifica un segundo polipéptido de fusión que comprende: los segmentos Nux y P2, en el que Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida, y P2 es un polipéptido de unión a ligando candidato para el ligando especificado por el usuario R2; y un sistema proteolítico dependiente de ubiquitina que comprende una proteasa específica de ubiquitina (UBP) de la regla del N-terminal; permitir que el ligando híbrido se una al polipéptido de unión a ligando P1 del primer polipéptido de fusión a través del primer ligando R1 y contacte con el polipéptido de unión a ligando candidato P2 del segundo polipéptido de fusión a través del ligando especificado por el usuario R2 de modo que, cuando R2 se une al polipéptido de unión a ligando candidato P2, los dominios Nux y Cub se asocian para formar un resto de ubiquitina reconstituida y la proteasa específica de ubiquitina escinde el enlace peptídico de Cub-Z para liberar un fragmento que contiene RM, siendo dicho fragmento susceptible a la degradación proteolítica dependiente de ubiquitina de la regla del N-terminal; mantener la célula cultivada en condiciones en las que la escisión del enlace de Cub-Z es necesaria para el crecimiento de la célula; e identificar la secuencia del gen quimérico que codifica el polipéptido de unión a ligando candidato P2, identificando de este modo una secuencia polipeptídica que se une a un ligando especificado por el usuario.
- [0048]** Otro aspecto relacionado de la invención proporciona un método de identificación de una secuencia polipeptídica que se une a un ligando especificado por el usuario que comprende: proporcionar un ligando híbrido que tiene la fórmula general R1-Y-R2, en la que R1 es un primer ligando, R2 es un ligando especificado por el usuario e Y es un enlazador; poner en contacto el ligando híbrido con una célula cultivada que comprende: un primer gen quimérico que codifica un primer polipéptido de fusión que comprende: los segmentos Nux y P1, en el que Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida, y P1 es un polipéptido de unión a ligando para el primer ligando R1, un segundo gen quimérico que codifica un segundo polipéptido de fusión que comprende: los segmentos P2, Cub-Z y RM en un orden en el que Cub-Z está más próximo al extremo N-terminal del segundo polipéptido de fusión que RM, en el que P2 es un polipéptido de unión a ligando candidato que se une al ligando especificado por el usuario R2, Cub es un subdominio carboxi-terminal de ubiquitina, Z es un resto aminoácido distinto de metionina y RM es un resto indicador; y un sistema proteolítico dependiente de ubiquitina que comprende una proteasa específica de ubiquitina de la regla del N-terminal; permitir que el ligando híbrido se una al polipéptido de unión a ligando P1 del primer polipéptido de fusión a través del primer ligando R1 y contacte con el polipéptido de unión a ligando candidato P2 del segundo polipéptido de fusión a través del ligando especificado por el usuario R2 de modo que, cuando R2 se une al polipéptido de unión a ligando candidato P2, los subdominios Nux

y Cub se asocian para formar un resto de ubiquitina reconstituida y la proteasa específica de ubiquitina escinde el enlace peptídico de Cub-Z para liberar un fragmento que contiene RM, siendo dicho fragmento susceptible a la degradación proteolítica dependiente de ubiquitina de la regla del N-terminal; mantener la célula cultivada en condiciones en las que la escisión del enlace de Cub-Z sea necesaria para el crecimiento de la célula; e identificar la secuencia del segundo gen quimérico que codifica el polipéptido de unión a ligando candidato P2, identificando de este modo una secuencia polipeptídica que se une a un ligando especificado por el usuario.

[0049] En una realización, P2 está codificado por un ácido nucleico de una biblioteca seleccionada del grupo que consiste en: una biblioteca de oligonucleótidos sintéticos, una genoteca de ADNc, una genoteca de fragmentos de ADN genómico bacteriano y una genoteca de fragmentos de ADN genómico eucariota. En otra realización, la secuencia de ácido nucleico que codifica la secuencia proteica de unión a ligando candidato representa una sola diana farmacológica seleccionada por el usuario. En otra realización, el primer ligando del ligando híbrido se une al polipéptido de unión a ligando con una gran afinidad. En otra realización, el primer ligando es metotrexato y el primer polipéptido de unión a ligando es DHFR. En otra realización, la afinidad de unión corresponde a una constante de disociación de ligando/proteína de unión a ligando de menos de $1 \mu\text{M}$. En otra realización, el primer ligando es capaz de formar un enlace covalente con el polipéptido de unión a ligando. En otra realización, Y es $(\text{CH}_2\text{OCH}_2)_3$. Preferiblemente, R1 es dexametasona, Y es $(\text{CH}_2\text{OCH}_2)_3$ y R2 es metotrexato o 2,4-diaminopteridina. En otra realización, el resto indicador (RM) es un marcador de selección negativa expresado en una célula que expresa el primer y segundo polipéptidos de fusión, y en la que una disminución en el nivel del resto indicador causa un aumento en el crecimiento de dicha célula. En otra realización, el resto indicador (RM) es un marcador de selección positiva expresado en una célula que expresa el primer y segundo polipéptidos de fusión, y en la que un aumento en la actividad del resto indicador causa un aumento en el crecimiento de dicha célula.

[0050] Otro aspecto relacionado de la invención proporciona un método de identificación de una secuencia polipeptídica que se une a un ligando especificado por el usuario que comprende: proporcionar un ligando híbrido que tiene la fórmula general R1-Y-R2, en la que R1 es un primer ligando, R2 es un ligando especificado por el usuario e Y es un enlazador; poner en contacto el ligando híbrido con una célula cultivada que comprende: un primer gen quimérico que codifica un primer polipéptido de fusión que comprende: los segmentos P1, Cub-Z y RM en un orden en el que Cub-Z está más próximo al extremo N-terminal del primer polipéptido de fusión que RM, en el que P1 es un polipéptido de unión a ligando que se une al primer ligando R1, Cub es un subdominio carboxi-terminal de ubiquitina, Z es metionina y RM es un resto indicador, un segundo gen quimérico que codifica un segundo polipéptido de fusión que comprende: los segmentos Nux y P2, en el que Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida, y P2 es un polipéptido de unión a ligando candidato para el ligando especificado por el usuario R2; y un sistema proteolítico dependiente de ubiquitina que comprende una proteasa específica de ubiquitina (UBP) de la regla del N-terminal; permitir que el ligando híbrido se una al polipéptido de unión a ligando P1 del primer polipéptido de fusión a través del primer ligando R1 y contacte con el polipéptido de unión a ligando candidato P2 del segundo polipéptido de fusión a través del ligando especificado por el usuario R2 de modo que, cuando R2 se une al polipéptido de unión a ligando candidato P2, los dominios Nux y Cub se asocian para formar un resto de ubiquitina reconstituida y la proteasa específica de ubiquitina escinde el enlace peptídico de Cub-Z para liberar un fragmento que contiene RM, siendo dicho fragmento que no es susceptible a la degradación proteolítica dependiente de ubiquitina de la regla del N-terminal funcional tras la escisión; mantener la célula cultivada en condiciones en las que la escisión del enlace de Cub-Z es necesaria para el crecimiento de la célula; e identificar la secuencia del gen quimérico que codifica el polipéptido de unión a ligando candidato P2, identificando de este modo una secuencia polipeptídica que se une a un ligando especificado por el usuario.

[0051] Otro aspecto de la invención proporciona un método de determinación de si un polipéptido P2 y un ligando R2 se unen entre sí que comprende: 1) proporcionar por traducción un primer polipéptido de unión a ligando que comprende los segmentos P1, Cub-Z y RM en un orden en el que Cub-Z está más próximo al extremo N-terminal del primer polipéptido de unión a ligando que RM, y un segundo polipéptido de unión a ligando que comprende los segmentos Nux y P2, en el que P1 y P2 son polipéptidos, Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida, Cub es el subdominio carboxi-terminal de una ubiquitina de tipo silvestre, Z es un resto aminoácido y RM es un resto indicador; 2) proporcionar un ligando híbrido representado por la fórmula general: R1-Y-R2, en la que R1 es un primer ligando que se une al primer polipéptido de unión a ligando en P1, R2 es un segundo ligando diferente de R1, al menos uno de R1 y R2 no es un péptido e Y es un enlazador, 3) permitir que el ligando híbrido contacte con el primer y segundo polipéptidos de unión a ligando; 4) detectar el grado de escisión por una proteasa específica de ubiquitina (UBP) del primer polipéptido de unión a ligando entre Cub y Z, en el que un aumento de la escisión es indicativo de la unión de polipéptido P2-ligando R2.

[0052] Otro aspecto de la invención proporciona un método de determinación de si un polipéptido P1 y un ligando R1 se unen entre sí, que comprende: 1) proporcionar por traducción un primer polipéptido de unión a ligando que comprende los segmentos P1, Cub-Z y RM, en un orden en el que Cub-Z esté más próximo al extremo N-terminal del primer polipéptido de unión a ligando que RM, y un segundo polipéptido de unión a ligando que comprende los segmentos Nux y P2, en el que P1 y P2 son polipéptidos, Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida, Cub es el subdominio carboxi-terminal de una ubiquitina de tipo silvestre, Z es un resto aminoácido y RM es un resto indicador; 2) proporcionar un ligando híbrido representado por la fórmula general: R1-Y-R2, en la que R1 es un primer ligando, R2 es un segundo ligando diferente de R1 que se une al segundo polipéptido de unión a ligando en P2, al menos uno de R1 y R2 no es un péptido e Y es un

enlazador; 3) permitir que el ligando híbrido contacte con el primer y segundo polipéptidos de unión a ligando; 4) detectar el grado de escisión por una proteasa específica de ubiquitina (UBP) del primer polipéptido de unión a ligando entre Cub y Z, en el que un aumento de la escisión es indicativo de unión de proteína P1-ligando R1.

[0053] En una realización, la etapa 1) implica el uso de una célula que proporciona un sistema de degradación de la regla del N-terminal. En una realización, el grado de escisión entre Cub y Z se determina por detección del grado de actividad del RM. En una realización, el grado de escisión entre Cub y Z se determina por detección del grado de actividad enzimática del RM. En una realización, el grado de escisión entre Cub y Z se determina detectando la cantidad de la forma escindida de RM.

[0054] Otro aspecto de la invención proporciona un método para inducir o permitir la detección de un acontecimiento biológicamente detectable, que comprende: 1) proporcionar al menos una célula que comprende al menos una secuencia de ácido nucleico que codifica un polipéptido de fusión que incluye: a) al menos un dominio de unión a ligando; y b) un dominio funcional que por sí mismo no es capaz de inducir o permitir la detección del acontecimiento detectable; 2) proporcionar un ligando híbrido de fórmula general R1-Y-R2, en la que R1 es diferente de R2, al menos uno de R1 y R2 no es un péptido, R1 o R2 representa un ligando que se une a dicho dominio de unión a ligando; Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO₂ y n es un número entero de 2 a 25; y en la que la unión de dicho ligando híbrido a dicho dominio de unión a ligando pone al primer dominio funcional en las proximidades de un segundo dominio funcional, induciendo o permitiendo de ese modo la detección del acontecimiento detectable; y 3) exponer dicha al menos una célula a una cantidad eficaz de dicho ligando híbrido para poner al primer dominio funcional en las proximidades de un segundo dominio funcional, induciendo o permitiendo de este modo la detección del acontecimiento detectable.

[0055] Otro aspecto de la invención proporciona un método de identificación de un ligando de un polipéptido especificado por el usuario, que comprende: 1) proporcionar al menos un ligando híbrido candidato que tiene la fórmula general R1-Y-R2, en la que R1 es un primer ligando, R2 es un ligando candidato e Y es un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO₂ y n es un número entero de 2 a 25; 2) introducir el ligando híbrido candidato en al menos una célula que contiene un sistema de exploración de ligandos híbridos, que incluye: a) un gen indicador unido operativamente a una secuencia reguladora de la transcripción, incluyendo dicha secuencia reguladora una secuencia de ADN que se une a un dominio de unión a ADN; b) un primer gen químico que codifica un primer polipéptido de fusión, que comprende: un dominio de unión a ligando y un dominio seleccionado del dominio de unión a ADN o un dominio de activación de la transcripción, en el que el dominio de unión a ligando se une al primer ligando R1; y c) un segundo gen químico que codifica un segundo polipéptido de fusión, que comprende: un dominio de unión a ligando especificado por el usuario para el ligando candidato R2 y un dominio seleccionado del dominio de unión a ADN o el dominio de activación de la transcripción; en el que uno de los dos polipéptidos de fusión contiene un dominio de unión a ADN y el otro polipéptido de fusión contiene un dominio de activación de la transcripción; 3) permitir que el ligando híbrido candidato se una al dominio de unión a ligando del primer polipéptido de fusión a través del primer ligando R1 y que contacte con el dominio de unión a ligando especificado por el usuario del segundo polipéptido de fusión a través del ligando candidato R2 de modo que, si el dominio de unión a ligando especificado por el usuario se une al ligando candidato R2, se produce un aumento en el nivel de transcripción del gen indicador; 4) identificar el ligando híbrido candidato que causa un aumento en el nivel de transcripción del gen indicador en la célula, identificando de este modo el ligando candidato en el ligando híbrido candidato como ligando para el polipéptido especificado por el usuario.

[0056] Un aspecto relacionado de la invención proporciona un método de identificación de un ligando que se une a un polipéptido especificado por el usuario, que comprende: proporcionar una población de ligando híbrido candidato que tiene la fórmula general R1-Y-R2, en la que R1 es un primer ligando, R2 es un ligando candidato e Y es un enlazador; poner en contacto cada ligando híbrido candidato individual con un sistema de unión a ligando híbrido de ubiquitina dividida, que comprende: un primer gen químico que codifica un primer polipéptido de fusión que comprende: los segmentos P1, Cub-Z y RM en un orden en el que Cub-Z está más próximo al extremo N-terminal del primer polipéptido de fusión que RM, en el que P1 es un polipéptido de unión a ligando que se une al primer ligando R1, Cub es un subdominio carboxi-terminal de ubiquitina, Z es un resto aminoácido distinto de metionina y RM es resto indicador, un segundo gen químico que codifica un segundo polipéptido de fusión, que comprende: los segmentos Nux y P2, en el que Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida y P2 es un polipéptido especificado por el usuario para el ligando candidato; y un sistema proteolítico dependiente de ubiquitina que comprende una proteasa específica de ubiquitina (UBP) de la regla del N-terminal; permitir que el ligando híbrido candidato se una al polipéptido de unión a ligando P1 del primer polipéptido de fusión a través del primer ligando R1 y que contacte con el polipéptido especificado por el usuario P2 del segundo polipéptido de fusión a través del ligando candidato R2 de modo que, cuando el polipéptido especificado por el usuario P2 se une al ligando candidato R2, los dominios Nux y Cub se asocian para formar un resto de ubiquitina reconstituida y la proteasa específica de ubiquitina escinde el enlace peptídico de Cub-Z para liberar un fragmento que contiene RM, siendo dicho fragmento susceptible a la degradación proteolítica dependiente de ubiquitina de la regla del N-terminal; medir el nivel del RM en presencia del ligando híbrido candidato en comparación con el nivel del RM en ausencia del ligando híbrido, en el que una disminución en el nivel del RM en presencia del ligando híbrido en comparación con el nivel del RM en ausencia del ligando híbrido indica que el polipéptido especificado por el usuario P2 se une al ligando candidato R2; identificar el ligando híbrido candidato que causa una disminución en nivel del RM en presencia del ligando híbrido en comparación con el nivel del RM en ausencia del ligando híbrido, identificando de este

modo un ligando que se une a un polipéptido especificado por el usuario.

[0057] Un aspecto relacionado de la invención proporciona un método de identificación de un ligando que se une a un polipéptido especificado por el usuario, que comprende: proporcionar una población de ligando híbrido candidato que tiene la fórmula general R1-Y-R2, en la que R1 es un primer ligando, R2 es un ligando candidato e Y es un enlazador; poner en contacto cada ligando híbrido candidato individual con un sistema de unión a ligando híbrido de ubiquitina dividida que comprende: un primer gen quimérico que codifica un primer polipéptido de fusión que comprende: los segmentos Nux y P1, en el que Nux es el subdominio amino-terminal de una ubiquitina de tipo silvestre o un subdominio amino-terminal de ubiquitina mutante de asociación reducida, y P1 es un polipéptido que se une al primer ligando R1 del ligando híbrido, un segundo gen quimérico que codifica un segundo polipéptido de fusión que comprende: los segmentos P2, Cub-Z y RM, en un orden en el que Cub-Z está más próximo al extremo N-terminal del primer polipéptido de fusión que RM, en el que P2 es un polipéptido de unión a ligando especificado por el usuario para el ligando candidato R2 del ligando híbrido, Cub es un subdominio carboxi-terminal de ubiquitina, Z es un resto aminoacídico distinto de metionina y RM es un resto indicador; y un sistema proteolítico dependiente de ubiquitina que comprende una proteasa específica de ubiquitina (UBP) de la regla del N-terminal; permitir que el ligando híbrido candidato se una al primer polipéptido de unión a ligando P1 del primer polipéptido de fusión a través del primer ligando R1 y que contacte con el polipéptido especificado por el usuario P2 del segundo polipéptido de fusión a través del ligando candidato R2 de modo que, cuando el polipéptido especificado por el usuario P2 se une al ligando candidato R2, los dominios Nux y Cub se asocian para formar un resto de ubiquitina reconstituida y la proteasa específica de ubiquitina escinde el enlace peptídico de Cub-Z para liberar un fragmento que contiene RM, siendo dicho fragmento susceptible a la degradación proteolítica dependiente de ubiquitina de la regla del N-terminal; medir el nivel del RM en presencia del ligando híbrido candidato en comparación con el nivel del RM en ausencia del ligando híbrido, en el que una disminución en el nivel del RM en presencia del ligando híbrido en comparación con el nivel del RM en ausencia del ligando híbrido indica que el polipéptido especificado por el usuario P2 se une al ligando candidato R2; identificar el ligando híbrido candidato que causa una disminución en el nivel del RM en presencia del ligando híbrido en comparación con el nivel del RM en ausencia del ligando híbrido, identificando de este modo un ligando que se une a un polipéptido especificado por el usuario.

[0058] En una realización, P2 está codificado por un ácido nucleico de una biblioteca seleccionada del grupo que consiste en: biblioteca de oligonucleótidos sintéticos, una genoteca de ADNc, una genoteca de fragmentos de ADN genómico bacteriano y una genoteca de fragmentos de ADN genómico eucariota. En una realización, el sistema de unión a ligando híbrido de ubiquitina dividida se proporciona por una célula.

[0059] Otro aspecto de la invención proporciona un método para investigar la relación de estructura-actividad de un ligando respecto a un dominio de unión a ligando, que comprende: 1) proporcionar un ligando híbrido R1-Y-R2, en el que a) R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, derivado de 2,4-diaminopteridina o ciclosporina, o un derivado de los mismos con modificaciones estructurales minoritarias; b) Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es número entero de 2 a 25; y c) R2 representa un segundo ligando especificado por el usuario que es diferente de R1 y se selecciona de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o un nucleótido; 2) proporcionar células que comprenden una proteína de fusión que incluye: a) al menos un dominio de unión a ligando; y b) un dominio funcional heterólogo respecto al dominio de unión a ligando que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional; 3) en el que en la etapa 1) se proporciona una pluralidad de ligandos híbridos que comprenden variantes estructurales de dicho segundo ligando R2, o en la etapa 2) se proporciona una pluralidad de proteínas de fusión que comprenden variantes estructurales de dicho dominio de unión a ligando; 4) exponer dichas células que comprenden cada proteína de fusión a una cantidad eficaz de cada ligando híbrido de modo que el primer dominio funcional puede ponerse en las proximidades de un segundo dominio funcional, induciendo o permitiendo de este modo la detección de un acontecimiento detectable; 5) medir la presencia, cantidad o actividad de cualquier acontecimiento detectable así inducido o permitido en la etapa 4) investigando de este modo la relación de estructura-actividad entre dicho segundo ligando y el dominio de unión a ligando.

[0060] En una realización, dicho primer dominio funcional de (b) se selecciona de: un dominio de unión a ADN, un dominio de activación de la transcripción, un subdominio carboxi-terminal de una ubiquitina de tipo silvestre, un subdominio amino-terminal de una ubiquitina o un subdominio amino-terminal de ubiquitina mutante de asociación reducida.

[0061] Otro aspecto de la invención proporciona un método para identificar un ligando híbrido que tiene la estructura general R1-Y-R2 adecuada para un ensayo *in vivo*, en el que dicho ensayo implica: 1) el uso de un ligando híbrido y 2) de al menos un polipéptido de fusión que incluye: a) al menos un dominio de unión a ligando P; y b) un dominio funcional que por sí mismo no es capaz de inducir o permitir la detección del acontecimiento detectable; y en el que

dicho método implica las etapas de: 3) sintetizar una pluralidad de ligandos híbridos R1-Y-R2 que difieren por una pluralidad de enlazadores diferentes Y, en el que R1 y R2 son diferentes y al menos uno de R1 y R2 no es un péptido; y 4) ensayar cada ligando híbrido en dicha pluralidad de ligandos híbridos individualmente para determinar su eficacia para inducir o permitir la detección del acontecimiento detectable; y 5) seleccionar un ligando híbrido con un enlazador particular que posea una eficacia adecuada para inducir o permitir la inducción del acontecimiento detectable.

[0062] En una realización, dicho enlazador tiene la estructura general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es número entero de 2 a 25, y la pluralidad de enlazadores difieren en n. En otra realización, R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, derivado de 2,4-diaminopteridina o ciclosporina, o un derivado de los mismos con modificaciones estructurales minoritarias.

[0063] Otro aspecto de la invención proporciona un kit que comprende al menos un polinucleótido que incluye un fragmento de ADN unido a una secuencia codificante de un dominio funcional heterólogo respecto al fragmento de ADN que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional; que comprende además instrucciones para sintetizar un ligando híbrido de estructura general R1-Y-R2 y para clonar un dominio de unión a ligando en el polinucleótido, y para ensayar la unión entre el ligando híbrido y el dominio de unión a ligando, en el que R2 es diferente de R1, uno de R1 y R2 es un ligando no peptídico, y en el que uno de R1 y R2 se une a o inhibe una quinasa.

[0064] Otro aspecto de la invención proporciona un kit que comprende al menos un polinucleótido que incluye un fragmento de ADN unido a una secuencia codificante de un dominio funcional heterólogo respecto al fragmento de ADN que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional; que comprende además instrucciones para sintetizar un ligando híbrido de estructura general R1-Y-R2, y para clonar un dominio de unión a ligando en el polinucleótido, y para ensayar la unión entre el ligando híbrido y el dominio de unión a ligando, en el que R2 es diferente de R1, uno de R1 y R2 es un ligando no peptídico, y en el que Y es de la estructura general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es número entero de 2 a 25.

[0065] Otro aspecto de la invención proporciona un kit que comprende al menos un polinucleótido que incluye un fragmento de ADN unido a una secuencia codificante de un dominio funcional heterólogo respecto al fragmento de ADN que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional; que comprende además instrucciones para sintetizar un ligando híbrido de estructura general R1-Y-R2, y para clonar un dominio de unión a ligando en el polinucleótido, y para ensayar la unión entre el ligando híbrido y el dominio de unión a ligando, en el que R2 es diferente de R1, uno de R1 y R2 es un ligando no peptídico, y en el que el dominio funcional es el dominio carboxi-terminal o amino-terminal de ubiquitina.

[0066] Otro aspecto de la invención proporciona un kit, que comprende: 1) un compuesto de estructura general R1-Y-L, en la que Y es de la estructura general $(CH_2-X-CH_2)_n$ y L es un grupo químico que se sustituye fácilmente por un grupo químico diferente y 2) instrucciones para usar el compuesto para la síntesis de un ligando híbrido R1-Y-R2, en el que R1 es diferente de R2 y al menos uno de R1 y R2 no es un péptido.

[0067] Otro aspecto de la invención proporciona un método de trabajar que comprende: 1) la identificación de polipéptidos que se unen a un ligando híbrido de fórmula general R1-Y-R2, en la que Y es de la fórmula general $(CH_2-X-CH_2)_n$, R1 es diferente de R2, y al menos uno de R1 y R2 no es un péptido, X = O, S, SO o SO_2 , y en el que no se sabía previamente que dichos polipéptidos se unieran a dicho ligando híbrido, y 2) proporcionar el acceso a los datos, ácidos nucleicos o polipéptidos así obtenidos a otro grupo para su consideración.

[0068] En una realización, dicha identificación de polipéptidos se realiza usando cualquiera de los métodos adecuados de la presente invención.

[0069] Un aspecto relacionado de la invención proporciona un método de trabajar que comprende: 1) la identificación de al menos un ligando que se une a un polipéptido especificado por el usuario usando una pluralidad de ligandos híbridos de fórmula general R1-Y-R2 que difieren en al menos uno de R1 y R2, en la que R1 y R2 son ligandos, R1 es diferente de R2, al menos uno de R1 y R2 no es un péptido, Y es de la estructura general $(CH_2-X-CH_2)_n$, X = O, S, SO o SO_2 , y en el que no se sabía previamente que dichos polipéptidos se unieran a dicho polipéptido, y 2) proporcionar el acceso a los datos y ligandos obtenidos a partir de dicha identificación a otro grupo para su consideración.

[0070] En una realización preferida, dicha identificación de ligandos se realiza usando cualquiera de los métodos adecuados de la presente invención.

Breve descripción de las figuras

[0071]

- Figura 1.** Esquemas sintéticos y representaciones estructurales para GPC 285937, 285985, 286004, 286026 y 285993.
- Figura 2.** Sensograma y posterior determinación de la constante de disociación K_D para la unión del complejo de quinasa dependiente de ciclina (CDK) 4/Ciclina D1 (CDK4/D1) a un ligando híbrido basado en metotrexato usando un Biosensor Biacore 2000-SPR. Se acopló covalentemente DHFR a la superficie de una microplaca SPR y se dejó que se uniera el ligando híbrido (GPC 285985). Posteriormente, se bombearon soluciones de diferentes concentraciones del complejo CDK4/D1 (mostrado por diferentes curvas) sobre la superficie de la microplaca durante 300 s, seguido de tampón de procesamiento para controlar la disociación. Las características de unión del metotrexato a DHFR se tuvieron en cuenta para estimar la k_{as} y k_{dis} del ligando híbrido a CDK4/D1 y la K_D calculada.
- Figura 3.** Representaciones estructurales de GPC 285937, GPC 285985 y GPC 285993.
- Figura 4.** Un ejemplo de un ensayo de crecimiento en halo. Un halo visible de crecimiento celular de levadura en medio que carece de histidina indica la activación del gen indicador *HIS3* causada por la dimerización de las proteínas de fusión LexBD-DHFR y GalAD-GR2 en presencia de GPC 285937, pero no en presencia de DMSO solamente.
- Figura 5.** Activación del gen indicador *HIS3* por dimerización inducida por compuesto de las proteínas de fusión Lex A-BD-DHFR y Gal4-AD-GR2 en presencia de un ligando híbrido de la invención (GPC 285937) en comparación con un ligando híbrido de la técnica anterior Mtx-mdbt-Dex (mdbt: metadibenzotioéster). Imágenes de microscopio de medios de cultivo en las que los objetos circulares son células de levadura individuales y los hilos lanosos oscuros son Mtx-mdbt-Dex precipitado. La precipitación de Mtx-mdbt-Dex se observa a 100 μ M.
- Figura 6.** Influencia de restos enlazadores diferentes de ligandos híbridos y sus efectos biológicos. Un ligando híbrido de la invención (GPC 285937) emplea 3 grupos de etilenglicol (EG) como enlazador, lo que ofrece una superioridad aumentada sobre el enlazador de metadibenzotioéster presente en el ligando híbrido de la técnica anterior Mtx-mdbt-Dex promoviendo un mejor crecimiento global de la colonia.
- Figura 7.** Diferencia en el crecimiento de colonias de levadura en la exploración de placas en presencia de GPC 285937 o Mtx-mdbt-Dex. Las colonias que crecían en medios con Mtx-mdbt-Dex eran difícilmente detectables, mientras que los clones crecían visiblemente mejor en medios que contenían GPC 285937.
- Figura 8.** Curvas de crecimiento de cultivos de levadura expuestos a diferentes concentraciones del ligando híbrido GPC 285985 en medio que carece de histidina según se midieron por el consumo de oxígeno usando un OxoPlate (PreSens, Alemania). Los cultivos de levadura que expresan la proteína de fusión de CDK2 muestran curvas de crecimiento típicas con el tiempo. Por el contrario, los cultivos de levadura que expresan una proteína de fusión de CDK4 sólo muestran crecimiento a las altas concentraciones del ligando híbrido, confirmando la especificidad del ligando híbrido hacia CDK2.
- Figura 9.** Una representación de la proteína de fusión Sec62-DHFR-Cub-PLV unida a la membrana del retículo endoplásmico (RE). Mientras está unido a la membrana, el factor de transcripción PLV es incapaz de activar un gen indicador. Sin embargo, en la escisión del Cub-PLV después de la formación de una molécula de ubiquitina casi nativa, el resto indicador de PLV escindido es capaz de trasladarse al núcleo y activar un gen indicador apropiado.
- Figura 10.** Un ensayo del ligando híbrido GPC 285985 usando un sistema de triple híbrido de levadura en un ensayo de halo. La fila superior muestra el crecimiento de células transformadas con pBTM118c-DHRF y pGAD426c-hCDK2 (parte superior izquierda) o pGAD426c-hCDK4 (parte superior derecha) después de dos días en medio que carece de trp, leu e his después de la adición de 1 μ l de una solución de GPC 285985 en DMSO 1 mM. La última fila muestra el crecimiento después de dos días en medio que carece de trp y leu, his después de la adición de GPC 285985. En el medio que carece de histidina, sólo las células transformadas con pGAD426c-hCDK2 muestran un crecimiento detectable, mientras que en el medio que carece sólo de trp y leu, las células transformadas tanto con pGAD426c-hCDK2 (parte inferior izquierda) como con pGAD426c-hCDK4 (parte inferior derecha) forman poblaciones densas.
- Figura 11.** Pueden detectarse interacciones débiles después de periodos de crecimiento más prolongados. En un experimento análogo al experimento mostrado en la Figura 10, las células transformadas con pBTM118c-DHRF y pGAD426c-hCDK2 (panel izquierdo) o pGAD526c-hCDK4 (panel derecho) se incubaron durante seis días a 30°C en medio que carecía de trp, leu e his después de la adición de 1 μ l de una solución 1 mM de GPC 285985 disuelto en DMSO en el centro de cada placa de petri. Después

de este tiempo de incubación, la interacción de baja afinidad (900 μ M) entre CDK4 y GPC 285985 era capaz de permitir un crecimiento débil pero detectable. Por el contrario, las células que expresaban la proteína de fusión de CDK2 formaban poblaciones densas en las mismas condiciones.

- 5 **Figura 12.** Resultados de un ensayo de halo de alto rendimiento usando clones recuperados de una exploración genética de triple híbrido. Se exploró una biblioteca de proteínas de fusión para aislar genes que codificasen proteínas que se unían al ligando híbrido GPC 285985. La tabla presenta una muestra del análisis realizado en 2811 clones positivos iniciales. 102 clones mostraban un crecimiento específico de compuesto. La identidad de todos los clones se confirmó por secuenciación y contenían genes que codifican CDK2 y otros genes.
- 10 **Figura 13.** Un plásmido aislado que codifica la proteína GPC761 expresada como proteína de fusión con GAL4 AD (aislado de una exploración genética de triple híbrido) se cotransformó con pBTM118c-DHFR en la cepa de levadura L40. Se realizó un ensayo de halo para validar y caracterizar adicionalmente e investigar la relación de estructura-actividad entre la interacción entre esta proteína y el ligando híbrido usado para la exploración inicial. Sólo el ligando híbrido que comprende el inhibidor de CDK2 activo GPC 285985 (panel izquierdo) permitía el crecimiento de células en medio que carece de trp, leu e his, mientras que la variante estructural GPC 285993 (que no se une a CDK2) era ineficaz promoviendo el crecimiento en este ensayo y, por lo tanto, no se unía a la proteína GPC761.
- 15 **Figura 14.** El rendimiento de los ligandos híbridos de la invención en células de mamíferos se ensayó como se ha descrito en el ejemplo 11. El gen indicador CAT se activa como se muestra por la presencia de un precipitado coloreado en el control positivo (Figura 14A). Las células que expresan las fusiones de DHFR y GR2 con el ligando híbrido de dimerización respectivo GPC 285937 (Figura 14B) también muestra un precipitado coloreado, pero no cuando está ausente GPC 285937 (Figura 14C).
- 20 **Figura 15.** Sistema de ensayo de triple híbrido basado en la tecnología detectora de proteína ubiquitina dividida. Se construyeron dos proteínas de fusión, una que consistía en la mitad N-terminal de ubiquitina (Nub) y una proteína presa (XY), y otra que consistía en la mitad C-terminal de ubiquitina (Cub), una proteína cebo (DHFR) y el resto del indicador (R). La asociación de presa y cebo por unión mutua a la molécula pequeña híbrida mtx-xy reconstituye una estructura de ubiquitina casi nativa (UBI) reconocida por la proteasa específica de ubiquitina (UBP), por lo que el resto indicador se escinde de la proteína de fusión. La escisión del resto indicador de la proteína de fusión puede detectarse por varias técnicas, por ejemplo, sin limitación, transferencia de Western, escisión o desestabilización del indicador a través de consideraciones de la regla del N-terminal (teniendo R un aminoácido distinto de metionina en su extremo N-terminal) o proporcionando un factor de transcripción como R y permitiendo su translocación al núcleo.
- 25 **Figura 16.** Efectos de la longitud del enlazador (número de repeticiones de PEG en el enlazador) sobre la funcionalidad según se midió por la actividad biológica en un ensayo de halo de triple híbrido. Sólo se observó crecimiento en halo de levadura en células en presencia de GPC 286026 (5 unidades de PEG como enlazador), pero no en presencia de GPC 286004 (3 unidades de PEG como enlazador).
- 30 **Figura 17.** Descripción del plásmido pACT2; una genoteca de ADNc de cerebro fetal humano se obtuvo comercialmente de Clontech, que se clonó en este vector y se usó posteriormente en experimentos de exploración. a. Un mapa de vector. b. Un mapa de restricción y sitio de clonación múltiple.
- 35 **Figura 17.** Descripción del plásmido pACT2; una genoteca de ADNc de cerebro fetal humano se obtuvo comercialmente de Clontech, que se clonó en este vector y se usó posteriormente en experimentos de exploración. a. Un mapa de vector. b. Un mapa de restricción y sitio de clonación múltiple.
- 40 **Figura 17.** Descripción del plásmido pACT2; una genoteca de ADNc de cerebro fetal humano se obtuvo comercialmente de Clontech, que se clonó en este vector y se usó posteriormente en experimentos de exploración. a. Un mapa de vector. b. Un mapa de restricción y sitio de clonación múltiple.

Mejor modo de llevar a cabo la invención

Descripción detallada de la invención

1. Visión de conjunto

- 45 **[0072]** En general, la invención proporciona un sistema de ensayo de triple híbrido y reactivos para la identificación del compañero de unión proteico de un agente farmacéutico pequeño seleccionado. Así mismo, la invención también proporciona métodos y reactivos para la identificación de un compañero de unión de agente farmacéutico pequeño de una proteína seleccionada. Una vez detectada, la invención proporciona además métodos para controlar la interacción del agente farmacéutico y su compañero de unión proteico, que pueden usarse para detectar competidores de la interacción.
- 50 **[0073]** De acuerdo con un aspecto de la invención, un compuesto que se une a un polipéptido diana conocido puede seleccionarse de una combinación/biblioteca de compuestos candidato. Preferiblemente, el compuesto es una molécula pequeña (véase la definición a continuación). En esta aspecto de la invención, cada molécula pequeña candidata (denominada "R2" en lo sucesivo) se une a una molécula pequeña conocida (denominada "R1" en lo sucesivo) a través de una secuencia enlazadora (denominada "Y" en lo sucesivo). Después, se deja que el compuesto R1-Y-R2 resultante
- 55 **[0073]** De acuerdo con un aspecto de la invención, un compuesto que se une a un polipéptido diana conocido puede seleccionarse de una combinación/biblioteca de compuestos candidato. Preferiblemente, el compuesto es una molécula pequeña (véase la definición a continuación). En esta aspecto de la invención, cada molécula pequeña candidata (denominada "R2" en lo sucesivo) se une a una molécula pequeña conocida (denominada "R1" en lo sucesivo) a través de una secuencia enlazadora (denominada "Y" en lo sucesivo). Después, se deja que el compuesto R1-Y-R2 resultante

- desencadena un acontecimiento biológicamente detectable. Si R2 interacciona con P2 con una afinidad lo bastante grande, entonces RS1 se pone en estrecha proximidad a RS2 mediante el efecto de formación de un puente del híbrido R1-Y-R2, desencadenando de este modo la activación de RS. Por lo tanto, poner en contacto el entorno (es decir, una célula) que contiene el RS, el P1-RS1-híbrido y el P2-RS2-híbrido con una combinación/biblioteca de híbridos de R1-Y-R2 y observar la activación de RS facilita el aislamiento de híbridos de R1-Y-R2, en los que R2 es capaz de unirse específicamente a P2.
- [0074]** En una realización, el RS es un sistema indicador basado en transcripción, tal como un sistema de doble híbrido de levadura. En otra realización relacionada, el RS es un sistema indicador basado en ubiquitina dividida.
- [0075]** En una realización, la secuencia enlazadora es particularmente adecuada para el uso *in vivo* del compuesto químico debido a su solubilidad aumentada y a su permeabilidad de membrana mejorada.
- [0076]** En una realización, la interacción de P1-R1 es una interacción no covalente. En una realización alternativa, la interacción de P1-R1 da como resultado un enlace covalente.
- [0077]** En una realización, la biblioteca química está sintetizada. En otra realización, la biblioteca química es de fuentes naturales.
- [0078]** De acuerdo con otro aspecto de la invención, un polipéptido que se une a una molécula pequeña diana conocida R2 puede seleccionarse de una biblioteca/bibliotecas de polipéptidos de ensayo. En este aspecto, la molécula pequeña diana R2 se une por una secuencia enlazadora Y a una molécula pequeña conocida R1 para formar un compuesto híbrido R1-Y-R2, que después se permite que contacte con el polipéptido P1, el compañero de unión conocido de la molécula pequeña conocida R1, fusionada a RS1, en un entorno adecuado. Se proporciona por traducción al mismo entorno una biblioteca o bibliotecas de polipéptidos de ensayo P2, cada uno fusionado a RS2. La unión entre la molécula pequeña diana R2 y cualquier polipéptido P2 miembro de la biblioteca/bibliotecas pondrá al híbrido P2-RS2 en las proximidades del híbrido P1-RS1, desencadenando de este modo la activación de un sistema indicador RS. Por lo tanto, poner en contacto células que contienen el RS, el híbrido P1-RS1 y una combinación/biblioteca de híbridos de P2-RS2 con el híbrido de R1-Y-R2 y observar la activación de RS facilita el aislamiento de híbridos de P2-RS2, en los que P2 es capaz de unirse específicamente a R2.
- [0079]** En una realización, el RS es un sistema indicador basado en transcripción, tal como un sistema de doble híbrido de levadura. En otra realización relacionada, el RS es un sistema indicador basado en ubiquitina dividida.
- [0080]** En una realización, la secuencia enlazadora es particularmente adecuada para el uso *in vivo* del compuesto químico debido a su solubilidad aumentada y a su permeabilidad de membrana mejorada.
- [0081]** En una realización, la interacción de P1-R1 es una interacción no covalente. En una realización relacionada, la interacción de P1-R1 da como resultado un enlace covalente.
- [0082]** En una realización, la biblioteca de polipéptidos es una genoteca de ADNc o una genoteca de ADN genómico. En otra realización, la biblioteca de polipéptidos se sintetiza aleatoriamente o semialeatoriamente. La biblioteca puede contener diferente número de miembros, preferiblemente de 2 a 10 miembros, o de 10 a 500 miembros, de 500 a 10.000 miembros o más de 10.000 miembros.
- [0083]** Los métodos descritos anteriormente no sólo son adecuados para identificar un miembro desconocido de una pareja de polipéptido-ligando (método de exploración), sino que también son adecuados para determinar si un polipéptido dado se une a un ligando dado (ensayo o método de ensayo).
- [0084]** De acuerdo con otro aspecto más de la invención, se proporciona un kit para detectar y/o seleccionar interacciones entre polipéptidos y moléculas pequeñas usando uno de los métodos mencionados anteriormente.
- [0085]** De acuerdo con otro aspecto de la invención, se proporciona un método para la investigación farmacéutica, en el que se controlan interacciones entre polipéptidos y moléculas pequeñas para facilitar la caracterización adicional y/o la optimización de la unión de al menos uno de los compañeros de unión identificados. Esto puede ser útil en una diversidad de situaciones. Por ejemplo, muchos fármacos o compuestos químicos tienen efectos secundarios indeseables perceptibles, a veces incluso graves. Esto está probablemente causado por el hecho de que el fármaco puede unirse indiscriminadamente a proteínas distintas de la diana deseada. La presente invención proporciona un método para identificar todos los compañeros de unión potenciales de un fármaco o compuesto químico dado, proporcionando de este modo una base para diseñar otros fármacos relacionados que no se unan a estas dianas no deseadas para evitar los efectos secundarios indeseables. En otros casos, un fármaco puede tener cierta eficacia para determinadas condiciones, pero se desconoce el mecanismo de acción del fármaco, por lo tanto, es difícil optimizar el fármaco para una mejor eficacia. La presente invención proporciona un método para identificar la diana del fármaco, ofreciendo de este modo un medio para un estudio adicional de la biología y de las rutas de señalización relacionadas, de modo que pueda conseguirse la optimización del fármaco basándose en el conocimiento obtenido a través de la investigación sobre esas rutas de señalización. Además, puede usarse la información sobre la unión de ligandos a dominios de unión a ligando polipeptídicos que se recoge poniendo en práctica los métodos de la invención para comprender o comprender adicionalmente la función o los efectos secundarios de un ligando en un entorno biológico o terapéutico. La información recogida de este modo puede usarse, por ejemplo, para proporcionar una prescripción más

- informada de medicamentos que comprenden el ligando o con medicamentos adicionales apropiados para proporcionar terapias de combinación más eficaces. Por lo tanto, la presente invención puede usarse para identificar o producir uno o más de cualquiera de los siguientes: un compuesto con un efecto biológico conocido, un compuesto con un mecanismo de acción desconocido, un compuesto que se une a más de un polipéptido, un compuesto candidato a fármaco o un compuesto que se une a una proteína desconocida.
- [0086]** La presente invención también proporciona ligandos híbridos que se unen a o inhiben una quinasa. Por ejemplo, R2 puede ser un compuesto seleccionado de la Tabla 2, que es una lista de compuestos que se sabe que se unen a o inhiben quinasas, o un derivado del mismo con modificaciones estructurales minoritarias. Una diana de quinasa típica puede ser una quinasa dependiente de ciclina.
- [0087]** Además, la presente invención también proporciona un método para identificar nuevos moduladores de ciertas proteínas conocidas y un método para producir formulaciones farmacéuticas de dichos moduladores.
- [0088]** Otro aspecto de la invención proporciona un método para identificar un compuesto que inhiba la interacción entre un ligando y un polipéptido, en el que la interacción se identifica usando cualquier método adecuado de la presente invención, que comprende: 1) identificar, por cualquiera de los métodos adecuados de la presente invención, un polipéptido que interaccione con un ligando especificado por el usuario, o identificar un ligando que interaccione con un polipéptido especificado por el usuario; 2) proporcionar un entorno en el que se produzca dicha interacción; 3) poner en contacto el entorno con un compuesto de ensayo; 4) determinar si dicho compuesto de ensayo inhibe dicha interacción, identificando de este modo un compuesto que inhiba la interacción entre un ligando y un polipéptido.
- [0089]** En una realización, el ligando es un ligando no peptídico. En una realización preferida, el ligando es de la estructura general R1-Y-R2, en la que R1, Y y R2 son como se han definido anteriormente.
- [0090]** En una realización, el compuesto de ensayo es de una biblioteca abigarrada que, por ejemplo, puede ser una genoteca de ácidos nucleicos (ADNc, ADN genómico, EST, etc.) que codifican polipéptidos; una biblioteca de polipéptidos (sintética, natural, aleatoria, semialeatoria, etc.); una biblioteca de químicos pequeños (naturales, sintéticos, etc.).
- [0091]** En una realización, el entorno es una célula. En una realización relacionada, el entorno contiene cualquiera de los sistemas de exploración de ligandos híbridos adecuados de la presente invención (incluyendo sistemas indicadores).
- [0092]** El efecto inhibitor del compuesto de ensayo puede evaluarse basándose en el cambio de estado del sistema indicador (véanse descripciones detalladas a continuación).
- [0093]** Este método puede ser útil en una diversidad de situaciones. Por ejemplo, si se identifica inicialmente que un compuesto químico pequeño posee cierta actividad biológica cuando se administra a una célula, su diana o dianas proteicas pueden identificarse. En el caso de que estén presentes múltiples dianas y sólo se desee una interacción diana (por ejemplo, otras interacciones con proteínas diana conducen a efectos secundarios indeseables), puede identificarse un compuesto de ensayo usando este método para que pueda bloquear específicamente aquellas interacciones indeseables, permitiendo al mismo tiempo que se produzca la interacción deseada. En otro escenario, después de la identificación de la diana polipeptídica de un ligando conocido, puede identificarse un compuesto usando el presente método para bloquear la interacción entre dicho ligando y polipéptido, para eliminar el efecto indeseable de la interacción de ligando-polipéptido o para controlar de forma reversible dicha interacción.
- [0094]** Otro aspecto de la invención proporciona un método para identificar una secuencia polipeptídica que se une a un ligando especificado por el usuario, que comprende: 1) proporcionar un ligando híbrido con la estructura general R-Y-R, en la que R es un ligando especificado por el usuario e Y es un enlazador, preferiblemente un enlazador que tiene la fórmula general $(-\text{CH}_2-\text{X}-\text{CH}_2-)_n$, en el que X y n son como se han definido anteriormente; 2) introducir el ligando híbrido en una población de células, conteniendo cada célula un sistema de exploración de ligando como se ha definido anteriormente, o un sistema basado en ubiquitina dividida Nux-Cub como se ha definido anteriormente, en el que tanto P1 como P2 (como se han definido anteriormente) representan el mismo polipéptido de ensayo; 3) permitir que el ligando híbrido contacte con P1 y P2 en dicho sistema de exploración de ligando; 4) identificar una célula de unión a ligando positiva en la que se produzca un cambio detectable en el estado del sistema indicador del sistema de exploración de ligando; identificar de este modo un ácido nucleico que codifique el polipéptido de ensayo.
- [0095]** En un aspecto relacionado de la invención, se proporciona un método para determinar si un ligando se une a un polipéptido, que comprende: 1) proporcionar un ligando híbrido con la estructura general R-Y-R, en la que R es un ligando especificado por el usuario e Y es un enlazador, preferiblemente un enlazador que tiene la fórmula general $(-\text{CH}_2-\text{X}-\text{CH}_2-)_n$, en la que X y n son como se han definido anteriormente; 2) introducir el ligando híbrido en un entorno que contenga un polipéptido de ensayo, en el que la multimerización (preferiblemente, dimerización) del polipéptido conduce a un cambio detectable; 3) determinar si se produce dicho cambio detectable, determinando de este modo si el ligando se une al polipéptido de ensayo.
- [0096]** En un aspecto relacionado, puede usarse un método similar para determinar si un polipéptido conocido interacciona con un ligando híbrido de ensayo.
- [0097]** En una realización, el cambio detectable es una actividad enzimática del polipéptido de ensayo, estando dicha

actividad solamente presente cuando dicho polipéptido está multimerizado (por ejemplo, dimerizado). En una realización relacionada, el polipéptido puede unirse a cualquiera de los sistemas de exploración de ligandos híbridos adecuados descritos anteriormente de modo que la multimerización del polipéptido por el ligando híbrido conduzca a la activación del sistema indicador.

[0098] En una realización, el polipéptido es una enzima que se inactiva como monómero y que sólo se activa como multímero, preferiblemente un dímero. En esta realización, puede ser suficiente usar solamente un único polinucleótido en un método de la invención. Por ejemplo, cuando se está buscando un nuevo ligando para un polipéptido de interés para el que ya se conoce un ligando, se podría usar un polinucleótido que codifica el polipéptido de interés fusionado a una enzima que se activa solamente como multímero, preferiblemente un dímero, y que no se dimeriza espontáneamente (por ejemplo, un mutante de afinidad reducida). Si este polipéptido de fusión se pone en contacto con un ligando híbrido R1-Y-R2 de la invención, donde R1 es el ligando conocido para el polipéptido de interés y R2 es un ligando de ensayo, la actividad de la enzima sólo se manifestará si el ligando de ensayo se une al polipéptido de interés.

[0099] En una realización, el entorno es una célula.

[0100] En una realización, el polipéptido comprende un receptor, preferiblemente un receptor que requiere multimerización para ser funcional o activarse, tal como un receptor que contenga un dominio citoplasmático de uno de los diversos receptores de membrana de superficie celular que se describen en el documento WO 94/18317. Por ejemplo, muchos de estos dominios son tirosina quinasas o están formando complejos con tirosina quinasas, por ejemplo, CD3 ζ , IL-2R, IL-3R, etc. Para una revisión véase Cantley, et al., Cell (1991) 64, 281. Los receptores de tirosina quinasa que se activan por entrecruzamiento, por ejemplo, dimerización (basándose en la nomenclatura propuesta por primera vez por Yarden y Ulrich, Annit. Rev. Bioclie7n. (1998) 57, 443) incluyen la subclase I: EGF-R, ATR2/neu, HER2/neu, Her3/c-erbB-3, Xmrk; subclase II: insulina-R, receptor del factor de crecimiento de tipo insulina IGF R, IRR; subclase III: PDGF-R-A, PDGF-R-B, CSF R (M-CSF/c-Fms), c-kit, STK-1/Flk-2; y subclase IV: FGF-R, flg [FGFJ ácido, bek [FGF básico]; tirosina quinasas neurotróficas: la familia de Trk incluye NGF-R, Ror1,2. Los receptores que se asocian con tirosina quinasas tras el entrecruzamiento incluyen la familia de CD3 ζ : CD3 ζ y CD3 η (que se encuentran principalmente en células T, se asocia con Fyn), cadenas β y γ de Fc ϵ R1 (que se encuentra principalmente en mastocitos y basófilos); cadena γ de Fc γ RIII/CD16 (que se encuentra principalmente en macrófagos, neutrófilos y células asesinas naturales); CD3 γ , δ y ε (que se encuentra principalmente en células T); Ig-a/MB-1 e Ig-P/B29 (que se encuentran principalmente en células B). Como alternativa, puede utilizarse un receptor de citocinas para detectar interacciones de ligando y receptor como se describe en Eyckerman et al (Nature Cell Biology 2001; 3: 1114-1119).

2. Definiciones

[0101] El término “agonista”, como se usa en la presente memoria, pretende referirse a un agente que mimetiza o regula positivamente (por ejemplo, potencia o complementa) la bioactividad de una proteína de interés, o un agente que facilita o promueve (por ejemplo, potencia o complementa) una interacción entre polipéptidos o entre un polipéptido y otra molécula (por ejemplo, un esteroide, hormona, ácido nucleico, molécula pequeña, etc.). Un agonista puede ser una proteína de tipo silvestre o derivado de la misma que tenga al menos una bioactividad de la proteína de tipo silvestre. Un agonista también puede ser una molécula pequeña que regule positivamente la expresión de un gen o que aumente al menos una bioactividad de una proteína. Un agonista también puede ser una proteína o molécula pequeña que aumente la interacción de un polipéptido de interés con otra molécula, por ejemplo, un ácido nucleico o péptido diana.

[0102] Un “antagonista”, como se usa en la presente memoria, pretende referirse a un agente que regula negativamente (por ejemplo, suprime o inhibe) la bioactividad de una proteína de interés, o un agente que inhibe/suprime o reduce (por ejemplo, desestabiliza o disminuye) la interacción entre polipéptidos u otras moléculas (por ejemplo, esteroides, hormonas, ácidos nucleicos, etc.). Un antagonista puede ser un compuesto que inhiba o disminuya la interacción entre una proteína y otra molécula, por ejemplo, un péptido diana, tal como la interacción entre ubiquitina y su sustrato. Un antagonista también puede ser un compuesto que regule negativamente la expresión de un gen de interés o que reduzca la cantidad de la proteína de tipo silvestre presente. Un agonista también puede ser una proteína o molécula pequeña que disminuya o inhiba la interacción de un polipéptido de interés con otra molécula, por ejemplo, un ácido nucleico o péptido diana.

[0103] El término “alelo”, que se usa indistintamente en la presente memoria con “variante alélica”, se refiere a las formas alternativas de un gen o porciones del mismo. Los alelos ocupan el mismo locus o posición en cromosomas homólogos. Cuando un sujeto tiene dos alelos idénticos de un gen, se dice que el sujeto es homocigoto para ese gen o alelo. Cuando el sujeto tiene dos alelos diferentes de un gen, se dice que el sujeto es heterocigoto para el gen. Los alelos de un gen específico pueden diferir entre sí en un solo nucleótido, o varios nucleótidos, y pueden incluir sustituciones, deleciones y/o inserciones de nucleótidos. Un alelo de un gen también puede ser una forma de un gen que contiene mutaciones.

[0104] La expresión “acontecimiento biológicamente detectable” es una expresión general usada para describir cualquier acontecimiento biológico que pueda detectarse en un sistema de ensayo, tal como, por ejemplo, sin limitación, en un ensayo de doble híbrido de levadura basado en transcripción, un ensayo de ubiquitina dividida, etc. Un acontecimiento biológicamente detectable se refiere a un acontecimiento que cambia una propiedad medible de un sistema biológico, por ejemplo, sin limitación, la absorbancia de luz a cierta longitud de onda, la emisión de luz después de la estimulación, la presencia/ausencia de determinado resto molecular en el sistema, la resistencia/capacitancia

eléctrica, etc., dependiendo dicho acontecimiento de otro, posiblemente una propiedad de interés del sistema biológico no medible o medible menos fácilmente, por ejemplo, sin limitación, la presencia o ausencia de una interacción entre dos proteínas. Preferiblemente, el cambio en la propiedad medible provocado por el acontecimiento biológicamente detectable es grande en comparación con variaciones naturales en la propiedad medible del sistema. Los ejemplos incluyen el color amarillo resultante de la acción de la β -galactosidasa sobre el o-nitrofenil-b-D-galactopiranosido (ONPG) (J. H. Miller, *Experiments in Molecular Genetics*, 1972) desencadenada por activación transcripcional del gen lacZ de *E. coli*, que codifica la β -galactosidasa, por reconstitución de un factor de transcripción tras la unión de las dos proteínas fusionadas a los dos dominios funcionales del factor de transcripción. Otros ejemplos de acontecimientos biológicamente detectables son claramente evidentes para el experto en la materia. Como alternativa, pueden inducirse y detectarse otras funciones biológicas después de la oligomerización, preferiblemente dimerización, de los dominios funcionales. Por ejemplo, regulación de la transcripción, modificación secundaria, localización celular, exocitosis, señalización celular, degradación o inactivación de proteínas, viabilidad celular, apoptosis regulada, velocidad de crecimiento, tamaño celular. Dichos acontecimientos biológicos también pueden controlarse por una diversidad de medios directos e indirectos, incluyendo actividades particulares asociadas con proteínas individuales tales como actividad proteína quinasa o fosfatasa, actividad reductasa, actividad ciclooxigenasa, actividad proteasa o cualquier otra reacción enzimática dependiente de la asociación de subunidades. Además, se puede proporcionar la asociación de proteínas G con una proteína receptora asociada con el ciclo celular, por ejemplo, ciclinas y cdc quinasas, o enzimas detoxificantes multiunitarias.

[0105] La “actividad biológica” o “bioactividad” o “actividad” o “función biológica”, que se usan indistintamente para los fines de la presente memoria, se refieren a una función catalítica, efectora, antigénica, de marcaje molecular o interacción molecular que se realiza directa o indirectamente por un polipéptido (ya sea en su conformación nativa o desnaturalizada) o por cualquier subsecuencia del mismo.

[0106] Las expresiones “muerte celular”, “destrucción celular” o “necrosis” se refieren al fenómeno de células que mueren como resultado de una pérdida extrínsecamente impuesta de una función celular particular esencial para la supervivencia de la célula.

[0107] “Células”, “células hospedadoras” o “células hospedadoras recombinantes” son expresiones usadas indistintamente en la presente memoria. Se entiende que dichas expresiones se refieren no sólo a una célula objeto particular, sino a la progenie o progenie potencial de dicha célula. Debido a que pueden producirse ciertas modificaciones en generaciones sucesivas debido a influencias ambientales o mutaciones, dicha progenie puede de hecho no ser idéntica a la célula parental, pero aun así se incluye dentro del alcance de la expresión según se usa en la presente memoria.

[0108] El término “caracterizar”, como se usa en la presente memoria, significa un estudio detallado de una molécula pequeña, un polipéptido o un ácido nucleico (polinucleótido) que codifica un polipéptido para revelar información biológica y química relevante. Esta información generalmente incluye una o más de, pero sin limitación, las siguientes: información de secuencia para proteína y ácido nucleico, información de estructura primaria, secundaria, terciaria y cuaternaria, peso molecular, solubilidad en diversos disolventes, actividad enzimática o de otro tipo, punto de concentración isoeléctrica, afinidad de unión a otras moléculas, compañeros de unión, estabilidad, patrón de expresión, distribución tisular, localización subcelular, regulación de la expresión, papeles evolutivos, fenotipos de animales transgénicos que sobreexpresan o carecen de un polipéptido o ácido nucleico, tamaño del ácido nucleico y propiedad de hibridación del ácido nucleico. Pueden usarse una diversidad de protocolos y metodologías de química, biología molecular y celular convencionales, tales como electroforesis en gel, electroforesis capilar, clonación, digestión con enzimas de restricción, generación de perfiles de expresión por hibridación, cromatografía de afinidad, HPLC, concentración isoeléctrica, espectrometría de masas, secuenciación automática y generación de animales transgénicos, cuyos detalles pueden encontrarse en muchos manuales de laboratorio de biología molecular y química convencionales (véanse a continuación). Las técnicas que emplean la hibridación de ácidos nucleicos pueden utilizar, por ejemplo, bibliotecas de ácidos nucleicos dispuestas en matrices, tales como oligonucleótidos, ADNc u otros (véase, por ejemplo, el documento US 5.837.832).

[0109] La expresión “químicamente similar” se usa para referirse a compuestos químicos con estructuras químicas y/o propiedades químicas similares. La similitud puede juzgarse por comparación entre dos compuestos de varias características, tales como la carga electrónica, el tamaño estérico, la estereoquímica, la capacidad donadora/aceptora de enlaces de hidrógeno y la polaridad (es decir, hidrofobicidad/hidrofiliidad). Por ejemplo, los aminoácidos químicamente similares tendrían cadenas laterales que, a juzgar por al menos tres, cuatro o preferiblemente las cinco características, se clasifican de la misma forma. Por ejemplo, en condiciones fisiológicas, la glicina y la alanina son similares a juzgar por las cinco características, la glicina y la fenilalanina difieren solamente a juzgar por el tamaño estérico, la glicina y la tirosina difieren por el tamaño estérico y la capacidad donadora de enlaces de hidrógeno, y la glicina y el ácido glutámico difieren por el tamaño estérico, la carga, la polaridad y la capacidad aceptora de enlaces de hidrógeno. Por ejemplo, los esteroides son generalmente similares en términos de conformación, polaridad, estereoquímica, carga, tamaño estérico, etc., aunque algunos esteroides (individualmente o como subclases) pueden diferir ligeramente de los esteroides “medios” (por ejemplo, los alcaloides esteroidales están típicamente cargados en condiciones fisiológicas).

- 5 **[0110]** En ciertas realizaciones, los compuestos moleculares pequeños químicamente similares comparten grupos funcionales y/o sistemas de anillos similares y, por lo tanto, muestran una combinación de elementos estructurales dispuestos en orientaciones o conformaciones similares, definiendo de este modo una clase estructural de compuestos que difieren ligeramente, por ejemplo, por sustituyentes adjuntos al núcleo estructural, o por ligeras variaciones en el núcleo estructural (tales como cambios en el tamaño del anillo, sustituciones de heteroátomos, homologación, etc.). Por ejemplo, los antibióticos beta-lactámicos comparten todos un anillo lactámico de cuatro miembros, los antibióticos macrólidos tienen una lactona macrocíclica (por ejemplo, de 10 a 18 miembros) sustituida con múltiples grupos metilo y/o hidroxilo (pudiendo algunos de los últimos estar hidroxilados), los péptidos son cadenas de alfa-aminoácidos unidos por enlaces amida, etc., y cada uno de dichos grupos de compuestos comprende miembros químicamente similares.
- 10 **[0111]** La expresión “derivado con modificaciones minoritarias” con respecto a un compuesto químico precursor, por ejemplo, una molécula pequeña, ligando, ligando híbrido, péptido o polipéptido, se usa para referirse a compuestos químicos que son químicamente similares al compuesto químico precursor. Preferiblemente, un derivado con modificaciones minoritarias tendrá modificaciones estructurales minoritarias y, por lo tanto, pueden considerarse como “variantes estructurales” del compuesto original. En general, dichas modificaciones estructurales minoritarias se realizan para obtener un compuesto con propiedades similares globales en comparación con el compuesto precursor, pero con un cambio con respecto a cierta propiedad del compuesto precursor que sea desventajosa o no deseada. Por ejemplo, puede añadirse una cadena lateral hidrófila a cierto compuesto químico para aumentar su solubilidad, conservando al mismo tiempo una actividad biológica deseada ya que la cadena lateral se añade de modo que no interfiera con la unión entre el compuesto y su diana biológica.
- 15 **[0112]** Un “polipéptido quimérico”, “polipéptido de fusión” o “proteína de fusión” es una fusión de una primera secuencia de aminoácidos que codifica un primer polipéptido con una segunda secuencia de aminoácidos que define un dominio (por ejemplo, porción polipeptídica) extraño para y no sustancialmente homólogo a ningún dominio del primer polipéptido. Dicha segunda secuencia de aminoácidos puede presentar un dominio que se encuentra (aunque en un polipéptido diferente) en un organismo que también expresa el primer polipéptido, o puede ser una fusión “interespecie”, “intergénica” etc. de estructuras polipeptídicas expresadas por diferentes clases de organismos. Al menos uno del primer y segundo polipéptidos puede ser también parcialmente o completamente sintético o aleatorio, es decir, no estar previamente identificado en ningún organismo.
- 20 **[0113]** El término “clonar”, como se usa en la presente memoria, como será evidente para el experto en la materia, puede entenderse como obtener copias exactas de una molécula polinucleotídica dada usando la tecnología de ADN recombinante. Además, la expresión “clonar en” puede entenderse como insertar una primera secuencia polinucleotídica dada en una segunda secuencia polinucleotídica, preferiblemente de modo que se obtenga como resultado una unidad funcional que combine las funciones del primer y segundo polinucleótidos, por ejemplo, sin limitación, un polinucleótido a partir del cual puede proporcionarse por traducción una proteína de fusión, comprendiendo dicha proteína de fusión secuencias de aminoácidos codificadas por la primera y segunda secuencias polinucleotídicas. Pueden encontrarse detalles de la clonación molecular en varios libros de protocolos de laboratorio usados comúnmente, tales como *Molecular Cloning: A Laboratory Manual*. 2ª Ed., ed. por Sambrook, Fritsch y Maniatis (Cold Spring Harbor Laboratory Press: 1989).
- 25 **[0114]** El término “clonar”, como se usa en la presente memoria, como será evidente para los expertos en la materia, también puede entenderse como obtener poblaciones idénticas o casi idénticas de células que poseen una propiedad dada común, tal como la presencia o ausencia de un marcador fluorescente, o un marcador de selección positiva o negativa. La población de células idénticas o casi idénticas obtenida por clonación también se denomina “clon”. Los métodos de clonación celular son bien conocidos en la técnica como se describe en muchos manuales de laboratorio comúnmente disponibles (véase *Current Protocols in Cell Biology*, Edición en CD-ROM, ed. por Juan S. Bonifacio, Jennifer Lippincott-Schwartz, Joe B. Hardford y Kenneth M. Yamada, John Wiley & Sons, 1999).
- 30 **[0115]** La expresión “exploración de complementación”, como se usa en la presente memoria, se refiere a la exploración genética para uno o varios genes o ADN de fuente que pueden conferir cierto fenotipo especificado que no existirá sin la presencia de dichos uno o varios genes o ADN de fuente. Habitualmente se realiza *in vivo*, introduciendo en células que carecen del fenotipo especificado una biblioteca de ADN de fuente a explorar e identificando las células que hayan obtenido un ADN de fuente y ahora muestren el fenotipo especificado. Como alternativa, podría realizarse *in vivo* inactivando aleatoriamente genes en el genoma de la célula que carece del fenotipo especificado e identificando las células que hayan perdido la función de ciertos genes y muestren el fenotipo especificado. Sin embargo, también puede realizarse una exploración de la complementación *in vitro* en sistemas sin células, ensayando cada candidato individualmente o como combinaciones de individuos.
- 35 **[0116]** La “recuperación de un clon de la célula... en condiciones en las que pueda seleccionarse una célula”, como se usa en la presente memoria, se entiende como la selección de una población de células, una subpoblación o una sola célula que posee una propiedad dada, tal como la presencia o ausencia de marcadores fluorescentes, o la presencia o ausencia de marcadores de selección positiva o negativa, y la obtención de un clon de cada célula seleccionada. Las células pueden seleccionarse en condiciones que eliminarán completamente o casi completamente cualquier célula que no tenga la propiedad deseada de las células a seleccionar. Por ejemplo, cultivando células en medios selectivos, sólo sobrevivirán las células que posean cierta propiedad deseada. Las células supervivientes pueden clonarse usando
- 40
- 45
- 50
- 55
- 60

protocolos de biología molecular y celular convencionales (véase Current Protocols in Cell Biology, Edición en CD-ROM, ed. por Juan S. Bonifacino, Jennifer Lippincott-Schwartz, Joe B. Hardford y Kenneth M. Yamada, John Wiley & Sons, 1999). Como alternativa, las células que poseen una propiedad deseada pueden seleccionarse de una población basándose en la observación de cierto fenotipo discernible, tal como la presencia o ausencia de marcadores fluorescentes. Las células seleccionadas pueden clonarse después usando protocolos de biología molecular y celular convencionales (véase Current Protocols in Cell Biology, Edición en CD-ROM, ed. por Juan S. Bonifacino, Jennifer Lippincott-Schwartz, Joe B. Hardford y Kenneth M. Yamada, John Wiley & Sons, 1999).

[0117] El término “equivalente” se entiende que incluye polipéptidos o secuencias de nucleótidos que son funcionalmente equivalentes o poseen una actividad equivalente en comparación con un polipéptido o secuencia de nucleótidos dada. Las secuencias de nucleótidos equivalentes incluirán secuencias que difieran por una o más sustituciones, adiciones o deleciones de nucleótidos, tales como variantes alélicas; y, por lo tanto, incluirán secuencias que difieran de la secuencia de nucleótidos de un gen particular debido a la degeneración del código genético. Los polipéptidos equivalentes incluirán polipéptidos que difieran por una o más sustituciones, adiciones o deleciones de aminoácidos, dejando dichas sustituciones, adiciones o deleciones de aminoácidos la función y/o actividad del polipéptido sustancialmente sin alteraciones. Un polipéptido equivalente a un polipéptido dado podría ser, por ejemplo, el polipéptido que realiza la misma función en otra especie. Por ejemplo, la ubiquitina murina se considera en la presente memoria un equivalente de la ubiquitina humana.

[0118] La expresión “derivado de FK506”, como se usa en la presente memoria, se refiere a un homólogo estructural del FK506 nativo en su sentido más amplio. Se ha descrito que el FKBP, el compañero de unión normal de FK506, puede modificarse para unirse a un derivado de FK506 de tal modo que el bolsillo de unión mutado sólo pueda alojar el derivado de FK506, pero no el FK506 de tipo silvestre (Clackson et al., 1998, Proc. Natl. Acad. Sci. U.S.A. 95: 10437-42 y Yang et al., 2000, J. Med. Chem. 43: 1135-42). Debería entenderse que la expresión “derivado de FK506” abarca al menos esta clase de derivados de FK506 en el contexto de la unión al FKBP mutante complementario. Además, los derivados de FK506 también pueden ser aquellos compuestos estructuralmente similares pero no idénticos que tengan esencialmente la misma función que FK506.

[0119] El “resto indicador”, como se usa en la presente memoria, se refiere a una característica que puede detectarse por ciertos medios. Por ejemplo, se consigue un ensayo de rutina para su detección por transferencia de Western usando un anticuerpo específico para una característica de proteína. Como alternativa, el resto indicador o resto que contiene un resto indicador puede ser capaz de presentar una función detectable deseada. Particularmente, la función puede suprimirse o inhibirse antes de que ocurra cierto acontecimiento (tal como la escisión del resto indicador del dominio Cub en un sistema de ubiquitina dividida) y la supresión o inhibición puede suprimirse después de que ocurra dicho acontecimiento. Por ejemplo, sin limitación, un resto indicador de transcripción puede volverse no funcional cuando se une a un resto Cub que está unido a una membrana fuera del núcleo de una célula diana. Puede volverse funcional después de la escisión del resto indicador del resto Cub cuando puede translocarse libremente al núcleo para ejercer su función de activación/supresión de la transcripción, actividad que a su vez es detectable midiendo la actividad de un gen indicador funcionalmente unido.

[0120] Como se usan en la presente memoria, las expresiones “gen”, “gen recombinante” y “construcción génica” se refieren a un ácido nucleico que comprende una fase de lectura abierta que codifica un polipéptido, incluyendo secuencias tanto exónicas como (opcionalmente) intrónicas. El término “intrón” se refiere a una secuencia de ADN presente en un gen dado que no se traduce en una proteína y que generalmente se encuentra entre exones.

[0121] La expresión “alta afinidad”, como se usa en la presente memoria, se refiere a una fuerte afinidad de unión entre moléculas con una constante de disociación K_D no superior a un 1 μ M. En un caso preferido, la K_D es inferior a 100 nM, 10 nM, 1 nM, 100 pM o incluso 10 pM o menos. En una realización más preferida, las dos moléculas pueden unirse covalentemente (K_D es esencialmente 0).

[0122] Los términos “homología” o “identidad” o “similitud” se refieren a la similitud de secuencia entre dos péptidos o entre dos moléculas de ácido nucleico, siendo la identidad una comparación más estricta. La homología y la identidad pueden determinarse cada una por comparación de una posición en cada secuencia que puede alinearse con fines comparativos. Cuando una posición en la secuencia comparada está ocupada por la misma base o aminoácido, entonces las moléculas son idénticas en esa posición. Un grado de homología o similitud o identidad entre secuencias de ácido nucleico está en función del número de nucleótidos idénticos o coincidentes en posiciones compartidas por las secuencias de ácido nucleico. Un grado de identidad de secuencias de aminoácidos está en función del número de aminoácidos idénticos en posiciones compartidas por las secuencias de aminoácidos. Un grado de homología o similitud de secuencias de aminoácidos está en función del número de aminoácidos, es decir, estructuralmente relacionados, en posiciones compartidas por las secuencias de aminoácidos. Una secuencia “no relacionada” o “no homóloga” comparte una identidad inferior al 40%, aunque preferiblemente una identidad inferior al 25% con otra secuencia.

[0123] El término “interaccionar”, como se usa en la presente memoria, pretende incluir todas las interacciones (por ejemplo, interacciones bioquímicas, químicas o biofísicas) entre moléculas, tales como interacciones de proteína-proteína, proteína-ácido nucleico, ácido nucleico-ácido nucleico, proteína-molécula pequeña, ácido nucleico-molécula pequeña o molécula pequeña-molécula pequeña.

[0124] El término “aislado”, como se usa en la presente memoria con respecto a ácidos nucleicos, tales como ADN o ARN, se refiere a moléculas separadas de otros ADN o ARN respectivamente, que están presentes en la fuente natural de la macromolécula. Por ejemplo, un ácido nucleico aislado que codifica uno de los polipéptidos objeto incluye preferiblemente no más de 10 kilobases (kb) de secuencia de ácido nucleico que flanquean inmediatamente de forma natural el gen en el ADN genómico, más preferiblemente no más de 5 kb de dichas secuencias flanqueantes de origen natural, y más preferiblemente, menos de 1,5 kb de dicha secuencia flanqueante de origen natural. El término aislado, como se usa en la presente memoria, también se refiere a un ácido nucleico o péptido que está sustancialmente libre de material celular, material viral o medio de cultivo cuando se produce por técnicas de ADN recombinante, o precursores químicos u otros productos químicos cuando se sintetiza químicamente. Además, un “ácido nucleico aislado” pretende incluir fragmentos de ácido nucleico que de forma natural no aparecen como fragmentos y que no se encontrarían en el estado natural. El término “aislado” también se usa en la presente memoria para referirse a polipéptidos que están aislados de otras proteínas celulares y se entiende que incluyen polipéptidos tanto purificados como recombinantes.

[0125] El término “kit”, como se usa en la presente memoria, se refiere a una colección de al menos dos componentes que constituyen el kit. En conjunto, los componentes constituyen una unidad funcional para un fin dado. Los componentes miembros individuales pueden envasarse físicamente juntos o por separado. Por ejemplo, un kit que comprende una instrucción para usar el kit puede o no incluir físicamente la instrucción con otros componentes miembros individuales. En su lugar, la instrucción puede suministrarse como un componente miembro separado, en forma de papel o en forma electrónica que puede suministrarse en un dispositivo de memoria legible por ordenador o descargarse de un sitio web de internet, o como una presentación grabada.

[0126] La “instrucción o instrucciones”, como se usa en la presente memoria, se refieren a documentos que describen los materiales o metodologías relevantes que pertenecen a un kit. Estos materiales pueden incluir cualquier combinación de lo siguiente: información antecedente, lista de componentes y su información de disponibilidad (información de compra, etc.), protocolos resumidos o detallados para usar el kit, resolución de problemas, referencias, soporte técnico y cualquier otro documento relacionado. Las instrucciones pueden suministrarse con el kit o con un componente miembro separado, en forma de papel o en forma electrónica que puede suministrarse en un dispositivo de memoria legible por ordenador o descargarse de un sitio web de internet, o como una presentación grabada. Las instrucciones pueden comprender uno o múltiples documentos, y pretenden incluir actualizaciones futuras.

[0127] El término “biblioteca”, como se usa en la presente memoria, generalmente se refiere a una multiplicidad de componentes miembros que constituyen la biblioteca cuyos componentes miembros difieren individualmente con respecto a al menos una propiedad, por ejemplo, una biblioteca de compuestos químicos. Particularmente, como será evidente para el experto en la materia, el término “biblioteca” se refiere a una pluralidad de ácidos nucleicos/polinucleótidos, preferiblemente en forma de vectores que comprenden elementos funcionales (promotor, sitios de unión a factor de transcripción, potenciador, etc.) necesarios para la expresión de polipéptidos, *in vitro* o *in vivo*, que están funcionalmente unidos a secuencias codificantes de polipéptidos. El vector puede ser un plásmido o un vector basado en virus adecuado para la expresión en procariotas o eucariotas, o en ambos, preferiblemente para la expresión en células de mamífero. Debería haber al menos uno, preferiblemente múltiples pares de sitios de clonación para la inserción de secuencias codificantes en la biblioteca, y para la recuperación o clonación posterior de esas secuencias codificantes. Los sitios de clonación pueden ser sitios de reconocimiento de endonucleasas de restricción u otras secuencias de reconocimiento basadas en recombinación, tales como secuencias loxP para la recombinasa Cre, o el sistema Gateway (Life Technologies, Inc), que se describe en la Patente de Estados Unidos Nº 5.888.732. Las secuencias codificantes de polipéptidos pueden ser ADNc, fragmentos de ADN genómico o polinucleótidos aleatorios/semialeatorios. Los métodos para la construcción de genotecas de ADNc o ADN genómico son bien conocidos en la técnica, y pueden encontrarse en varios manuales de biología molecular de laboratorio usados comúnmente (véase a continuación).

[0128] El término “modulación”, como se usa en la presente memoria, se refiere tanto a la regulación positiva (es decir, activación o estimulación, por ejemplo, por agonismo o potenciación) como a la regulación negativa (es decir, inhibición o supresión, por ejemplo, por antagonismo, disminución o inhibición) de una actividad.

[0129] Los términos “mutación” o “mutado” en referencia a un gen o ácido nucleico se refieren a una forma alélica o modificada de un gen o ácido nucleico, que presenta una secuencia de nucleótidos diferente y/o una propiedad física o química alterada en comparación con el gen o ácido nucleico de tipo silvestre. Generalmente, la mutación podría alterar la secuencia reguladora de un gen sin afectar a la secuencia polipeptídica codificada por el gen de tipo silvestre. Pero más comúnmente, un gen o ácido nucleico mutado perderá completamente la capacidad para codificar un polipéptido (mutación nula) o codificar un polipéptido con una propiedad alterada, incluyendo un polipéptido con una actividad biológica reducida o aumentada, un polipéptido con una nueva actividad biológica o un polipéptido que interfiera con la función del polipéptido de tipo silvestre correspondiente. Como alternativa, una mutación puede aprovechar la degeneración del código genético sustituyendo un codón de triplete por un codón de triplete diferente que, no obstante, codifique el mismo aminoácido que el codón de triplete de tipo silvestre. Dicha sustitución puede conducir, por ejemplo, a una estabilidad aumentada del gen o ácido nucleico en determinadas condiciones. Además, una mutación puede comprender un cambio de nucleótido en una sola posición del gen o ácido nucleico, o en varias posiciones, o deleciones o adiciones de nucleótidos en una o varias posiciones.

[0130] La expresión “mutante de asociación reducida”, como se usa en la presente memoria se refiere a un polipéptido mutante que muestra una afinidad reducida por su compañero de unión normal. Por ejemplo, un mutante de asociación

reducida del extremo N-terminal de ubiquitina (Nux) es un polipéptido que muestra una afinidad reducida por su compañero de unión normal —la mitad C-terminal de ubiquitina (Cub), hasta el punto en que mostrará una asociación reducida o no se asociará con un Cub de tipo silvestre y formará una “ubiquitina casi silvestre” sin la afinidad de unión complementada entre dos polipéptidos fusionados a Nux y Cub, respectivamente. En una realización preferida de la invención, dichas mutaciones en Nux son ciertas mutaciones con sentido erróneo introducidas en el 3º o 13º resto aminoacídico de la ubiquitina de tipo silvestre. Diferentes mutaciones con sentido erróneo en estas posiciones pueden afectar de forma diferencial a la afinidad/asociación entre Nux y Cub, proporcionado de este modo una sensibilidad diferente del ensayo que se describe en la presente invención. Estas mutaciones puntuales con sentido erróneo pueden introducirse de forma rutinaria en genes clonados usando protocolos de biología molecular convencionales, tales como mutagénesis dirigida usando PCR.

[0131] Como se usa en la presente memoria, la expresión “ácido nucleico”, en su sentido más amplio, se refiere a polinucleótidos tales como ácido desoxirribonucleico (ADN) y, cuando es apropiado, a ácido ribonucleico (ARN). El término también debería entenderse que incluye, como equivalentes, análogos de ARN o ADN generados a partir de análogos de nucleótidos generados a partir de análogos de nucleótidos y, como es aplicable a la realización que se está describiendo, polinucleótidos mono- (con sentido o antisentido) o bicatenarios.

[0132] En concreto, la expresión “ácido nucleico o ácidos nucleicos” puede referirse a polinucleótidos que contienen la información necesaria para la transcripción y/o traducción de polipéptidos codificados por los polinucleótidos. Estos incluyen, pero sin limitación, plásmidos que comprenden señales de transcripción (por ejemplo, sitios de unión a factores de transcripción, promotores y/o potenciadores) unidas funcionalmente a secuencias codificantes cadena abajo de polipéptidos, fragmentos de ADN genómico que comprenden señales de transcripción (por ejemplo, sitios de unión a factores de transcripción, promotores y/o potenciadores) unidas funcionalmente a secuencias codificantes cadena abajo para polipéptidos, fragmentos de ADNc (lineal o circular) que comprenden señales de transcripción (por ejemplo, sitios de unión a factores de transcripción, promotores y/o potenciadores) unidas funcionalmente a secuencias codificantes cadena debajo de polipéptidos, o moléculas de ARN que comprenden elementos funcionales para la traducción *in vitro* o *in vivo*, o ambas, que están unidas funcionalmente a secuencias que codifican polipéptidos. Estos polinucleótidos también debería entenderse que incluyen, como equivalentes, análogos de ARN o ADN generados a partir de análogos de nucleótidos y, como es aplicable a la realización que se está describiendo, polinucleótidos mono- (con sentido o antisentido) y bicatenarios. Estos polinucleótidos pueden estar en forma aislada, por ejemplo, un vector aislado, o incluidos en el episoma o el genoma de una célula.

[0133] Como se usa en la presente memoria, el término “promotor” se refiere a una secuencia de ADN que regula la expresión de una secuencia de ADN seleccionada unida operativamente al promotor, y que ejerce la expresión de la secuencia de ADN seleccionada en las células. El término incluye promotores “específicos de tejido”, es decir, promotores que ejercen la expresión de la secuencia de ADN seleccionada sólo en células específicas (por ejemplo, células de un tejido específico). El término también abarca los denominados promotores “débilmente específicos”, que regulan la expresión de un ADN seleccionado principalmente en un tejido, pero causan expresión también en otros tejidos. El término también incluye promotores inespecíficos de tejido y promotores que expresan constitutivamente o que son inducibles (es decir, pueden controlarse los niveles de expresión).

[0134] Los términos “proteína”, “polipéptido” “péptido” se usan indistintamente en la presente memoria cuando se hace referencia a un producto génico natural o recombinante o fragmento del mismo que no es un ácido nucleico.

[0135] La expresión “proteína recombinante” se refiere a un polipéptido que se produce por técnicas de ADN recombinante, en las que, generalmente, se inserta ADN que codifica un polipéptido en un vector de expresión adecuado que a su vez se usa para transformar una célula hospedadora para producir el polipéptido codificado por dicho ADN. Este polipéptido puede ser uno que se exprese de forma natural por la célula hospedadora, o puede ser heterólogo respecto a la célula hospedadora, o la célula hospedadora puede haberse modificado por ingeniería genética para perder la capacidad de expresar el polipéptido que de otro modo se expresa en formas de tipo silvestre de la célula hospedadora. El polipéptido también puede ser un polipéptido de fusión. Además, la expresión “derivado de” con respecto a un gen recombinante pretende incluir dentro del significado de “proteína recombinante” a aquellas proteínas que tienen una secuencia de aminoácidos de un polipéptido nativo, o una secuencia de aminoácidos similar a la misma que se genera por mutaciones, incluyendo sustituciones, deleciones y truncamiento de una forma de origen natural del polipéptido.

[0136] La “molécula pequeña”, como se usa en la presente memoria, pretende referirse a una composición o compuesto que tiene un peso molecular de menos de aproximadamente 5 kD y, más preferiblemente, de menos de aproximadamente 4 kD. Las moléculas pequeñas pueden ser ácidos nucleicos, péptidos, polipéptidos, peptidomiméticos, carbohidratos, lípidos u otras moléculas orgánicas (que contienen carbono) o inorgánicas. Muchas compañías farmacéuticas tienen amplias bibliotecas de mezclas químicas y/o biológicas, con frecuencia extractos fúngicos, bacterianos o de algas, que pueden explorarse potencialmente con métodos de la invención relacionando dichos productos químicos con un ligando común que se usa en la presente invención.

[0137] La “transcripción” es un término genérico usado a lo largo de toda la memoria descriptiva para referirse a un proceso de síntesis de moléculas de ARN de acuerdo con sus secuencias de molde de ADN correspondientes, que puede incluir señales de inicio, potenciadores y promotores que inducen o controlan la transcripción de secuencias codificantes de proteína con las que están unidas operativamente. La expresión “represor de la transcripción”, como se

- usa en la presente memoria, se refiere a cualquiera de diversos polipéptidos de origen procariota o eucariota, o que son construcciones químicas artificiales sintéticas capaces de represión en solitario o junto con otros polipéptidos, y que reprimen la transcripción de una forma activa o pasiva. También se entenderá que la transcripción de un gen recombinante puede estar bajo el control de secuencias reguladoras de la transcripción que son iguales o diferentes de aquellas secuencias que controlan la transcripción de las formas de origen natural del gen recombinante, o sus componentes.
- [0138]** La “traducción”, como se usa en la presente memoria, es un término genérico usado para describir la síntesis de una proteína o polipéptido en un molde, tal como ARN mensajero (ARNm). Es la generación de una secuencia proteica/polipeptídica por traducción del código genético de una molécula de ARNm asociada con un ribosoma. El proceso completo puede realizarse *in vivo* en el interior de una célula usando la maquinaria de traducción de proteínas de la célula, o realizarse *in vitro* usando sistemas sin células, tales como lisados de reticulocitos o cualquier otro equivalente. El molde de ARN para traducción puede proporcionarse por separado o directamente como ARN, o indirectamente como el producto de transcripción a partir de un molde de ADN proporcionado, tal como un plásmido.
- [0139]** La expresión “proporcionar por traducción” se refiere a proporcionar un polipéptido/proteína por medio de traducción. Como se ha definido anteriormente, la traducción es un proceso que puede realizarse *in vivo* en el interior de una célula usando la maquinaria de traducción de proteínas de la célula, o realizarse *in vitro* usando sistemas sin células, tales como lisados de reticulocitos o cualquier otro equivalente. El molde de ARN para traducción puede proporcionarse por separado directamente como ARN o indirectamente como el producto de transcripción a partir de un molde de ADN proporcionado, tal como un plásmido. El ADN de molde puede introducirse en una célula hospedadora/diana mediante una diversidad de procedimientos de biología molecular convencionales, tales como transformación, transfección, conjugación o fusión celular, o puede proporcionarse a una reacción de traducción *in vitro* directamente.
- [0140]** Los términos “transfección” y “transformación” se usan indistintamente en la presente memoria para denominar la introducción de un ácido nucleico, por ejemplo, sin limitación, por medio de un vector de expresión, en una célula destinataria.
- [0141]** El término “tratar”, como se usa en la presente memoria, pretende incluir la curación, así como la mejoría de al menos un síntoma de la afección o enfermedad.
- [0142]** El término “vector” se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico con el que se ha unido. Un tipo de vector preferido es un episoma, es decir, un ácido nucleico capaz de replicación extra-cromosómica. Los vectores preferidos son aquellos capaces de una replicación y/o expresión autónoma de los ácidos nucleicos con los que están unidos. Los vectores capaces de dirigir la expresión de genes con los que están unidos operativamente se denominan en la presente memoria “vectores de expresión”. En general, los vectores de expresión de utilidad en técnicas de ADN recombinante están con frecuencia en forma de “plásmidos”, que se refieren en general a bucles de ADN bicatenarios circulares que, en su forma de vector, no están unidos al cromosoma. En la presente memoria descriptiva, los términos “plásmido” y “vector” se usan indistintamente, ya que el plásmido es la forma de vector usada más comúnmente. Sin embargo, la invención pretende incluir otras formas tales de vectores de expresión que sirvan a funciones equivalentes y que se conozcan en la técnica posteriormente a la misma.
- [0143]** Las “ubiquitinas” son una clase de proteínas que se encuentran en todas las células eucariotas. El polipéptido de ubiquitina se caracteriza por un resto de glicina carboxi-terminal que se activa por ATP a un intermedio de tiol-éster de alta energía en una reacción catalizada por una enzima de activación de ubiquitina (E1). La ubiquitina activada se transfiere a un polipéptido de sustrato a través de un enlace isopeptídico entre el extremo carboxi-terminal de la ubiquitina y el grupo épsilon-amino de (a) un resto o restos de lisina en el sustrato proteico. Esta transferencia requiere la acción de enzimas de conjugación con ubiquitina, tales como actividades de E2 y, en algunos casos, E3. El sustrato modificado con ubiquitina se altera de este modo en su función biológica y, en algunos casos, se convierte en un sustrato para componentes de la maquinaria proteolítica dependiente de ubiquitina que incluye tanto enzimas UBP como proteínas proteolíticas que son subunidades del proteosoma. Como se usa en la presente memoria, el término “ubiquitina” incluye dentro de su alcance todos los homólogos de ubiquitina eucariotas conocidos, así como no identificados, de origen de vertebrados o invertebrados, que pueden clasificarse como equivalentes de ubiquitina humana. Los ejemplos de polipéptidos de ubiquitina mencionados en la presente memoria incluyen el polipéptido de ubiquitina humana que está codificado por la secuencia de ácido nucleico codificante de ubiquitina humana (Números de Acceso de GenBank: U49869, X04803). Se entiende que las secuencias de nucleótidos codificantes de polipéptidos de ubiquitina equivalentes incluyen aquellas secuencias que difieren por una o más sustituciones, adiciones o deleciones de nucleótidos, tales como variantes alélicas; así como secuencias que difieren de la secuencia de nucleótidos de la secuencia codificante que codifica la ubiquitina humana debido a la degeneración del código genético. Otro ejemplo de un polipéptido de ubiquitina mencionado en la presente memoria es ubiquitina murina que está codificada por la secuencia de ácido nucleico codificante de ubiquitina murina (Número de Acceso de GenBank: X51730). Será fácilmente evidente para el experto en la materia cómo modificar los métodos y reactivos proporcionados por la presente invención para el uso de polipéptidos de ubiquitina distintos de ubiquitina humana.
- [0144]** La expresión “proteína tipo ubiquitina”, como se usa en la presente memoria, se refiere a un grupo de proteínas de origen natural que no pueden describirse de otro modo como equivalentes de ubiquitina, pero que no obstante muestran una fuerte homología de aminoácidos con ubiquitina humana. Como se usa en la presente memoria, este

término incluye los polipéptidos NEDD8, UBL1, NPVAC y NPVOC. Estas “proteínas tipo ubiquitina” tienen una identidad de secuencia de al menos más del 40% con el polipéptido de ubiquitina humana y contienen un par de restos de glicina carboxi-terminales que funcionan en la activación y transferencia de ubiquitina a sustratos diana como se ha descrito anteriormente.

5 **[0145]** Como se usa en la presente memoria, la expresión “proteína relacionada con ubiquitina”, como se usa en la presente memoria, se refiere a un grupo de proteínas de origen natural que no pueden describirse de otro modo como equivalentes de ubiquitina, pero que no obstante muestran cierto grado relativamente bajo (identidad <40%) de homología de aminoácidos con la ubiquitina humana. Estas proteínas “relacionadas con ubiquitina” incluyen la proteína de reactividad cruzada con ubiquitina humana (UCRP, identidad del 36% con huUb, N° de Acceso P05161), FUBI
10 (identidad del 36% con huUb, N° de Acceso GenBank AA449261) y Sentrin/Sumo/Pic1 (identidad del 20% con huUb, N° de Acceso GenBank U83117). La expresión “proteína relacionada con ubiquitina”, como se usa en la presente memoria, se refiere además a polipéptidos que poseen un par carboxi-terminal de restos de glicina y que funcionan como etiquetas proteicas por medio de la activación del resto de glicina carboxi-terminal y la posterior transferencia a un sustrato proteico.

15 **[0146]** La expresión “proteína homóloga a ubiquitina”, como se usa en la presente memoria, se refiere a un grupo de proteínas de origen natural que no pueden describirse de otro modo como equivalentes de ubiquitina o proteínas tipo ubiquitina o relacionadas con ubiquitina, que parecen ser funcionalmente distintas de la ubiquitina en su capacidad para actuar como etiquetas proteicas, pero que no obstante muestran cierto grado de homología con la ubiquitina humana (identidad del 34-41%). Estas “proteínas homólogas a ubiquitina” incluyen RAD23A (identidad del 36% con huUb, N° de
20 Acceso SWISS-PROT. P54725), RAD23B (identidad del 34% con huUb, N° de Acceso SWISS-PROT. P54727), DSK2 (identidad del 41% con huUb, N° de Acceso GenBank L40587) y GDX (identidad del 41% con huUb, N° de Acceso GenBank J03589). La expresión “proteína homóloga a ubiquitina”, como se usa en la presente memoria, pretende además significar una clase de polipéptidos homólogos a ubiquitina cuya similitud con la ubiquitina no incluye restos de glicina en las posiciones de restos penúltima y carboxi-terminal. Dichas proteínas parecen ser funcionalmente distintas
25 de ubiquitina, así como de polipéptidos tipo ubiquitina y relacionados con ubiquitina, en el sentido de que, de acuerdo con su ausencia de una glicina carboxi-terminal conservada para usarse en una reacción de activación, no se ha demostrado que sirvan como etiquetas para otras proteínas por enlace covalente.

[0147] La expresión “maquinaria de conjugación con ubiquitina”, como se usa en la presente memoria, se refiere a un grupo de proteínas que funcionan en la activación dependiente de ATP y la transferencia de ubiquitina a proteínas de sustrato. El término incluye por lo tanto: enzimas E1, que transforman la glicina carboxi-terminal de ubiquitina en un intermedio de tiol de alta energía mediante una reacción dependiente de ATP; enzimas E2 (los genes UBC), que transforman el conjugado activado por E1-S-ubiquitina en un intermedio de E2-S-ubiquitina que actúa como donador de ubiquitina a un sustrato, otro resto de ubiquitina (en una reacción de poliubiquitinación) o una E3; y las enzimas E3 (o
30 ligasas de ubiquitina), que facilitan la transferencia de una molécula de ubiquitina activada de un E2 a una molécula de sustrato o a otro resto de ubiquitina como parte de una cadena de poliubiquitina. La expresión “maquinaria de conjugación con ubiquitina”, como se usa en la presente memoria, pretende además incluir todos los miembros conocidos de estos grupos, así como los miembros que todavía están por descubrirse o caracterizarse, pero que están suficientemente relacionados por homología con enzimas de conjugación con ubiquitina conocidas para permitir que un individuo experto en la materia los identifique fácilmente como miembros de este grupo. La expresión, como se usa en
40 la presente memoria, pretende incluir nuevas enzimas activadoras de ubiquitina que todavía están por descubrirse, así como aquellas que funcionan en la activación y conjugación de polipéptidos de tipo ubiquitina o relacionados con ubiquitina con sus sustratos y con cadenas proteicas de tipo poliubiquitina o relacionadas con poliubiquitina.

[0148] La expresión “maquinaria proteolítica dependiente de ubiquitina”, como se usa en la presente memoria, se refiere a enzimas proteolíticas que funcionan en las rutas bioquímicas de la ubiquitina, proteínas tipo ubiquitina y relacionadas con ubiquitina. Dichas enzimas proteolíticas incluyen las hidrolasas C-terminales de ubiquitina, que hidrolizan el enlace entre el resto de glicina carboxi-terminal de la ubiquitina y diversos aductos; UBP, que hidrolizan el enlace de glicina 76-lisina 48 entre restos de ubiquitina entrecruzados en conjugados de poliubiquitina, así como otras enzimas que funcionan en la eliminación de conjugados de ubiquitina de sustratos ubiquitinados (generalmente denominadas “enzimas desubiquitinizantes”). Las actividades de proteasa mencionadas anteriormente funcionan en la eliminación de
50 unidades de ubiquitina de un sustrato ubiquitinado después de o durante la degradación dependiente de ubiquitina, así como en ciertas funciones de corrección de errores en las que se eliminan polipéptidos de ubiquitina libres de proteínas incorrectamente ubiquitinadas. La expresión “maquinaria proteolítica dependiente de ubiquitina”, como se usa en la presente memoria, también pretende incluir las subunidades proteolíticas del proteosoma (incluyendo las subunidades del proteosoma humano C2, C3, C5, C8 y C9). La expresión “maquinaria proteolítica dependiente de ubiquitina”, como se usa en la presente memoria, incluye además dos clases de proteasas: las enzimas desubiquitinizantes y las subunidades del proteosoma. No se sabe que las funciones de proteasa de las subunidades del proteosoma ocurran fuera del contexto del proteosoma ensamblado, sin embargo, no se ha excluido un funcionamiento independiente de
55 estos polipéptidos.

[0149] El término “quinasa”, como se usa en la presente memoria, se refiere a una enzima que transfiere un grupo fosfato de un nucleósido trifosfato a otra molécula. Preferiblemente, la quinasa se selecciona de la lista siguiente: AMP-PK (proteína quinasa activada por AMP, acetil-CoA carboxilasa quinasa-3, HMG-CoA reductasa quinasa, lipasa quinasa sensible a hormonas), ACK2 (acetil-CoA carboxilasa quinasa-2), AFK (actina-fragmina quinasa), APL-A1 (PK 1
60

dependiente de AMPc de *Aplysia Californica*), APL-A2 (PK 2 dependiente de AMPc de *Aplysia Californica*), CAK (quinasa activadora de Cdk), CAMII (= CaM-II), beta-ARK1 (quinasa 1 de receptor beta-adrenérgico = GRK2), beta-ARK2 (quinasa 2 de receptor beta-adrenérgico = GRK3), c-Abl (Abl celular), c-Raf (Raf celular), c-Src (Src celular), Cdk (quinasa dependiente de ciclina), cdc2 (proteína quinasa del ciclo de división celular), CK (caseína quinasa), CK-1 o CKI (caseína quinasa I), CK-II o CKII (caseína quinasa II), CTD quinasa ((ARN polimerasa II) quinasa de dominio carboxi-terminal), CAM-I (proteína quinasa I dependiente de calmodulina), CaM-II (proteína quinasa II dependiente de calmodulina, multiproteína quinasa dependiente de calmodulina, CaM-MPK), CaM-III (proteína quinasa III dependiente de calmodulina, EF-2 quinasa), DNA-PK (proteína quinasa dependiente de ADN), ADN-bc quinasa (proteína quinasa activada por ADN bicatenario), ARN-bc quinasa (proteína quinasa activada por ARN bicatenario, p68 quinasa), EGF-R o EGFR (receptor del factor de crecimiento epidérmico), ERK (quinasa regulada por señales extracelulares = MAPK), ERT PK (quinasa regulada por factor de crecimiento), FAK (quinasa de adhesión focal), GRK1 (quinasa receptora 1 acoplada a proteína G = RK), GRK2 (quinasa receptora 2 acoplada a proteína G = beta-ARK1), GRK3 (quinasa receptora 3 acoplada a proteína G = beta-ARK2), GRK4 (quinasa receptora 4 acoplada a proteína G), GRK5 (quinasa receptora 5 acoplada a proteína G), GRK6 (quinasa receptora 5 acoplada a proteína G), GSK1 (glucógeno sintasa quinasa 1 = PKA), GSK2 (glucógeno sintasa quinasa 2 = PHK), GSK3 (glucógeno sintasa quinasa 3), GSK4 (glucógeno sintasa quinasa 4), GSK5 (glucógeno sintasa quinasa 5 = CKII), H1-HK (quinasa de histona H1 asociada al crecimiento (MPF), proteína quinasa cdc2+/CDC28), H4-PK (proteína quinasa activada por proteasa específica de histona H4), H4-PK-I (quinasa I de histona H4), H4-PK-II (quinasa II de histona H4), HCR (represor controlado por hemo, eIF-2-alfa quinasa regulada por hemo), HKII (histona quinasa II), INS-R o INSR (receptor de insulina), Jak1 (proteína tirosina quinasa Janus 1), Jak2 (proteína tirosina quinasa Janus 2), LCK/FYN (PROTEÍNA TIROSINA QUINASA ESPECÍFICA DE LINFOCITOS P56LCK), MAPK (proteína quinasa activada por mitógeno (MAP quinasa) = ERK), MAPKAPK-1 (proteína quinasa 1 activada por MAP quinasa = S6K-II), MAPKAPK-2 (proteína quinasa 2 activada por MAP quinasa), MEK (MAP, Erk quinasa, MAP quinasa quinasa), MFPK (proteína quinasa multifuncional), MHCK (quinasa de cadena pesada de miosina), MLCK (quinasa de cadena ligera de miosina), p135tyk2 (tirosina proteína quinasa tyk2 de 135 kD), p34cdc2 (proteína quinasa del ciclo de división celular de 34 kD), p42cdc2 (proteína quinasa del ciclo de división celular de 42 kD), p42mapk (isoforma de MAP quinasa de 42 kD), p44mpk (proteína quinasa básica de mielina activada por meiosis de 44 kD = ERK1), p60-src (tirosina proteína quinasa src), p74raf 1 (isoforma de proteína quinasa Raf de 74 kDa), PDGF-R o PDGFR (receptor del factor de crecimiento derivado de plaquetas), PHK (fosforilasa quinasa), PI-3 quinasa (fosfatidilinositol 3' quinasa), PKA (proteína quinasa dependiente de AMPc, proteína quinasa A), PKC (proteína quinasa C), PKG (proteína quinasa dependiente de GMPc), PRK1 (quinasa relacionada con PKC activada por lípidos), Raf (proteína quinasa Raf), RK (rodopsina quinasa = GRK1), RS quinasa (proteína quinasa unida a envuelta nuclear), S6K (S6 quinasa), S6K-II (S6-quinasa 2 = MAPKAPK-1), v-Src (Src viral).

[0150] La expresión “unirse a o inhibir una quinasa” se refiere a la capacidad de determinados compuestos para unirse a quinastas con gran afinidad y a la propiedad adicional de ciertos compuestos para disminuir la actividad de una quinasa. El “o” en la misma no pretende ser excluyente, es decir, un compuesto puede tanto unirse a quinasa como inhibirla, o puede sólo unirse o puede sólo inhibir dicha quinasa, según sea el caso.

3. Sistema transcripcional y otros sistemas indicadores

[0151] De acuerdo con la invención, se usa un sistema indicador para detectar la proximidad de dos polipéptidos P1 y P2 (como se han definido anteriormente) cuando está presente un compuesto molecular pequeño, de modo que el compuesto molecular pequeño o uno de los polipéptidos pueda identificarse y caracterizarse adicionalmente.

[0152] Las secciones siguientes describirán una diversidad de sistemas indicadores que pueden usarse en la invención. Será fácilmente evidente para el experto que la invención más inmediata también puede usarse junto con otros sistemas indicadores, incluso los que se desarrollen en el futuro.

3.1 Sistemas indicadores de ubiquitina dividida

[0153] En parte, la invención se basa en el descubrimiento de que pueden detectarse incluso interacciones transitorias usando un nuevo método de selección de la asociación polipeptídica basado en ubiquitina dividida. El método de ubiquitina dividida se ha usado para demostrar, por ejemplo, la asociación de Sec63p con diversas otras proteínas de membrana de levadura que transitan por el retículo endoplásmico (RE) y el aparato de Golgi y se dirigen a la membrana plasmática.

[0154] Se entiende que la invención incluye modificaciones y extensiones de los ejemplos descritos anteriormente, como se describe a continuación.

[0155] La invención proporciona una proteína de fusión que comprende el polipéptido P1-Cub-Z-RM, en el que P1 es un primer polipéptido, Cub es un subdominio C-terminal de ubiquitina, Z es un resto aminoacídico y RM es un resto indicador, en el que la proteína de fusión puede escindirse por una proteasa específica de ubiquitina en presencia de una forma mutante o de tipo silvestre de interacción del subdominio Nub de ubiquitina fusionado a un segundo polipéptido P2 (fusión P2-Nux), y da como resultado la liberación del resto indicador. Dependiendo de la identidad del resto Z, el RM liberado puede ser estable si Z es Met e inestable si Z es un aminoácido amino-terminal distinto de metionina, por lo tanto la actividad de dicho resto indicador puede cambiarse antes y/o después de dicha liberación. La afinidad entre Cub y Nub puede modularse introduciendo mutaciones puntuales (por ejemplo, en los restos 3 ó 13 o en ambas posiciones) en Nub, de modo que Cub y Nub (o sus formas mutantes derivadas “Nux”) no pueden interactuar

entre sí sin la presencia de otras fuerzas estabilizantes, tales como la proporcionada por la interacción entre P1 y P2, en este caso indirectamente, a través de un ligando compuesto. Debería entenderse que debido a la naturaleza simétrica del sistema, la designación de P1/P2 y R1/R2 es arbitraria. El resto indicador de estas proteínas de fusión puede ser una diversidad de proteínas, incluyendo, pero sin limitación: un marcador de selección negativa, un marcador de selección positiva, un marcador metabólico, un factor de transcripción y un marcador fluorescente. En aplicaciones preferidas, el indicador es un marcador de selección que es capaz de una selección tanto positiva como negativa, tal como URA3, HygTk, Tkneo, TkBSD, PACTk, HygCoda, Codaneo, CodaBSD y PACCoda. Otros indicadores incluyen LYS2, HIS3 y GPT de mamífero. El resto indicador también puede ser un marcador fluorescente, un factor de transcripción, por ejemplo, PLV (Stagljar et al, PNAS, 1998, 95: 5187-92) o DHFR.

[0156] La invención usa bibliotecas de péptidos expresados como proteínas de fusión. Dichas bibliotecas de péptidos pueden ser bibliotecas de péptidos sintéticas, naturales, aleatorias, parcialmente aleatorias, restringidas, no restringidas y combinatorias. En ciertos casos, las bibliotecas de péptidos se proporcionan por expresión de la construcción o construcciones de ácido nucleico que codifican los polipéptidos. Las bibliotecas de ADN pueden ser una construcción o construcciones de ácido nucleico oligonucleotídico o genómico, sintético, parcialmente aleatorio, aleatorio o ADNc que codifique polipéptidos.

[0157] La invención proporciona además un método de detección de la unión de un compuesto químico a una proteína que comprende: proporcionar una primera proteína como una primera fusión de polipéptido que comprende el polipéptido de estructura P1-Cub-Z-RM, en el que P1 es un primer polipéptido, Cub es un subdominio C-terminal de ubiquitina, Z es un resto aminoacídico y RM es un resto indicador; proporcionar una segunda proteína de fusión como una segunda fusión de polipéptido que comprende la estructura P2-Nux, en la que P2 es un segundo polipéptido y Nux es una forma mutante o de tipo silvestre de un subdominio amino-terminal de ubiquitina, proporcionar un compuesto químico de fórmula general R1-Y-R2, en la que R1 es un ligando conocido para P1, R2 es un ligando potencial para P2 e Y es una secuencia enlazadora; permitir que el compuesto químico se ponga en estrecha proximidad con el primer polipéptido de fusión y el segundo polipéptido de fusión en condiciones en las que si R2 interacciona con P2, y la escisión de la primera proteína de fusión da como resultado la liberación del resto indicador que tiene el resto aminoacídico amino-terminal Z; proporcionar condiciones que permitan la detección de la actividad del resto indicador, en las que la presencia o ausencia de una señal detectable del resto indicador indica que el compuesto químico R2 se une a P2. Debería entenderse que, debido a la naturaleza simétrica del sistema, la designación de P1/P2 y R1/R2 es arbitraria y que P1 o P2 puede fusionarse a Cub-Z-RM. De forma similar, en la proteína de fusión P1-Nux, debería entenderse que, a menos que se especifique específicamente, P1-Nux se refiere a cualquiera de las dos configuraciones posibles de la proteína de fusión, en concreto P1-Nux (fusión N-terminal) o Nux-P1 (fusión C-terminal). Además, se entiende que P1-Cub-Z-RM incluye todas las configuraciones posibles de la proteína de fusión siempre que sea un orden en el que Cub-Z esté más próximo al extremo N-terminal de la proteína de fusión que RM (por ejemplo, P1-Cub-Z-RM, Cub-Z-P1-RM y Cub-Z-RM-P1 son todas configuraciones posibles).

[0158] En una realización preferida, se sabe que P1 y R1 interaccionan entre sí, mientras que el ligando que se une a una proteína P2 conocida o la proteína P2 que se une a un ligando R2 conocido pueden identificarse y caracterizarse adicionalmente.

[0159] Este método de la invención puede realizarse en un formato *in vitro* o *in vivo*. Los formatos *in vivo* pueden utilizar una célula hospedadora tal como una célula eucariota. Las células eucariotas adecuadas incluyen células de mamífero, incluyendo células humanas, de ratón, de rata y de hámster; células de vertebrados, incluyendo células de pez cebra; células de invertebrados, incluyendo células de nematodos y *Drosophila*; y células fúngicas, incluyendo células de *S. pombe* y *S. cerevisiae*. En realizaciones *in vivo* preferidas del método de la invención, el resto indicador es un marcador de selección positiva. El indicador también puede ser un marcador de selección negativa. El marcador puede ser un marcador metabólico, un factor de transcripción, un marcador de selección tanto positiva como negativa, un marcador fluorescente, un factor de transcripción o DHFR. El método proporciona el uso de diversos restos aminoacídicos que se van a modificar por ingeniería genética en el presunto extremo amino-terminal del indicador o proteína marcadora de selección. En una realización, este aminoácido es arginina, sin embargo, también puede ser otro aminoácido distinto de metionina —por ejemplo, lisina o histidina. En otras realizaciones, Z puede ser metionina u otros aminoácidos estables en un entorno dado (véase a continuación).

[0160] El método de la invención usa un primer y/o segundo polipéptidos, P1 y/o P2, que pueden suministrarse como bibliotecas de péptidos sintéticas, naturales, aleatorias, parcialmente aleatorias, restringidas, no restringidas y combinatorias. Estas bibliotecas pueden proporcionarse por expresión de una construcción o construcciones de ácido nucleico que codifican dicho primer y/o segundo polipéptidos. El método de la invención también usa una proteína de fusión que comprende P2 y Nux, en la que Nux se fusiona al extremo N-terminal del segundo polipéptido P2 o al extremo C-terminal del segundo polipéptido P2.

[0161] El método de la invención proporciona el compuesto químico R1-Y-R2 que puede suministrarse como bibliotecas de compuestos químicos sintéticos o naturales o de otro tipo.

3. 1. 1 Marcadores de selección

[0162] El principio establecido de la tecnología detectora de proteína ubiquitina dividida actual emplea dos vectores lanzadera de levadura/*E. coli* que codifican las proteínas de fusión “cebo-Cub-indicador” y “Nub-presa”, en las que Nub y

Cub representan las mitades N- y C-terminales respectivas del monómero de ubiquitina (Johnsson y Varshavsky, 1994, Proc. Natl. Acad. Sci. U. S. A. 91: 10.340-10.344).

[0163] Tras la interacción entre *cebo* y *presa* a través de un compuesto químico R1-Y-R2, las mitades de ubiquitina se ponen en estrecho contacto y se vuelven a asociar para formar una unidad que es suficientemente bien reconocida por UBP (proteasas específicas de ubiquitina). Este acontecimiento de reconocimiento conduce a la escisión proteolítica y posterior liberación del indicador fusionado C-terminalmente.

[0164] En una estrategia de triple híbrido típica, la reasociación de las mitades de ubiquitina con la posterior liberación del indicador dependería de una interacción de molécula pequeña-proteína, más que de una interacción de proteína-proteína. La construcción de *cebo* emplearía una fusión de "receptor-Cub-indicador" (P1-Cub-RM). De forma similar a la tecnología detectora de proteína ubiquitina dividida, las construcciones de "receptor-Cub-indicador" y de Nub-*cebo* se expresan a partir de 2 vectores lanzadera separados. La molécula pequeña a investigar se fusiona con un grupo funcional común que se une al "receptor". El receptor puede ser DHFR (deshidrofolato reductasa). Aquí, la DHFR funciona como receptor para el grupo funcional común metotrexato (Mtx). El Mtx o sus derivados con un grupo funcional similar (tales como 2,4-diaminopteridina) se fusionarán a diversas moléculas pequeñas con numerosas moléculas enlazadoras diferentes. La propia molécula pequeña se analizará para determinar su interacción con proteínas presentes en una biblioteca de Nub-*presa*. La interacción del compuesto con una *presa* conducirá a la formación de un puente de R-Cub-DHFR::Mtx-molécula pequeña::presa-Nub, poniendo de este modo a Cub y Nub (o Nux) en estrecho contacto, conduciendo a la liberación del resto indicador RM.

[0165] El resto indicador puede desencadenar cualquier clase de cambio detectable, es decir, puede depender de la detección de productos de corte y empalme proteolítico por electroforesis en gel y/o análisis de transferencia de Western, lectura de fluorescencia o enzimática, complementación nutricional u otras formas de lectura transcripcional.

[0166] El resto indicador puede ser un factor de transcripción unido a una membrana celular que impide su entrada en el núcleo y la activación de la transcripción. Sólo tras la reasociación de las mitades de ubiquitina después de la interacción de compuesto-proteína, el resto indicador se liberará y se translocará al núcleo, donde puede activarse la transcripción de un gen indicador. Los genes indicadores pueden ser enzimas, marcadores fluorescentes o marcadores nutricionales (por ejemplo, lacZ, proteína verde fluorescente GFP/proteína roja fluorescente optimizada para codones de levadura yRFP, HIS/URA) (Stagljär et al. (1998) Proc. Natl. Acad. Sci. U.S.A., 95: 5187-92).

[0167] La invención usa genes marcadores de selección negativa o "indicadores de selección" que pueden usarse en una célula hospedadora eucariota, preferiblemente una célula de levadura o de mamífero, o una célula procariota, y que pueden seleccionarse en contra en condiciones apropiadas. En realizaciones preferidas, el indicador de selección se proporciona como un polipéptido de fusión con un subdominio carboxi o C-terminal de ubiquitina (o Cub) y se altera para codificar un resto aminoacídico distinto de metionina en la unión con el Cub. El resto aminoacídico distinto de metionina es preferiblemente un aminoácido que se reconoce por el sistema de proteasa de ubiquitina de la regla del N-terminal (por ejemplo, un resto de arginina, lisina, histidina, fenilalanina, triptófano, tirosina, leucina o isoleucina) y que, cuando está presente en el extremo amino-terminal del marcador de selección negativa, dirige al marcador de selección negativa para una degradación proteolítica rápida.

[0168] Un ejemplo preferido de un gen marcador de selección para usar en levaduras es el gen URA3 que puede tanto seleccionarse a favor (selección positiva) por cultivo de cepas de levadura auxótrofas *ura3* en ausencia de uracilo, como seleccionarse en contra (selección negativa) por cultivo de células en medios que contienen ácido 5-fluoroórótico (5-FOA) (véase Boeke, et al (1987) Methods Enzymol 154: 164-75). La concentración de 5-FOA puede optimizarse por valoración para seleccionar lo máximo a favor de células en las que el indicador URA3 esté, por ejemplo, inactivado por degradación proteolítica en cierto grado preferido. Por ejemplo, pueden usarse concentraciones relativamente elevadas de 5-FOA que permitan que sobrevivan únicamente las células que expresen niveles en estado estacionario muy bajos de indicador URA3. Dichas células corresponderán a aquellas en las que la primera y segunda proteínas de fusión de subdominio de ubiquitina tengan una afinidad relativamente alta entre sí, dando como resultado un reensamblaje eficaz de los fragmentos Nub y Cub y una liberación correspondientemente eficaz del marcador inestabilizado Z-URA3. Por el contrario, pueden usarse menores concentraciones de 5-FOA para seleccionar a favor de compañeros de unión proteicos con afinidades relativamente débiles entre sí. Además, puede usarse prolina en los medios como fuente de nitrógeno para hacer a las células hipersensibles a los efectos tóxicos del 5-FOA (McCusker y Davis (1991), Yeast 7: 607-8). Por consiguiente, las concentraciones de prolina, así como las concentraciones de 5-FOA pueden valorarse para obtener una selección óptima para células deficientes en indicador URA3. Por lo tanto, el uso de URA3 como marcador de selección negativa permite una amplia variedad de rigurosidades selectivas que pueden adaptarse para minimizar las interferencias de fondo de falsos positivos y/o optimizar la selección para interacciones de unión de alta afinidad. Otros marcadores de selección negativa que funcionan en levaduras y que pueden adaptarse al método de la invención se incluyen dentro del alcance de la invención.

[0169] Se conocen en la técnica numerosos marcadores de selección que funcionan en células de mamífero y pueden adaptarse al método de la invención para permitir la selección negativa directa de proteínas de interacción en células de mamífero. Los ejemplos de marcadores de selección negativa de mamífero incluyen la timidina quinasa (Tk) (Wigler et al (1977) Cell 11: 223-32; Borrelli et al (1988) Proc Natl Acad Sci. USA 85: 7572-76) del virus herpes simple, el gen humano para la hipoxantina fosforribosil transferasa (HPRT) (Lester et al (1980) Somatic Cell Genet. 6: 241-59; Albertini et al. (1985) Nature 316: 369-71) y la citidina desaminasa (codA) de *E. coli* (Mullen et al (1992) Proc. Natl. Acad. Sci.

USA. 89: 33-37; Wei y Huber (1996) J. Biol. Chem. 271: 3812-16). Por ejemplo: el gen Tk puede seleccionarse en contra usando ganciclovir (GANC) (por ejemplo, usando una concentración 1 μ M) y el gen *codA* puede seleccionarse en contra usando 5-fluorocitidina (5-FIC) (por ejemplo, usando una concentración de 0,1-1,0 mg/ml). Además, se han descrito ciertos marcadores de selección quiméricos (Karreman (1998) Gene 218: 57-61) en los que un marcador de selección negativa de mamífero funcional se fusiona a un marcador de selección positiva de mamífero funcional, tal como resistencia a higromicina (Hyg^R), resistencia a neomicina (neo^R), resistencia a puomicina (PAC^R) o resistencia a blasticidina S (Bla^R). Estos producen diversos marcadores de selección positiva/negativa basados en Tk para células de mamífero tales como HygTk, Tkneo, TkBSD y PACTk, así como diversos marcadores de selección positiva/negativa basados en *codA* para células de mamífero, tales como HygCoda, Codaneo, CodaBSD y PACCoda. También se han descrito recientemente indicadores Tk-neo que incorporan luciferasa, proteína verde fluorescente y/o beta-galactosidasa (Strathdee et al (2000) BioTechniques 28: 210-14). Estos vectores tienen la ventaja de permitir una fácil exploración del marcador/indicador "positivo" por microscopía fluorescente y/o inmunofluorescente. El uso de dichos marcadores de selección positiva/negativa proporciona las ventajas mencionadas anteriormente para URA3 como indicador en levaduras, en vista de que permiten que las células de mamífero se evalúen por métodos de selección tanto positiva como negativa para determinar la expresión y el nivel en estado estacionario relativo de la fusión indicadora. Otras ventajas de estas construcciones de marcador de selección e indicador de mamífero serán evidentes para el experto en la materia.

3.1.2 Componentes de la ruta proteolítica de la regla del N-terminal

[0170] El sistema de la "regla del N-terminal" para la degradación proteolítica es una ramificación particular de la ruta proteolítica mediada por ubiquitina presente en células eucariotas (Bachmair et al (1986) Science 234: 179-86). Este sistema funciona degradando un polipéptido celular a una velocidad dependiente del resto aminoacídico amino-terminal de ese polipéptido. La traducción de proteína se inicia comúnmente con un codón de metionina ATG y, de ahí que la mayoría de polipéptidos tengan un resto de metionina amino-terminal y que típicamente sean relativamente estables *in vivo*. Por ejemplo, en la levadura *S. cerevisiae*, un polipéptido de beta-galactosidasa con un extremo amino-terminal de metionina tiene una semivida de >20 horas (Varshavsky (1992) Cell 725-35). En determinadas circunstancias, sin embargo, pueden crearse polipéptidos que posean un resto amino-terminal distinto de metionina. Por ejemplo, cuando una endoproteasa hidroliza y, por lo tanto, escinde un enlace polipeptídico único (A-B) interno a un polipéptido, da como resultado la liberación de dos polipéptidos separados —uno de los cuales posee un aminoácido amino-terminal, Z, que puede no ser metionina. Por ejemplo, la endoproteasa proteasa específica de ubiquitina, que es un componente preferido de la presente invención, escindirán un enlace polipeptídico carboxi-terminal con el resto de glicina final (codón 76) independientemente de cuál sea el codón siguiente. En la función normal de la célula, esta proteasa específica sirve para escindir un precursor de poliubiquitina en unidades de ubiquitina individuales. Sin embargo, también puede usarse para generar un polipéptido diana con prácticamente cualquier resto amino-terminal, fusionando simplemente el polipéptido diana en fase de lectura con un codón correspondiente al aminoácido amino-terminal deseado (Z), estando dicho codón, a su vez, fusionado cadena abajo de la ubiquitina (típicamente contiguo al codón Gly 76 de ubiquitina). La construcción quimera de gen diana resultante tiene la fórmula general ubiquitina-Z-diana. Las construcciones diana preferidas comprenden además una etiqueta epitópica (Ep) de modo que la construcción quimera de gen diana resultante tiene la fórmula general ubiquitina-Z-Ep-diana, que da como resultado la producción final de un polipéptido de fórmula general Z-Ep-diana. Las actividades de proteasa específica de ubiquitina constitutivamente activa presentes en células eucariotas darán como resultado el procesamiento endoproteolítico del polipéptido de ubiquitina-Z-diana en ubiquitina y entidades Z-diana. Se actúa adicionalmente sobre el polipéptido Z-diana por los componentes del sistema de la regla del N-terminal que se describen a continuación. Si el polipéptido diana es un marcador de selección negativa (NSM) y si Z es un resto aminoacídico (tal como arg) que potencia una degradación rápida por el sistema de la regla del N-terminal, entonces las células que expresan ubiquitina-Z-NSM intacta pueden seleccionarse en contra, mientras que las células en las que la fusión está prendida a un polipéptido Z-NSM relativamente inestable pueden seleccionarse a favor.

[0171] Se ha determinado, con una fiabilidad razonable, el efecto relativo de un resto amino-terminal dado, Z, sobre la estabilidad del polipéptido diana. Por ejemplo, cuando los 20 restos aminoacídicos amino terminales posibles se ensayaron para determinar su efecto sobre la estabilidad de la beta-galactosidasa (utilizando una fusión quimérica de ubiquitina-Z-beta-galactosidasa) en *Saccharomyces cerevisiae*, se descubrieron diferencias drásticas (véase Varshavsky (1992) Cell 69: 725-35). Por ejemplo, cuando Z era, met, cys, ala, ser, thr, gly, val o pro, el polipéptido resultante era muy estable (semivida de >20 horas). Cuando Z era tyr, ile, glu o gln, el polipéptido resultante poseía una estabilidad de proteína moderada (semivida de 10-30 minutos). Por el contrario, los restos arg, lys phe, leu, trp, his, asp y asn conferían todos una baja estabilidad al polipéptido de beta-galactosidasa (semivida de <3 minutos). El resto de arginina (arg), cuando se localizaba en el extremo amino-terminal de un polipéptido, parecía conferir generalmente, la menor estabilidad. Por lo tanto, son realizaciones generalmente preferidas de la presente invención construcciones quiméricas y los polipéptidos de fusión correspondientes que emplean un resto arg en la posición Z.

[0172] Los experimentos descritos anteriormente que establecen las semividas relativas conferidas por cada uno de los 20 posibles restos amino-terminales constituyen la base

[0173] de la regla del N-terminal. Los componentes del sistema de la regla del N-terminal son aquellos productos génicos que actúan provocando la proteólisis rápida de polipéptidos que poseen restos amino-terminales que confieren inestabilidad. El sistema de la regla del N-terminal para proteólisis en eucariotas parece ser una parte de las rutas del

sistema proteolítico dependiente de ubiquitina general que poseen aparentemente todas las células eucariotas. En resumen, este sistema implica el marcaje covalente de un polipéptido diana en uno o más restos de lisina mediante un marcador polipeptídico de ubiquitina (para formar un enlace covalente de diana (lys)-épsilon amino-gly(76) ubiquitina). Posteriormente pueden conjugarse restos de ubiquitina adicionales con el polipéptido diana y el polipéptido diana “ubiquitinado” resultante se somete después a destrucción proteolítica completa por un complejo multiproteico grande (26S) conocido como proteosoma. Las enzimas que conjugan los restos de ubiquitina con la proteína diana incluyen las funciones E2 y E3 (o ligasa de ubiquitina). Se cree que las enzimas E2 y E3 poseen la mayor parte de la especificidad por procesos proteolíticos dependientes de ubiquitina.

[0174] Un componente clave de la ruta proteolítica de la regla del N-terminal en levadura es *UBR1* (Bartel, et al (1990), EMBO J. 9: 3179-89), un gen que codifica una función tipo E3 que parece reconocer polipéptidos que poseen restos amino-terminales susceptibles y de este modo facilita la ubiquitinación de dichos polipéptidos (Dohmen et al (1991) Proc Natl. Acad. Sci. USA. 88: 7351-55). Por consiguiente, puede usarse *UBR1* como componente de la regla del N-terminal regulable que es el efector de la degradación proteolítica del polipéptido del gen diana. El gen de *UBR1* se ha clonado ahora a partir de un organismo de mamífero (Kwon et al. (1998) Proc. Natl. Acad. Sci. USA. 95: 7893-903), así como a partir de levadura. Por lo tanto, la construcción de una línea celular de ratón *knock out* para *UBR1* es inminente y, de este modo, el control de la inestabilidad de fusiones de Z-indicador puede manipularse adicionalmente controlando el nivel de *UBR1* expresado.

[0175] El gen de *UBR1* es particularmente fundamental para ciertos aspectos de la presente invención porque puede usarse selectivamente junto con cualquiera de los restos desestabilizantes amino-terminales “Z” distintos de metionina descritos anteriormente, incluyendo: el más desestabilizante —arg; restos fuertemente desestabilizantes —tales como lys, phe, leu, trp, his, asp y asn; y restos moderadamente desestabilizantes —tales como tyr, ile, glu o gln. De hecho, un objeto de ciertas realizaciones de la presente invención es proporcionar un medio, cuando se desee, para no desactivar completamente la función de un marcador de selección negativa, sino simplemente atenuarlo en cierta medida fijada. Esto puede conseguirse usando el método de la presente invención de cualquiera de varias formas. Por ejemplo, puede desplegarse un resto amino-terminal moderadamente desestabilizante (Z = tyr, ile, glu o gln) en el indicador de polipéptido diana —dando como resultado una eliminación menos rápida de la combinación de polipéptidos diana.

[0176] Otros componentes de la regla del N-terminal para usar en la presente invención incluyen *UBC2* (*RAD6*) de *S. cerevisiae*, que codifica una función de conjugación de ubiquitina E2 que coopera con la E3 de la regla del N-terminal codificada por *UBR1* para promover la multiubiquitinación y posterior degradación de sustratos de la regla del N-terminal (Dohmen et al (1991) Proc. Natl. Acad. Sci. USA. 88: 7351-55). Por lo tanto, la proteólisis dirigida por la regla del N-terminal no se producirá en ausencia de *UBR1* o *UBC2*. Esto permite que se use cualquier gen como “efector de proteólisis dirigida” inducible por métodos de la presente invención. De hecho, un polipéptido de gen diana que posea un aminoácido amino-terminal desestabilizante de la regla del N-terminal (tal como arg) será estable hasta que se induzca la expresión del *UBR1* (E3) o del *UBC2* (E2) a partir de la construcción promotora inducible afín.

[0177] Tanto *UBR1* como *UBC2* pueden usarse junto con cualquiera de los restos desestabilizantes amino-terminales “Z” descritos anteriormente, incluyendo: el más desestabilizante —arg; restos fuertemente desestabilizantes —tales como lys, phe, leu, trp, his, asp y asn; y restos moderadamente desestabilizantes —tales como tyr, ile, glu o gln. Otras realizaciones alternativas más del componente de la regla del N-terminal de la presente invención son componentes del sistema de la regla del N-terminal que afectan sólo a un subconjunto de los restos desestabilizantes. Por ejemplo, la *NTA1* desamidasa (Baker y Varshavsky (1995) J Biol Chem 270: 12065-74) funciona desaminando los restos de asn o gln amino-terminales (para formar polipéptidos con restos amino-terminales asp o glu, respectivamente). Las cepas de levadura que albergan alelos nulos de *nta1* son incapaces de degradar sustratos de la regla del N-terminal que lleven restos de asn o gln amino-terminales. Por lo tanto, el gen de *NTA1* es una realización alternativa del componente de la regla del N-terminal de la presente invención, pero se usa preferiblemente junto con un polipéptido de gen diana (Z-diana), en el que Z es asn o gln. De forma similar, la *ATE1* transferasa (Balzi et al (1990) J. Biol Chem 265: 7464-71) es una enzima que actúa transfiriendo el resto de arg de un ARNt activado ARNt-Arg a un polipéptido que lleve glu o asp amino-terminales. El polipéptido con arg-glu resultante y los productos polipeptídicos con arg-asp son entonces susceptibles a los procesos proteolíticos dependientes de la regla del N-terminal mediados por E2/E3 descritos anteriormente. Por lo tanto, la *ATE1* transferasa es una realización alternativa del componente de la regla del N-terminal de la presente invención, pero su uso está preferiblemente relacionado con polipéptidos de gen diana (Z-diana), en los que Z es asp, glu, asn o gln. Los polipéptidos que llevan los últimos dos restos amino-terminales se convierten primero en polipéptidos que llevan uno de los primeros dos restos amino-terminales por la función *NTA1* desamidasa descrita anteriormente.

[0178] Es importante señalar en la presente memoria que, como es el caso para el represor que se somete a la inducción por un promotor inducible, el componente de la regla del N-terminal debe estar disponible como un clon de modo que pueda ponerse bajo el control de un promotor inducible (usando métodos de subclonación convencionales conocidos en la técnica). Esto puede conseguirse introduciendo primero copias modificadas por ingeniería genética del represor inducible y de las construcciones de componentes de la regla del N-terminal inducibles, y delecionando posteriormente las copias cromosómicas normales de estos genes del hospedador por métodos de inactivación génica “*knock out*”. Dichos métodos que se señalan en la presente memoria están bien desarrollados en la técnica — particularmente en el caso de tanto la levadura *Saccharomyces cerevisiae* como el mamífero ratón. Sin embargo, es más conveniente la disponibilidad de la tecnología de inserción génica “*knock in*” que permite que la copia cromosómica

existente del gen se modifique de modo que se deleccione su promotor nativo y se inserte un promotor inducible en una sola etapa.

3.1.3 Secuencias polipeptídicas de ubiquitina

[0179] Se proporciona una descripción completa y detallada de las construcciones de Cub y Nub que pueden usarse en el método de la presente invención en las Patentes de Estados Unidos Nº 5.503.977 y 5.585.245. Se le supone formación respecto a la biología molecular del sistema proteolítico de ubiquitina en general y el sistema de la regla del N-terminal y del ensayo de asociación detector de ubiquitina al experto en la materia que busca la práctica de la presente invención. En resumen, la ubiquitina (Ub) es una proteína de un solo dominio de 76 restos cuyo acoplamiento covalente a otras proteínas produce conjugados de Ub-proteína ramificados, y desempeña un papel en varios procesos celulares, principalmente a través de rutas que implican la degradación de proteínas. A diferencia de los conjugados de Ub ramificados, que se forman postraduccionalmente, los aductos de Ub lineales son los productos de traducción de fusiones de Ub naturales u obtenidas por ingeniería genética. Se ha demostrado que, en eucariotas, las fusiones de Ub recién formadas se escinden rápidamente en la unión de Ub-polipéptido por proteasas específicas de Ub (UBP). En la levadura *Saccharomyces cerevisiae*, hay al menos cinco especies de UBP. Trabajos recientes han demostrado que la escisión de una fusión de Ub por UBP requiere la conformación plegada de Ub, debido a que se observa escasa o ninguna escisión con fusiones cuyo resto Ub estaba conformacionalmente desestabilizado por sustituciones de un solo resto o una delección distante del sitio de escisión por UBP.

[0180] La presente invención depende en parte del sistema detector de proteína ubiquitina dividida descrito anteriormente (véanse las Patentes de Estados Unidos Nº 5.503.977 y 5.585.245 y el documento WO 02/12902). En resumen, se ha demostrado que un subdominio de ubiquitina N-terminal y un subdominio de ubiquitina C-terminal, llevando el último una extensión indicadora en su extremo C-terminal, cuando se coexpresan en la misma célula por técnicas de ADN recombinante como entidades diferentes, tiene la capacidad de asociarse, reconstituyendo una molécula de ubiquitina que se reconozca y se escinda por proteasas de procesamiento específicas de ubiquitina que están presentes en todas las células eucariotas. Esta molécula de ubiquitina reconstituida que se reconoce por proteasas específicas de ubiquitina se denomina en la presente memoria resto de ubiquitina casi nativa. Como se describe en la presente memoria, las proteasas específicas de ubiquitina reconocen la conformación plegada de la ubiquitina. Sorprendentemente, las proteasas específicas de ubiquitina conservaban su actividad de escisión y su especificidad de reconocimiento del resto de ubiquitina que se había reconstituido a partir de dos subdominios de ubiquitina no unidos.

[0181] La ubiquitina es una proteína de un solo dominio de 76 restos que comprende dos subdominios que son relevantes para la presente invención —el subdominio N-terminal y el subdominio C-terminal. La proteína ubiquitina se ha estudiado ampliamente y se ha publicado la secuencia de ADN que codifica la ubiquitina (Ozkaynak et al., EMBO J. 6: 1429 (1987)). El subdominio N-terminal (Nub), mencionado en la presente memoria, es esa porción de la molécula de ubiquitina nativa que se pliega en la única hélice alfa de ubiquitina que interacciona con dos cadenas beta. Hablando en general, este subdominio comprende los restos aminoacídicos de aproximadamente el resto número 1 a aproximadamente el resto número 36.

[0182] El subdominio C-terminal de ubiquitina (Cub), mencionado en la presente memoria, es esa porción de la ubiquitina que no es una porción del subdominio N-terminal definido en el párrafo anterior. Hablando en general, este subdominio comprende los restos aminoacídicos de aproximadamente 37 a aproximadamente 36. Debería reconocerse que usando solamente experimentación de rutina será posible definir con precisión los requisitos mínimos en ambos extremos del subdominio N-terminal y del subdominio C-terminal que son necesarios para que sean útiles en relación con la presente invención.

[0183] Es importante señalar que el Nub se refiere, en realizaciones preferidas de la invención, a la unidad del subdominio de ubiquitina amino-terminal que se ha mutado para disminuir su afinidad de unión, haciendo de este modo a la asociación de Cub/Nub dependiente de la unión de una segunda pareja proteica fusionada a las subunidades Cub y Nub. Se describen a continuación formas adecuadas de Nub y otras más están fácilmente disponibles para el experto en la materia por métodos de exploración y mutación de rutina.

[0184] Para estudiar la interacción entre un ligando híbrido y una pareja de dominios de unión a ligando, un miembro de la pareja se fusiona al subdominio N-terminal de ubiquitina y el otro miembro de la pareja se fusiona al subdominio C-terminal de ubiquitina. Puesto que los miembros de la pareja de unión específica (unidos a subdominios de ubiquitina) tienen afinidad por el ligando híbrido, esta afinidad aumenta la concentración (local) "eficaz" de los subdominios N-terminal y C-terminal de ubiquitina, promoviendo de este modo la reconstitución de un resto de ubiquitina casi nativa. Por comodidad, la expresión "resto de ubiquitina casi nativa" se usará en la presente memoria para indicar un resto reconocible como sustrato por proteasas específicas de ubiquitina. A la luz del hecho de que los subdominios N-terminal y C-terminal de ubiquitina se asocian para formar un resto de ubiquitina casi nativa incluso en ausencia de fusión de los dos subdominios con miembros individuales de la pareja de dominios de unión a ligando, puede imponerse un requisito adicional en ciertas realizaciones de la presente invención para aumentar la capacidad de resolución del método para estudiar dichas interacciones. Este requisito adicional preferido es que el subdominio N-terminal de ubiquitina puede ser un aliado de mutación alterado para reducir su capacidad para producir, a través de la asociación con Cub, un resto de ubiquitina casi nativa. Un experto en la materia reconocerá que los estudios de interacciones de unión descritos en la presente memoria se llevan a cabo en condiciones apropiadas para la interacción de proteína/ligando. Tales

condiciones se proporcionan *in vivo* (es decir, en condiciones fisiológicas en el interior de células vivas) o *in vitro*, cuando parámetros tales como la temperatura, el pH y la concentración salina se controlan de una forma destinada a imitar las condiciones fisiológicas.

[0185] La alteración por mutación de un subdominio de ubiquitina amino-terminal para usarse con la presente invención es preferiblemente una mutación puntual. A la luz del hecho de que es esencial que el resto de ubiquitina reconstituida debe “verse y sentirse” como la ubiquitina nativa para una proteasa específica de ubiquitina, deben evitarse alteraciones mutacionales que se esperaría que afectasen en extremo a la estructura del subdominio que lleva la mutación. Se han descrito varias proteasas específicas de ubiquitina y también se conocen las secuencias de ácido nucleico que codifican dichas proteasas (véase, por ejemplo, Tobias et al., J. Biol. Chem. 266: 12021 (1991); Baker et al., J. Biol. Chem. 267: 23364 (1992)). Debería añadirse que todas de las al menos cinco proteasas específicas de ubiquitina en la levadura *S. cerevisiae* requieren una conformación plegada de ubiquitina para su reconocimiento como sustrato. Amplias deleciones dentro del subdominio N de ubiquitina son un ejemplo del tipo de alteración mutacional que se esperaría que afectase en extremo a la estructura del subdominio y, por lo tanto, son ejemplos de tipos de alteraciones mutacionales que deberían evitarse.

[0186] A la luz de esta consideración, la alteración mutacional preferida dentro de la subunidad Nub es una mutación en la que se efectúa una sustitución de aminoácido. Por ejemplo, se prefiere la sustitución de un aminoácido que tiene propiedades químicas similares al aminoácido sustituido (por ejemplo, una sustitución conservativa). En concreto, la perturbación leve deseada de la interacción del subdominio de ubiquitina se consigue sustituyendo un resto aminoacídico químicamente similar que difiera principalmente en el tamaño de su cadena lateral. Se espera que dicha perturbación estérica introduzca una desestabilización conformacional (leve) deseada de un subdominio de ubiquitina. El objetivo es reducir la afinidad de los subdominios N-terminal y C-terminal entre sí, sin eliminar necesariamente esta afinidad.

[0187] Por ejemplo, la alteración mutacional puede introducirse en el subdominio N-terminal de ubiquitina. Más específicamente, un primer resto aminoacídico neutro puede sustituirse con un segundo aminoácido neutro que tenga una cadena lateral que difiera en tamaño de la cadena lateral del primer resto aminoacídico neutro para conseguir la disminución deseada de la afinidad. Por ejemplo, el primer resto aminoacídico neutro isoleucina (resto 3 ó 13 de la ubiquitina de tipo silvestre) puede sustituirse con un aminoácido neutro que tenga una cadena lateral que difiera en tamaño de la isoleucina, tal como glicina, alanina o valina (véase Johnsson y Varshavsky, 1994, Proc. Natl. Acad. Sci. U.S.A. 91: 10340-10344).

[0188] Pueden usarse una amplia diversidad de combinaciones de construcciones de fusión en los métodos de esta invención. Un requisito estricto que es aplicable a todas las combinaciones de construcciones de fusión N- y C-terminales es que el subdominio C-terminal debe llevar una extensión de aminoácidos (por ejemplo, péptido, polipéptido o proteína). Este requisito se basa en el hecho de que la detección de la interacción entre dos proteínas de interés unidas a dos subdominios de ubiquitina se consigue a través de la escisión después del resto C-terminal del resto de ubiquitina casi nativa, con la formación de una proteína (o péptido) indicadora libre que se había unido previamente a un subdominio C-terminal de ubiquitina. Las proteasas específicas de ubiquitina escinden una fusión de ubiquitina lineal entre el resto C-terminal de ubiquitina y el resto N-terminal del compañero de fusión de ubiquitina, pero no escinden una fusión de otro modo idéntica cuyo resto de ubiquitina esté conformacionalmente alterado. En particular, no reconocen como sustrato un subdominio C-terminal de ubiquitina unido a una secuencia indicadora “cadena abajo” a menos que este subdominio C-terminal se asocie con un subdominio N-terminal de ubiquitina para dar un resto de ubiquitina casi nativa.

[0189] Además, las características de la extensión de aminoácidos C-terminal del subdominio de ubiquitina C-terminal deben ser tales que los productos de la proteína de fusión escindida sean distinguibles de la proteína de fusión no escindida. En la práctica, esto se logra generalmente controlando una propiedad física o actividad de la extensión C-terminal que se libera por escisión del resto de ubiquitina C-terminal. Generalmente, se controla una propiedad de la extensión C-terminal libre como indicio de que se ha formado una ubiquitina casi nativa, ya que el control del resto de ubiquitina casi nativa directamente es difícil en células eucariotas debido a la presencia de ubiquitina nativa. Aunque no es necesario para la práctica de la presente invención, por su puesto sería apropiado controlar directamente la presencia de la ubiquitina casi nativa también, con tal de que este control pudiera llevarse a cabo en ausencia de interferencia de la ubiquitina nativa (por ejemplo, en células procariotas que carecen de forma natural de ubiquitina).

[0190] El tamaño de la extensión C-terminal que se libera después de la escisión del resto de ubiquitina casi nativa dentro de una fusión con indicador por una proteasa específica de ubiquitina es una característica particularmente conveniente a la luz del hecho de que es relativamente fácil controlar cambios en el tamaño usando, por ejemplo, métodos electroforéticos. Por ejemplo, si la extensión indicadora C-terminal tiene un peso molecular de aproximadamente 20 kD, los productos de escisión serán distinguibles del resto de ubiquitina casi nativa no escindido en virtud de la aparición de una banda de 20 kD específica de indicador previamente ausente después de la escisión de la fusión con indicador.

[0191] A la luz del hecho de que la escisión puede tener lugar, por ejemplo, en extractos celulares brutos o *in vivo*, generalmente no es posible controlar tales cambios en el peso molecular de los productos de escisión tiñendo simplemente un electroforetograma con un colorante que tiña proteínas inespecíficamente, porque hay demasiadas proteínas en la mezcla para analizarlas de esta forma. Un método de análisis preferido es la inmunotransferencia. Este

es un método analítico convencional en el que los productos de escisión se separan electroforéticamente, generalmente en una matriz de gel de poliacrilamida, y posteriormente se transfieren a un soporte sólido cargado (por ejemplo, nitrocelulosa o una membrana de nylon cargada). Un anticuerpo que se une al indicador de los productos de escisión de proteasa específica de ubiquitina se emplea después para detectar los productos de escisión transferidos usando métodos de rutina para la detección del anticuerpo unido.

[0192] Otro método útil es la inmunoprecipitación de una fusión que contiene un indicador con subdominios C-terminales de ubiquitina o del indicador libre (liberado a través de la escisión por proteasas específicas de ubiquitina tras la reconstitución de un resto de ubiquitina casi nativa) con un anticuerpo contra el indicador. Las proteínas que se van a inmunoprecipitar se marcan primero *in vivo* con un aminoácido radiactivo tal como ³⁵S-metionina usando métodos de rutina en la técnica. Después, se prepara un extracto celular y las proteínas que contienen indicador se precipitan a partir del extracto usando un anticuerpo anti-indicador. Las proteínas inmunoprecipitadas se fraccionan por electroforesis en un gel de poliacrilamida, seguido de detección de especies proteicas radiactivas por autorradiografía o fluorografía.

[0193] Un diseño experimental preferido consiste en prolongar el subdominio C-terminal de ubiquitina con un péptido que contiene un epítipo extraño para el sistema en el que se esté llevando a cabo el ensayo. También es preferible diseñar el experimento de modo que la extensión indicadora C-terminal del subdominio C-terminal de ubiquitina sea suficientemente grande, es decir, fácilmente detectable por el sistema electroforético empleado. En esta realización preferida, la extensión indicadora C-terminal del subdominio C-terminal debería verse como un marcador de peso molecular. Las características de la extensión, aparte de su peso molecular y de su reactividad inmunológica no son de una importancia particular. Por lo tanto, se reconocerá que esta extensión C-terminal puede representar una amalgama que comprenda prácticamente cualquier combinación de secuencias de aminoácidos fusionada a un epítipo para el que esté disponible un anticuerpo que se una específicamente. Por ejemplo, la extensión C-terminal del subdominio de ubiquitina C-terminal puede ser una combinación del epítipo "ha" fusionado a DHFR de ratón (está fácilmente disponible un anticuerpo contra el epítipo "HA").

[0194] Aparte del peso molecular de la extensión de aminoácidos C-terminal del subdominio de ubiquitina C-terminal, también pueden controlarse otras características para detectar la escisión de un resto de ubiquitina casi nativa. Por ejemplo, la actividad enzimática de algunas proteínas puede suprimirse prolongando su extremo N-terminal. Dicha enzima "indicadora" que, en su forma nativa, muestra una actividad enzimática que se suprime cuando la enzima se prolonga N-terminalmente, también puede servir como indicador C-terminal unido al subdominio de ubiquitina C-terminal.

[0195] En este esquema de detección, cuando el indicador está presente como una fusión con el subdominio de ubiquitina C-terminal, la proteína indicadora es inactiva. Sin embargo, si el subdominio de ubiquitina C-terminal y el subdominio de ubiquitina N-terminal se asocian para reconstruir un resto de ubiquitina casi nativa en presencia de una proteasa específica de ubiquitina, la proteína indicadora se liberará, con la restauración concomitante de su actividad enzimática.

[0196] En realizaciones preferidas, la proteína indicadora es un marcador de selección negativa eucariota (NSM) que se ha modificado por ingeniería genética para procesarse y liberarse como una fusión de Z-NSM inestable a la regla del N-terminal después de la escisión proteolítica por proteasa específica de ubiquitina. Los marcadores de selección negativa (NSM) para usar en la invención se describen en otra parte. La ventaja de usar una fusión de Z-NSM es que la interacción de la pareja de unión específica puede seleccionarse directamente (en lugar de explorarse) en virtud del hecho de que sólo las células en las que se ha liberado Z-NSM sobrevivirán a la selección negativa.

[0197] El indicador de gen diana (marcador de selección negativa) debe fusionarse cadena abajo de un codón que codifique un resto susceptible a la regla del N-terminal (Z, como se ha descrito anteriormente), y este resto, en término, debe fusionarse en fase de lectura con el extremo carboxi-terminal de una secuencia codificante de ubiquitina (generalmente el extremo carboxi-terminal de un subdominio de ubiquitina C-terminal (Cub) que corresponde a la gly76 de la ubiquitina intacta). La razón para construir esta construcción génica quimérica amplia es aprovechar la capacidad de proteasas de ubiquitina constitutivas para escindir cualquier enlace peptídico que esté carboxi-terminal a la gly76 de una unidad de ubiquitina intacta. Esta proteasa específica de ubiquitina normalmente funciona procesando cadenas de poliubiquitina (el producto de traducción de las secuencias codificantes de ubiquitina en tándem de genomas eucariotas) en restos de ubiquitina separados (normalmente de 76 aminoácidos) que se usan en rutas del sistema de ubiquitina. En el método de la presente invención, las proteasas específicas de ubiquitina sirven como medio conveniente para generar polipéptidos de genes diana que lleven restos amino-terminales específicos (Z). No obstante, se entiende que existen otras alternativas a la ubiquitina de mamíferos o levadura que pueden funcionar en el método de la presente invención. Dichos equivalentes de ubiquitina incluyen, por ejemplo, mutantes de ubiquitina, proteínas tipo ubiquitina, proteínas relacionadas con ubiquitina y proteínas homólogas a ubiquitina. Por ejemplo, las proteínas tipo ubiquitina tales como NEDD8, UBL1, FUBI y UCRP, así como proteínas relacionadas con ubiquitina análogas, tales como SUMO/Sentrin/Pic1, pueden usarse como equivalentes de ubiquitina en el método de la invención. Otras proteínas relacionadas con ubiquitina, pero que son algo menos homólogas a la misma, incluyen proteínas homólogas a ubiquitina tales como Rad23 y Dsk2, cuya similitud con la ubiquitina no incluye la presencia de un par de glicinas carboxilo-terminales. Estas proteínas de tipo ubiquitina comparten las características comunes de estar relacionadas con la ubiquitina por homología de secuencia de aminoácidos y, con la excepción evidente de las proteínas homólogas a ubiquitina, de transferirse covalentemente a dianas proteicas celulares postraduccionalmente.

[0198] De hecho, en algunas realizaciones, el alcance deseado de la invención inmediata incluye cualquier medio conocido en la técnica por el que pueda generarse un polipéptido diana que lleve un resto susceptible a la regla del N-terminal (Z = arg, lys, his, leu, phe, try, ile, trp, asn, gln, asp o glu). Se conocen bien en la técnica métodos generales para modificar por ingeniería genética dichos restos N-terminales en vectores de expresión quimera de ubiquitina-indicador (por ejemplo, el método de "PCR de fusión"; véase Karreman (1988) *BioTechniques* 24: 736-42).

[0199] La descripción resumida del párrafo anterior no analiza ciertas consideraciones experimentales importantes. Por ejemplo, para dos proteínas de interacción, P1 (fusionada a Nub) y P2 (fusionada a Cub) se incluyen las consideraciones adicionales siguientes en el alcance de la invención. A la luz de su papel como componente de afinidad, se reconocerá que P1 puede fusionarse al extremo N-terminal o al extremo C-terminal del subdominio de ubiquitina N-terminal. De forma similar, P2 puede fusionarse al extremo N-terminal o al extremo C-terminal del subdominio de ubiquitina C-terminal. Si P2 se fusiona al extremo C-terminal del subdominio de ubiquitina C-terminal, se eliminará por escisión por la proteasa específica de ubiquitina, con tal de que los subdominios de ubiquitina se asocien para formar un resto de ubiquitina casi nativa. De acuerdo con la descripción resumida en el párrafo anterior, si el resto P2 se fusiona al extremo C-terminal del subdominio de ubiquitina C-terminal, también puede usarse como indicador para detectar la reconstitución de un resto de ubiquitina casi nativa. Además, la posición de P2 dentro de la región que contiene el indicador C-terminal de la fusión no es una consideración crítica.

3.1.4 Detección de la escisión del resto indicador

[0200] La forma más directa para detectar la escisión del resto indicador es detectando la presencia del "RM libre" escindido. Un ensayo de rutina para ese tipo de detección se consigue por transferencia de Western usando un anticuerpo específico para el RM. No es necesaria una actividad adicional del RM siempre que sea razonablemente estable. Por esta razón, deberá estar presente una Met en el extremo N-terminal del RM escindido. Como alternativa, si el extremo N-terminal del RM escindido tiene un aminoácido desestabilizante y, por lo tanto, la forma de RM libre se degradará, la detección del RM no escindido unido a Cub también será capaz de evaluar el grado de escisión que se ha producido. Para evitar la necesidad de un anticuerpo para cada RM particular, puede fusionarse una etiqueta epitópica (tal como HA, myc o cualquier otra etiqueta usada rutinariamente contra la que puedan existir anticuerpos disponibles en el mercado) al RM en una localización apropiada, tal como el extremo C-terminal. La transferencia de Western es bien conocida en la técnica y puede encontrarse en varios manuales de laboratorio.

[0201] Si el RM tiene una actividad enzimática que sólo está presente cuando el RM se elimina por escisión de la fusión de Cub-RM, el grado de escisión también puede determinarse indirectamente ensayando la actividad enzimática del RM libre. Por ejemplo, algunas quinasas pueden ser inactivas cuando están fusionadas a un dominio inhibidor N-terminal y activarse después de eliminar el dominio inhibidor. Dichas quinasas pueden usarse como RM para esta realización de la invención. Una Met constituirá preferiblemente el extremo N-terminal del RM libre.

[0202] De forma similar, si un RM se inactiva/degrada enzimáticamente cuando se elimina por escisión de la fusión, también puede usarse un ensayo de la actividad enzimática para determinar el grado de escisión. Para ese ensayo, un aminoácido distinto de Met es preferiblemente el primer aminoácido del RM escindido.

[0203] Otras actividades del RM pueden ser útiles para detectar la escisión. Por ejemplo, si el RM es una proteína fluorescente, entonces el RM escindido puede degradarse por UBP si el primer aminoácido no es Met. Pueden medirse cambios en la intensidad fluorescente para indicar el grado de escisión.

[0204] Si el RM es un factor de transcripción (por ejemplo PLV, Staglar et al. (1998) *Proc. Natl. 20 Acad. Sci. U.S.A.*, 95: 5187-92), el RM escindido puede ahora volver a trasladarse al núcleo y estar disponible para la activación transcripcional de un gen indicador, cuya actividad sirve a su vez como indicador del grado de escisión. Si el RM no escindido es capaz de servir como factor de transcripción, entonces se espera que el nivel global de transcripción disminuya si el RM libre escindido es inestable, según se determina por la regla del N-terminal.

[0205] Los métodos de detección ejemplares anteriores son con fines ilustrativos solamente. Un experto en la materia será capaz de imaginar métodos equivalentes de estos ejemplos y, por lo tanto, esos métodos equivalentes se incluyen también en el alcance de la presente invención.

3.2 Sistemas indicadores basados en transcripción

[0206] De acuerdo con la invención, puede usarse un sistema indicador basado en transcripción para detectar si P1 y P2 están a una distancia corta entre sí. Un sistema indicador basado en transcripción típico es el sistema de doble híbrido de levadura, que es bien conocido en la técnica (véase a continuación). En ese sentido, P1 y P2 se sintetizan ambos como proteínas de fusión, una fusionada a un dominio de unión a ADN, la otra fusionada a un dominio de activación de la transcripción. El dominio de unión a ADN se unirá a la región promotora de un gen indicador. Si P1 y P2 están a una distancia corta entre sí (por medio de unión a R1-Y-R2), entonces el dominio de activación de la transcripción será capaz de activar la transcripción de un gen indicador, que facilitará la identificación de la proteína de ensayo o del compuesto químico pequeño de ensayo. Debido a la naturaleza asimétrica del sistema, no habrá limitaciones en lo que se refiere a si P1 o P2 se fusiona al dominio de unión a ADN o al dominio de activación de la transcripción. Además, tanto P1 como P2 pueden sintetizarse como proteínas de fusión N- o C-terminales.

[0207] La descripción detallada de diversos componentes del sistema de doble híbrido de levadura puede encontrarse fácilmente en otra parte. Por ejemplo, *The Yeast Two-Hybrid System (Advances In Molecular Biology)*, Ed. Paul L. Battel y Stanley Fields, Oxford University Press, 1997, es un libro dedicado únicamente al sistema de doble híbrido de levadura. Los pioneros en el campo proporcionan protocolos detallados, consejos prácticos sobre la resolución de problemas y sugerencias para un desarrollo futuro. Además, ilustran cómo construir una biblioteca híbrida de dominios de activación, cómo identificar mutaciones que alteren una interacción y cómo usar el sistema en células de mamífero. Los temas de capítulos incluyen la caracterización de complejos de hormona/receptor; la identificación de ligandos peptídicos; y el análisis de las interacciones mediadas por modificaciones de proteínas. También pueden encontrarse técnicas de doble híbrido y variaciones igualmente valiosas en *Yeast hybrid technologies* (Zhu, L., y Hannon, G.J., Eds., Biotechniques Press, Westborough, MA, Estados Unidos, 2000). Un tercer libro, *Two-Hybrid Systems: Methods and Protocols (Methods in Molecular Biology Vol. 177)*, Ed. Paul MacDonald, Humana Press, 2001, proporciona algunas actualizaciones recientes al campo del ensayo de doble híbrido de levadura.

[0208] Se describe también otra versión de los sistemas de doble híbrido de levadura. Por ejemplo, el sistema de doble híbrido de levadura inverso se describe en las Patentes de Estados Unidos Nº 5.955.280 y 5.965.368. Estas patentes describían métodos para identificar interacciones moleculares (por ejemplo, interacciones de proteína/proteína, proteína/ADN, proteína/ARN o ARN/ARN), empleando todas una selección y contraselección, y al menos dos moléculas híbridas. De forma similar al sistema de doble híbrido de levadura convencional, los sistemas de doble híbrido inversos también implican moléculas que interaccionan para reconstituir un factor de transcripción y dirigir la expresión de un gen indicador, cuya expresión se ensaya después. También se describen por estas patentes construcciones genéticas que son útiles en la práctica de los métodos de la invención.

[0209] Licitra y Liu (documento WO 97/41255 y Patente de Estados Unidos Nº 5.928.868) también describen un “ensayo de exploración de triple híbrido” en el que se aplica el sistema de ensayo de doble híbrido de levadura básico. La diferencia significativa es: en lugar de depender de la interacción entre una proteína denominada “cebo” y una denominada “presa”, la transcripción del gen indicador depende de la proximidad de las dos proteínas, cada una de las cuales puede unirse específicamente a uno de los dos restos de un ligando híbrido pequeño. El ligando híbrido pequeño constituye el “tercer” componente del sistema de ensayo híbrido. En ese sistema, un resto conocido del ligando híbrido se unirá a la proteína “cebo”, mientras que la interacción entre el otro resto y la proteína “presa” puede aprovecharse para explorar para una proteína que pueda unirse a un resto conocido, o un resto pequeño (compuesto farmacológico o fármaco) que pueda unirse a una diana proteica conocida.

[0210] Por ejemplo, con respecto a tecnologías de interacción de proteínas, Bartel y Fields resumen muchas estrategias/variaciones diferentes de los sistemas de doble híbrido disponibles en *The yeast-two-hybrid system* (Bartel, P.L., y Fields, S., Eds., Oxford University Press, Nueva York, NY, Estados Unidos, 1997). También pueden encontrarse técnicas de doble híbrido y variaciones igualmente valiosas en *Yeast hybrid technologies* (Zhu, L., y Hannon, G.J., Eds., Biotechniques Press, Westborough, MA, Estados Unidos, 2000). Sistemas adicionales incluyen los documentos WO 96/02561 (The General Hospital Corporation; Brent et al, Two hybrid system using conformationally constrained proteins as one of the hybrids); EP 0646644 (Bristol Myers Squibb, Menzel, un sistema de interacción unido a membrana periplásmica); WO 9825947 (Bristol Myers Squibb, Kornacker, un sistema de doble híbrido procarionta que usa *E. coli* y otras células); WO 9807845 (Dove, un sistema de trampa de interacción o “ITS” que se obtiene usando células procariontas obtenidas recombinantemente por ingeniería genética); WO 9834120 (Michnick, describe una estrategia para diseñar y aplicar ensayos de complementación de fragmentos proteicos (PCA) para detectar interacciones biomoleculares *in vivo* e *in vitro* —el sistema de exploración de la interacción de proteína DHFR. El diseño, la puesta en práctica y las amplias aplicaciones de esta estrategia se ilustran con un gran número de enzimas, proporcionándose detalles particulares para el ejemplo de la dihidrofolato reductasa (DHFR) murina. Se coexpresaron péptidos de fusión que consistían en fragmentos N- y C-terminales de la DHFR murina fusionados a secuencias de cremallera de leucina de GCN4 en *Escherichia coli* cultivadas de medio mínimo, en el que la actividad de DHFR endógena se inhibía con trimetoprim. La coexpresión de los productos de fusión complementarios restauraba la formación de colonias. La supervivencia sólo se daba cuando estaban presentes ambos fragmentos de DHFR y contenían secuencias formadoras de cremalleras de leucina, demostrando que la reconstitución de la actividad enzimática requiere la ayuda de la formación de cremalleras de leucina. Mutaciones puntuales en la superficie de contacto de los fragmentos de DHFR de gravedad creciente (de Ile a Val, Ala y Gly) dieron como resultado un aumento secuencial en los tiempos de duplicación de *E. coli*, ilustrando el reensamblaje con éxito de fragmentos de DHFR más que interacciones inespecíficas entre fragmentos. Este ensayo podía usarse para estudiar el equilibrio y los aspectos cinéticos de interacciones moleculares que incluyen interacciones de proteína-proteína, proteína-ADN, proteína-ARN, proteína-carbohidrato y proteína-molécula pequeña, para explorar genotecas de ADNc por la unión de una proteína diana con proteínas desconocidas o bibliotecas de moléculas orgánicas pequeñas para determinar su actividad biológica. Los criterios de selección y diseño aplicados en la presente memoria se desarrollan para numerosos ejemplos de ensayos de selección clonal, colorimétricos, fluorimétricos y de otro tipo basados en enzimas cuyos productos pueden medirse. Se demuestra que el desarrollo de dichos sistemas de ensayo es simple y proporciona un conjunto diverso de aplicaciones de complementación de fragmentos proteicos); WO 9839483 (Ventana, Alexander Kamb, se describen métodos para identificar secuencias de ácido nucleico que afectan a un fenotipo celular; el método usa un gen indicador cuyo nivel de expresión se correlaciona con el fenotipo junto con un método o dispositivo para medir el nivel de expresión de indicador); WO 9844350 (Helen Blau, ensayo de la complementación enzimática en el que se proporciona métodos y

composiciones para detectar interacciones moleculares, particularmente interacciones de proteína-proteína. La invención permite la detección de dichas interacciones en células vivas o *in vitro*. La detección de interacciones moleculares en células vivas no se limita al compartimento nuclear, sino que puede lograrse en el citoplasma, la superficie celular, los orgánulos o entre estas entidades. En una realización, el método utiliza nuevas composiciones que comprenden proteínas de fusión entre las moléculas de interés y dos o más mutantes de β -galactosidasa inactivos de débil complementación. La asociación entre las moléculas de interés aproxima los mutantes de β -galactosidasa de complementación de modo que se produce la complementación y se produce una β -galactosidasa activa. La β -galactosidasa activa puede detectarse por métodos bien conocidos en la técnica; Van Ostade et al., J. Interf. Cytok. Res. 20, 79-87 (2000) y los documentos WO00/06722, WO 01/90188 (Un bioensayo para ligandos que señalizan a través de agrupamientos de receptores, denominado MAPPIT. En concreto, la invención se refiere a un receptor recombinante que comprende un dominio de unión a ligando extracelular y un dominio citoplasmático que comprende un polipéptido cebo heterólogo, activándose dicho receptor por unión de un ligando a dicho dominio de unión a ligando y por unión de un polipéptido presa a dicho péptido cebo heterólogo. La invención también se refiere a un método para detectar la unión de compuesto-compuesto usando dicho receptor recombinante); documentos WO9418317, WO9613613, WO9941258 (Schreiber, métodos para inducir un acontecimiento biológico por dimerización inducida por compuesto) y Ghosh et al., J. Am. Chem. Soc., 2000, 122: 5658-9 (reconstitución de la fluorescencia a partir de una proteína verde fluorescente dividida).

[0211] También se han descrito sistemas para estudiar interacciones de proteína-proteína en células de mamífero. Por ejemplo, Fearon et al. (Karyoplasmic interaction selection strategy: A general strategy to detect protein-protein interactions in mammalian cells, Proc. Natl. Acad. Sci. USA 89: 7958-7962, 1992) describen una estrategia y reactivos para el estudio de interacciones de proteína-proteína en células de mamífero, denominada estrategia de selección de interacción carioplásmica (KISS). Con esta estrategia, se identifican interacciones de proteína-proteína específicas por reconstitución de la actividad funcional del activador de la transcripción de levadura GAL4 y la transcripción resultante de un gen indicador regulado por GAL4. La reconstitución de la función de GAL4 se obtiene como resultado de la interacción específica entre dos proteínas de fusión: una contiene el dominio de unión a ADN de GAL4; la otra contiene un dominio de activación de la transcripción. La transcripción del gen indicador se produce si las dos proteínas de fusión pueden formar un complejo que reconstituye las funciones de unión a ADN y activación de la transcripción de GAL4. Usando el sistema KISS, Fearon et al. demuestran interacciones específicas para secuencias de tres parejas de proteínas diferentes que forman complejos en el citoplasma. Además, demuestran que los genes indicadores que codifican marcadores de superficie celular o de resistencia a fármacos pueden activarse específicamente como resultado de interacciones de proteína-proteína. Con estos marcadores de selección, el sistema KISS puede usarse para explorar genotecas de ADNc especializadas para identificar nuevas interacciones proteicas.

[0212] Un experto en la materia será capaz de identificar los componentes del sistema de doble híbrido de levadura adecuados para usar con la presente invención sin una experimentación innecesaria. Estos incluirán, pero sin limitación, vectores de expresión para genes indicadores y sus métodos de ensayo/detección, vectores de expresión para la expresión de una proteína de fusión que comprende una proteína de unión a ADN y P1/P2, y vectores de expresión para la expresión de una proteína de fusión que comprende un dominio de activación de la transcripción y P1/P2. En ciertas realizaciones, P2 es de una biblioteca o bibliotecas de polipéptidos, de modo que el vector seleccionado para la expresión de la fusión de P2 será apropiado para la construcción de la biblioteca. Un experto en la materia deberá ser capaz de utilizar cualquiera de las tecnologías/métodos descritos anteriormente, o una combinación de los mismos, para poner en práctica la presente invención.

3.3 Genes indicadores

[0213] En un sistema indicador basado en la activación transcripcional de un gen indicador, se tiene que seleccionar un gen indicador apropiado para el tipo de célula hospedadora y formato de ensayo previsto. La célula hospedadora de elección tiene que proporcionar la maquinaria transcripcional apropiada, la selección del gen indicador dependerá del método seleccionado para detectar y potencialmente cuantificar la transcripción del gen indicador, por ejemplo, por transferencia de Western, métodos colorimétricos o fluorimétricos o un ensayo de inhibición del crecimiento en medios selectivos o contraselectivos, o un marcador de superficie celular.

[0214] El experto en la materia conocerá una amplia variedad de genes indicadores adecuados para usar en los métodos de la presente invención, y será fácilmente capaz de seleccionar el gen indicador apropiado para un formato de ensayo dado. Dicho gen indicador puede ser un gen marcador de selección positiva que puede seleccionarse en condiciones apropiadas. En principio, cualquier gen no redundante en una ruta sintética que sea esencial para la supervivencia de la célula puede usarse para la construcción de un marcador de selección positiva auxotrófico, pero dichos marcadores frecuentemente usados incluyen, sin limitación, HIS3, LYS2, LEU2, TRP2, ADE2. Habitualmente, se construye una línea celular que es deficiente en el gen marcador y que sólo puede crecer en medios complementados con el producto metabólico correspondiente, es decir, histidina, lisina, leucina, triptófano o adenina. Cuando se usa para selección, un fenotipo deseable, es decir, la expresión de un gen recombinante deseado, está vinculado a la expresión del gen en el que es deficiente la célula. Otros marcadores de selección positiva incluyen marcadores de resistencia a antibióticos, por ejemplo, resistencia a higromicina (Hyg^R), resistencia a neomicina (neo^R), resistencia a puomicina (PAC^R) o resistencia a blastidina S (Blas^R), o cualquier otro marcador de resistencia a antibióticos. Aquí, la expresión de un gen recombinante deseado está vinculada a la expresión del marcador de resistencia a antibiótico transformando células con construcciones génicas que comprenden tanto el gen recombinante deseado como una forma recombinante

del gen marcador de resistencia a antibiótico. Después, la selección se lleva a cabo en medios que contienen el antibiótico, por ejemplo, higromicina, neomicina, puromicina o blastidina S.

[0215] Además, el gen indicador puede codificar una proteína detectable que, tras la activación transcripcional de dicho gen indicador, permita que las células hospedadoras se diferencien visualmente de las células hospedadoras en las que no se ha activado dicho gen indicador. Dicha proteína detectable está preferiblemente codificada por al menos uno de los genes *lacZ*, *gfp*, *yfp*, *bfp*, *cat*, *luxAB*, *HPRT* o un gen marcador de superficie celular. Existen otros genes similares y el experto en la materia identificará fácilmente otros genes de este tipo que pueden emplearse de acuerdo con esta realización.

[0216] El documento WO 9825947 describe un sistema de ensayo de doble híbrido procariota que también proporciona detalles acerca de genes indicadores bacterianos que pueden usarse con la presente invención. Los marcadores de selección para usar en células bacterianas incluyen marcadores de resistencia a antibióticos, por ejemplo, *bla* (gen de resistencia a beta-lactamasa), *cam* (gen de cloranfenicol acetil transferasa) o *kan* (gen de kanamicina fosforil transferasa), marcadores de luminiscencia tales como *gfp*, marcadores de inducción de color, por ejemplo *lacZ*, marcadores auxotróficos (cualquier gen de biosíntesis de aminoácidos) y marcadores de resistencia a metales pesados. Pueden encontrarse marcadores de selección adicionales en: *Escherichia coli and Salmonella: Cellular and Molecular biology*, Segunda edición, F. C. Neidhart, et al. (Eds.), 1996. ASM Press, Washington, DC, Estados Unidos.

[0217] Además, pueden emplearse genes indicadores de selección negativa que pueden usarse en una célula y en contra de los cuales puede seleccionarse en las condiciones apropiadas. En aplicaciones preferidas, el indicador es un marcador de selección que es capaz de selección tanto positiva como negativa. Por ejemplo, el gen indicador puede seleccionarse de la lista de URA3, HIS3, LYS2, HygTk, Tkneo, TkBSD, PACTk, HygCoda, Codaneo, CodaBSD, PACCoda, Tk, codA y GPT2. El resto indicador también puede ser TRP1, CYH2, CAN1, HPRT.

[0218] Un ejemplo preferido de un gen marcador de selección negativa para usar en levadura es el gen URA3 que puede tanto seleccionarse a favor (selección positiva) cultivando cepas de levadura auxótrofas *ura3* en ausencia de uracilo, como seleccionarse en contra (selección negativa) cultivando células en medios que contienen ácido 5-fluoroorótico (5-FOA) (Boeke, et al, 1987, Methods Enzymol 154: 164-75). La concentración de 5-FOA puede optimizarse por valoración para seleccionar lo máximo células en las que el indicador URA3 esté inactivado por degradación proteolítica en cierta medida preferida. Por ejemplo, pueden usarse concentraciones relativamente elevadas de 5-FOA que permitan que sobrevivan sólo las células que expresen niveles en estado estacionario muy reducidos de indicador URA3. Por el contrario, pueden usarse menores concentraciones de 5-FOA para seleccionar compañeros de unión con afinidades relativamente débiles entre sí. Además, puede usarse prolina en los medios como fuente de nitrógeno para hacer a las células hipersensibles a los efectos tóxicos del 5-FOA (McCusker y Davis (1991), Yeast 7: 607-8). Por consiguiente, las concentraciones de prolina, así como las concentraciones de 5-FOA, pueden valorarse para obtener una selección óptima para células deficientes en indicador URA3. Por lo tanto, el uso de URA3 como marcador de selección negativa permite una amplia variedad de rigurosidades selectivas que pueden adaptarse para minimizar las interferencias de fondo de falsos positivos y/o para optimizar la selección de interacciones de unión de alta afinidad. Otros marcadores de selección negativa que pueden adaptarse a los métodos de la invención se incluyen en el alcance de la invención.

[0219] Otro ejemplo de un gen marcador de selección negativa para usar en levaduras es el gen de TRP1 que puede tanto seleccionarse a favor (selección positiva) por cultivo de cepas de levadura auxótrofas para *trp1* en ausencia de triptófano, como seleccionarse en contra (selección negativa) por cultivo de células en medios que contienen ácido 5-fluoroantránico (5-FAA) (Toyn et al, 2000, Yeast, 16: 553-560).

[0220] Otros dos genes marcadores de selección negativa para el uso en levaduras son CYH2 y CAN1, pudiendo ambos seleccionarse en contra (selección negativa) cultivando las células en medios que contienen cicloheximida o canavanina (*The Yeast Two-Hybrid System (Advances in Molecular Biology)*, Ed. Paul. L. Bartel y Stanley Fields, Oxford University Press, 1997).

[0221] Los marcadores de selección en contra para usar en bacterias incluyen *sacB* (gen de *B. subtilis* que codifica la levanosacarasa, que convierte la sacarosa en levanos, que es perjudicial para las bacterias), *rpsL* (*strA*) (codifica la diana proteica de subunidad ribosómica (S12) de estreptomycin), *tetA^R* (confiere resistencia a tetraciclina pero sensibilidad a compuestos lipófilos, por ejemplo, ácidos fusárico y quinálico), *phe^S* (codifica las subunidades de Phe-ARNt sintetasa, que hace a las bacterias sensibles a p-clorofenilalanina, un análogo de fenilalanina), *thyA* codifica la timidilato sintasa, que confiere sensibilidad a trimetoprim y compuestos relacionados, *lacY* (codifica la lactosa permeasa, que hace a las bacterias sensibles a t-o-nitrofenil-D-galactopiranosido), *gata-1* (codifica una proteína de unión a ADN de dedos de zinc que inhibe el inicio de la replicación bacteriana), *ccdB* (codifica una proteína de destrucción celular que es un potente veneno de girasa bacteriana). Pueden encontrarse marcadores de selección en contra adicionales en: *Escherichia coli and Salmonella: Cellular and molecular biology, Segunda edición*, F. C. Neidhardt, et al. (Eds.), 1996. ASM Press, Washington, DC, Estados Unidos.

[0222] Se conocen en la técnica numerosos marcadores de selección que funcionan en células de mamífero y pueden adaptarse al método de la invención para permitir la selección negativa directa de proteínas de interacción en células de mamífero. Los ejemplos de marcadores de selección negativa de mamífero incluyen timidina quinasa (Tk) (Wigler et al, 1977, Cell 11: 223-32; Borrelli et al, 1988, Proc. Natl. Acad. Sci. USA. 85: 7572-76) del virus herpes simple, el gen

humano para la hipoxantina fosforribosil transferasa (HPRT) (Lester et al, 1980, Somatic Cell Genet. 6: 241-59; Albertini et al, 1985, Nature 316: 369-71) y la citidina desaminasa (codA) de *E. coli* (Mullen et al, 1992, Proc. Natl. Acad. Sci. USA. 89: 33-37; Wei y Huber, 1996, J. Biol Chem. 271: 3812-16). Por ejemplo: el gen de Tk puede seleccionarse en contra usando ganciclovir (GANC) (por ejemplo, usando una concentración 1 μ M) y el gen *codA* puede seleccionarse en contra usando 5-fluorocitidina (5-FIC) (por ejemplo, usando una concentración de 0,1-1,0 mg/ml). Además, se han descrito ciertos marcadores de selección quiméricos (Karreman, 1998, Gene 218: 57-61) en los que un marcador de selección negativa de mamífero funcional se fusiona a un marcador de selección positiva de mamífero funcional, tal como resistencia a higromicina (Hyg^R), resistencia a neomicina (neo^R), resistencia a puromicina (PAC^R) o resistencia a blasticidina S (Bla^R). Estos producen diversos marcadores de selección positiva/negativa basados en Tk para células de mamífero, tales como HygTk, Tkneo, TkBSD y PACTk, así como diversos marcadores de selección positiva/negativa basados en *codA* para células de mamífero tales como HygCoda, Codaneo, CodaBSD y PACCoda. También se han descrito recientemente indicadores de Tk-neo que incorporan luciferasa, proteína verde fluorescente y/o beta-galactosidasa (Strathdee et al, 2000, BioTechniques 28: 210-14). Estos vectores tienen la ventaja de permitir la exploración rápida del marcador/indicador "positivo" por microscopía fluorescente y/o inmunofluorescente. El uso de dichos marcadores de selección positiva/negativa proporciona las ventajas mencionadas anteriormente para URA3 como indicador en levadura, en la medida en que permite que las células de mamífero se evalúen por métodos de selección tanto positiva como negativa para la expresión y un nivel en estado estacionario relativo de la fusión con indicador. Por ejemplo, Rojo-Niersbach et al describieron el uso de GPT2 (guanina fosforil transferasa 2) en células de mamífero como base para la selección de interacciones de proteínas (Biochem. J. 348: 585-590, 2000).

[0223] La lista anterior de genes adecuados para usar como genes indicadores en los métodos de la presente invención no pretende ser exhaustiva ni limitante. El experto puede conocer otros, o enterarse de sistemas recién descubiertos o desarrollados adecuados para usar como genes indicadores en los métodos de la presente invención. El alcance de la presente invención pretende incluir su uso.

3.4 El ensayo de crecimiento en halo

[0224] Puede usarse un ensayo de crecimiento en halo en varias realizaciones de la presente invención. En general, este tipo de ensayo proporciona la determinación cualitativa del efecto de diferentes concentraciones de un compuesto sobre el crecimiento celular. En esencia, un ensayo de crecimiento en halo comprende la distribución de una solución diluida de las células bajo investigación en una placa de agar, seguida de la colocación de una gota de una solución que contiene el compuesto bajo investigación en un punto predeterminado en el agar (por ejemplo, la mitad de una placa de petri). Posteriormente, la placa de agar se cultiva en condiciones que conducen a crecimiento celular, y el crecimiento se evalúa un tiempo después predeterminado. Durante este tiempo, el compuesto difundirá a través del agar, formando un gradiente de concentración con su mayor concentración en el punto de aplicación, disminuyendo radialmente hacia fuera de este punto. Si el agar está preparado para servir de soporte al crecimiento celular y el compuesto no tiene efecto, debería encontrarse una alfombra de células uniforme. Por el contrario, si el agar está preparado para suprimir el crecimiento celular, por ejemplo, agar que carece de un componente esencial para el crecimiento celular, y el compuesto no tiene efecto, no debería aparecer crecimiento celular. Si el compuesto tiene un efecto tóxico sobre las células, no deberían observarse cambios con el agar de supresión del crecimiento, pero en agar de mantenimiento del crecimiento debería aparecer un área circular (halo) sin crecimiento en el agar de mantenimiento del crecimiento alrededor del punto de aplicación, disminuyendo gradualmente hacia dentro el crecimiento hacia este punto. Cuando un compuesto tiene un efecto beneficioso sobre el crecimiento, tal como complementar la ausencia de un componente esencial en un agar de supresión del crecimiento, debería aparecer un halo circular de crecimiento alrededor del punto de aplicación, disminuyendo gradualmente hacia fuera el crecimiento desde este punto. Un experto en la materia estará familiarizado con dichos ensayos de halo. Sin embargo, pueden usarse igualmente métodos alternativos que satisfagan las mismas necesidades.

[0225] En ciertas realizaciones de la invención, puede ser ventajoso realizar un gran número de dichos ensayos para un solo experimento, preferiblemente más de aproximadamente 10, 100, 1000 ó más de 10000 ensayos. Puede contribuirse a dicho número de ensayos mediante el uso de placas de petri o de agar de aproximadamente 70, 300, 480 o más de 500 cm² de área superficial sobre las que se ponen las células y ligandos híbridos/compuestos de la invención. De hecho, para maximizar el rendimiento y minimizar el coste de realizar uno solo de dichos ensayos, es preferible reducir la escala del ensayo. Pueden realizarse ensayos minimizados, por ejemplo, usando una placa de microtitulación de preferiblemente 96, 384, 1536 o más de 1536 pocillos. Como alternativa, dichos ensayos pueden realizarse en agar de cultivo sólido en el que las células y los ligandos híbridos/compuestos se ponen en grandes cantidades o densidades. Por ejemplo, pueden realizarse aproximadamente 10, 100, 1000 o más de 10000 ensayos separados en una o más placas de petri o de agar, en las que un ensayo particular se separa de otro ensayo por una distancia de aproximadamente 1, 3, 10 o más de 30 mm. En ciertas realizaciones, es ventajoso que los ensayos se pongan en un patrón regular de modo que pueda realizarse más fácilmente el análisis del crecimiento posterior por visualización a simple vista o mediante máquina. Dichas cantidades, densidades o patrones de ensayos pueden formarse por varios métodos, como será evidente para un experto en la materia. Por ejemplo, pueden usarse pipetas multicanal de 8, 12 ó 16 vías o replicadores de 96/384 pocillos (Genetix). Como alternativa, si se desea un alto rendimiento o precisión, puede emplearse un dispositivo automático. El experto en la materia conocerá muchos dispositivos automáticos adecuados y se incluyen, sin limitación, unidades de pipeteo automático con 1, 2, 4, 8, 12, 96 o más de 96 puntas de pipeteo tal como se comercializan por varios fabricantes incluyendo el MultiProbe II o MultiTrack (Packard), Hamillton, Quadra 96 o 384 (Tomtec), CyBio etc. También pueden emplearse otros dispositivos automáticos

que transfieran con precisión un gran número de pequeñas cantidades de materiales biológicamente activos. Por ejemplo, pueden emplearse robots de cuadrícula tales como el Qbot (Genetix, Reino Unido), BioGrid (BioRobotics, Reino Unido) o los descritos en Maier et al 1997 (en Automation for genome characterisation. Ed TJ Beueldijk. J Wiley Nueva York).

5 3.5 El ensayo de crecimiento por detección de la fluorescencia

[0226] Un ensayo de crecimiento que puede realizarse en un formato de placa de microtitulación es ventajoso. Por ejemplo, pueden manipularse fácilmente MTP en grandes cantidades, usarse relativamente poco material por ensayo y, por lo tanto, pueden realizarse un gran número de ensayos usando automatización de laboratorio convencional. Los inventores desarrollaron un ensayo tal basado en el principio de que las células que crecen en suspensión consumen oxígeno del medio circundante. Sin embargo, el uso de este principio no pretende limitar el alcance de la invención, ya que el experto en la materia podrá apreciar otros métodos de evaluación del crecimiento de células en placas de microtitulación.

[0227] Con un detector de oxígeno integrado incorporado en la parte inferior de la placa, la OxoPlate (PreSens Precision Sensing GmbH, Regensburg, Alemania) es capaz de medir la concentración de oxígeno en la solución en cada pocillo de una placa de 96 pocillos a tiempo casi real (tiempo de respuesta <30 s). La medición se basa en la emisión de fluorescencia de dos colorantes en un detector en la parte inferior de cada pocillo, uno de los cuales puede inactivarse por oxígeno, mientras que la fluorescencia del segundo colorante no se ve afectada por el oxígeno y se usa como referencia interna. Ambos colorantes tienen una excitación igual (540 nm), pero diferentes desplazamientos de Stokes y longitudes de onda de emisión (colorante inactivable: 590 nm, colorante no inactivable: 650 nm). La relación de las emisiones a 650 nm y 590 nm ($I_{\text{inactivable}}/I_{\text{no inactivable}}$) se toma como medida de la concentración de oxígeno. Cuando se reduce la presión parcial de oxígeno en la solución en el pocillo, aumentará la intensidad de emisión del colorante que puede inactivarse por oxígeno, mientras que la intensidad de emisión del segundo colorante se mantendrá constante. El uso de dicha referencia interna hace a este ensayo independiente de muchas fuentes de error potenciales, tales como la inestabilidad del sistema óptico. También evita la necesidad de pocillos de calibración separados y, por lo tanto, los 96 pocillos de una placa de 96 pocillos pueden usarse para muestras. Este método usa un lector de placas que puede leer desde la parte inferior de una placa de microtitulación, y que puede medir en modo de cinética doble, es decir, tomando varias mediciones a dos longitudes de onda diferentes. Un experto en la materia conocerá bien lectores adecuados e incluyen, sin limitación, el contador HTS multimarcador Wallac Victor2 V 1420 de Perkin Elmer (Perkin Elmer, Wellesley, MA, Estados Unidos).

[0228] Cuando se siembran células adecuadas en los pocillos de una OxoPlate en un medio conducente a crecimiento, se producirá crecimiento celular logarítmico, se agotará el oxígeno y la presión parcial de oxígeno puede volverse limitante. A medida que disminuye adicionalmente el nivel de oxígeno, el crecimiento celular podría dificultarse, hasta que la presión parcial de oxígeno alcance casi cero, punto en el cual puede cesar el crecimiento celular. Este patrón de crecimiento se refleja en una curva sigmoidea de la relación de la intensidad de emisión de fluorescencia de los dos colorantes. Por el contrario, si el medio en un pocillo suprime el crecimiento, no se usará oxígeno y las mediciones de la relación de la intensidad de emisión de fluorescencia darán una línea constante próxima al valor para medio sin células.

4. Moléculas pequeñas híbridas

[0229] Se conocen en la técnica ensayos de triple híbrido de levadura que usan compuestos de ligando híbridos diferentes de los de la presente invención (Véase, por ejemplo: Crabtree et al. documento WO 94/18317; Schreiber et al. documento WO 96/13613, Holt et al. documento WO 96/06097, Licitra y Liu, documento WO 97/41255; Bergmann et al. J. Steroid Biochem Molec. Biol. 1994, 49:139-52; Lin et al, J. Am. Chem. Soc. 2000.122: 4247-8). Sin embargo, los compuestos de ligando híbridos de acuerdo con la presente invención poseen propiedades ventajosas que los separan claramente de los descritos en la técnica anterior. Por ejemplo, Lin et al. usaron un metadibenzotioéster como enlazador entre R1 y R2, confiriendo rigidez, lipofiliidad y baja solubilidad en agua a su compuesto de ligando híbrido Mtx-mdbt-Dex. Para atravesar membranas celulares, es deseable cierta lipofiliidad. Sin embargo, para llegar a la membrana, dicho compuesto tiene que atravesar primero un compartimento acuoso por difusión. Si su solubilidad en agua es demasiado baja, demasiado poco compuesto puede alcanzar la membrana y ejercer su efecto en el interior de la célula.

4. 1 Secuencias enlazadoras

[0230] En ciertas realizaciones, puede usarse cualquier enlazador químico Y (incluyendo polipéptidos sintéticos, véase a continuación) para unir R1 a R2, con tal de que la presencia de la secuencia enlazadora no interfiera significativamente con el sistema indicador cuando P1 se une a R1 y P2 se une a R2. Además, la presencia del enlazador no debería afectar demasiado perjudicialmente a las afinidades entre P1 y R1 o entre P2 y R2.

[0231] Como tal, para confirmar la idoneidad de un ligando híbrido dado como compuesto dimerizante de estructura general R1-Y-R2 para los usos propuestos en la presente memoria, puede ser útil caracterizar las propiedades de unión de dicho ligando híbrido a sus compañeros de unión P1 y P2 en la medida en que éstos se conozcan, y posiblemente comparar estas características de unión con las de los compuestos no relacionados R1 y R2, respectivamente. Preferiblemente, el ligando híbrido debería mostrar propiedades de unión similares a las propiedades de unión de los compuestos no relacionados. Sin embargo, el aumento de peso molecular provocado por la unión, así como los efectos electrónicos y estéricos causados por la unión del enlazador a un grupo funcional de los compuestos no relacionados, pueden alterar las características de unión. Por lo tanto, aunque no es esencial, es preferible realizar dicha

caracterización sobre un ligando híbrido recién sintetizado. Sin embargo, esto no debería interpretarse como limitante del alcance de la invención.

[0232] La afinidad de ligandos híbridos hacia sus compañeros de unión correspondientes puede determinarse, por ejemplo, usando un sistema de ensayo BIACORE™ (Biacore AB, Uppsala, SE). Otros sistemas que producen un resultado cualitativamente similar, por ejemplo, los desarrollados por Affinity Sensors (Cambridge, Reino Unido), serán fácilmente evidentes para los expertos en la materia. Además, pueden emplearse otras metodologías de interacción que miden las afinidades de unión entre un ligando híbrido y sus proteínas de unión.

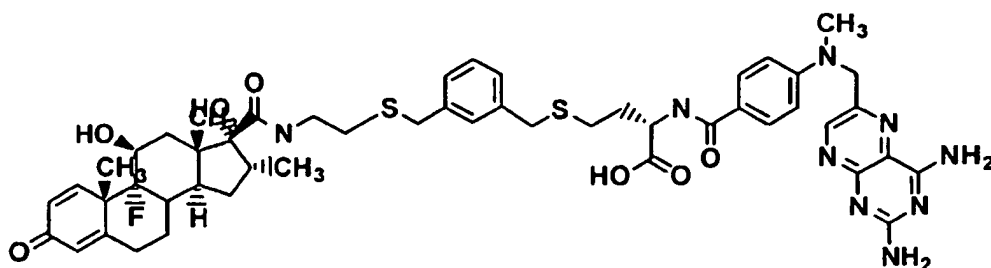
[02333] Los restos enlazadores (Y) no tienen que contener elementos esenciales para la unión a las proteínas P1 y/o P2, y para ciertas realizaciones de la presente invención pueden seleccionarse de un intervalo muy amplio de tipos estructurales. Los restos preferidos incluyen estructuras de alquilo, arilo o dialquilarilo C₂-C₂₀ en las que el alquilo y el 2,5 arilo están definidos como anteriormente. Los restos enlazadores pueden unirse convenientemente a monómeros R1 y R2 a través de grupos funcionales tales como éteres, amidas, ureas, carbamatos y ésteres; o a través de enlaces de carbono-carbono de alquilo-alquilo, alquilo-arilo o arilo-arilo. Además, los restos enlazadores pueden optimizarse (por ejemplo, por modificación de la longitud de cadena y/o de sustituyentes) para mejorar las propiedades farmacocinéticas del agente multimerizante. Holt et al. (documento WO 96/06097) y Kathryn et al. (J. Steroid Biochem. Molec Biol., 49:139-152) describen varios restos enlazadores que pueden usarse para construir los ligandos híbridos de la presente invención (R1-Y-R2).

[0234] En otras realizaciones, las secuencias enlazadoras están específicamente diseñadas para que se obtenga como resultado una solubilidad aumentada y una permeabilidad mejorada. Esto es importante ya que los componentes de la molécula híbrida, R1 y R2, son moléculas orgánicas con una solubilidad en agua potencialmente reducida. Por unión de dos moléculas pequeñas, el peso molecular obviamente se aumenta, disminuyendo potencialmente además la solubilidad en agua y el coeficiente de difusión. Por diseño de un enlazador que aumente la solubilidad y mejore la permeabilidad del híbrido, el híbrido R1-Y-R2 disponible en solución y, en última instancia, en el interior de la célula se aumenta eficazmente, de modo que puede conseguirse una sensibilidad significativamente superior de todo el sistema. En una realización, pueden usarse de 2 a 25 repeticiones de grupos de polietilenglicol (PEG) de fórmula general CH_2XCH_2 , en la que X representa O, S, SO o SO_2 . El número de repeticiones está preferiblemente en el intervalo de 3-25, 5-25, 9-25, 2-15, 3-15, 5-15 ó 9-15 y, más específicamente, es preferiblemente de 25, 20, 15, 10, 9, 8, 7, 6, 5, 4 ó 2. En una realización más preferida, se usan tres grupos de polietilenglicol como enlazador, que ofrecen una solubilidad y permeabilidad de membrana significativamente mejores (véase el Ejemplo 7 y el GPC 285937 a continuación). En otros casos en los que se desea una solubilidad y/o permeabilidad de membrana aún más claramente mejoradas, pueden usarse cinco repeticiones. Además, debería entenderse que pueden conseguirse fácilmente modificaciones de las cadenas laterales del enlazador sin que afecten perjudicialmente a la solubilidad, permeabilidad de membrana y/o actividad biológica global del compuesto y, por lo tanto, dichas unidades de secuencia enlazadora derivadas también se incluyen en el alcance de la invención.

[0235] A continuación se presentan varios ejemplos para moléculas híbridas como se prevén en la presente invención. Grupos $(CH_2XCH_2)_n$, en los que X representa O, $n = 3 \text{ ó } 5$, se emplearon para estos ejemplos sin limitación. El aumento de la longitud de la secuencia enlazadora parece aumentar la eficacia del compuesto en al menos algunos ensayos de triple híbrido, lo que se debe muy probablemente a una solubilidad o permeabilidad de membrana o flexibilidad de la molécula mejoradas, o a una combinación de las mismas. Por ejemplo, se predice que el coeficiente de reparto de n-octanol-agua (clogP) del compuesto Mtx-mdbt-Dex por cálculos basados en la estructura usando el programa Kowwin (Syracuse Research Corporation) es de 3,62, y su solubilidad en agua se sitúa en el intervalo de 0,00035 mg/l, mientras que se estima por el mismo método que el clogP para GPC 285937, idéntico al Mtx-mdbt-Dex excepto por el enlazador sustituido, es de -1,71, y su solubilidad como de 0,13 mg/l, que corresponde a un factor de solubilidad aumentada de aproximadamente 300.

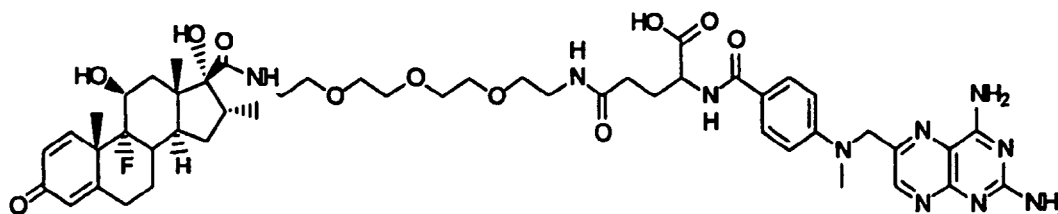
Estructura de Mtx-mdbt-Dex (R1 = metotrexato, R2 = dexametasona, Y = metadibenzotioéster)

[0236]



Estructura de GPC 285937 (R1 = metotrexato, R2 = dexametasona, Y = (CH₂-CH₂-O)₃)

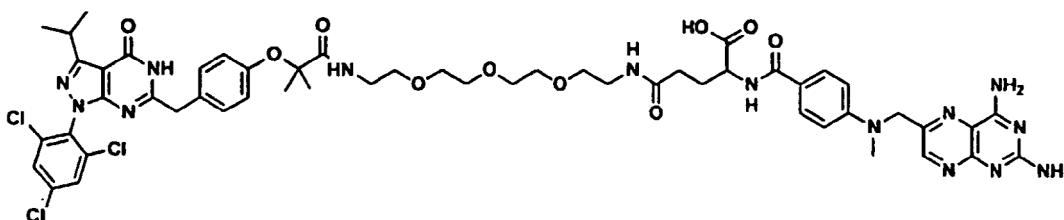
[0237]



[0238] ácido 4-(N-{2-[2-(2-{2-[(2S,11S,15S,17S,1R,13R,14R)-1-fluoro-14,17-dihidroxi-2,13,15-trimetil-5-oxotetraciclo[8.7.0.0<2,7>.0<11,15>]heptadeca-3,6-dien-14-il)carbonilamino]etoxi}etoxi]etil}carbamoil)-2-[(4-{[(2,4-diaminopteridin-6-il)metil]metilamino}fenil)carbonilamino]butanoico.

Estructura de GPC 285985 (R1 = metotrexato, Y = (CH₂-CH₂-O)₃, R2 es un inhibidor de CDK2 activo)

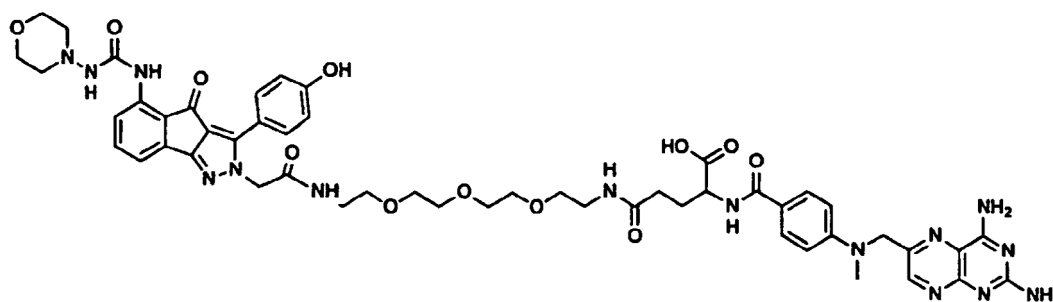
[0239]



[0240] ácido 2-[(4-{[(2,4-diaminopteridin-6-il)-metil]metilamino}fenil)carbonilamino]-4-(N-{2-[2-(2-{2-metil-2-(4-{[(3-(metiletil)-4-oxo-1-(2,4,6-triclorofenil)(5-hidropirazol[5,4-d]pirimidin-6-il)]metil}fenoxi)propanoilamino]etoxi}etoxi]etil}carbamoil)butanoico.

Estructura de GPC 285993 (R1 = metotrexato, Y = (CH₂-CH₂-O)₃, R2 es inactivo como inhibidor de CDK2)

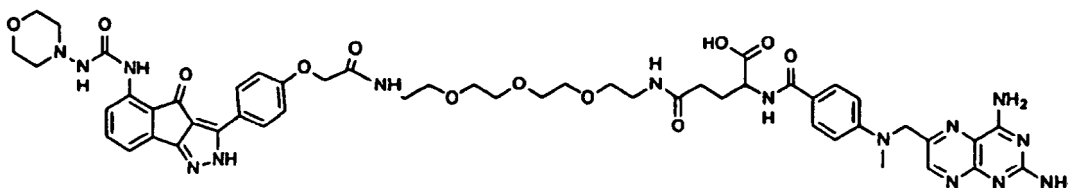
[0241]



[0242] ácido 2-[(4-{[(2,4-diaminopteridin-6-il)metil]metilamino}fenil)carbonilamino]-4-{N-[2-(2-{2-[2-(2-{3-(4-hidroxifenil)-5-[(morfolin-4-ilamino)carbonilamino]-4-oxoindeno[3,2-c]pirazol-2-il]acetilamino]etoxi}etoxi]etil}carbamoil)butanoico.

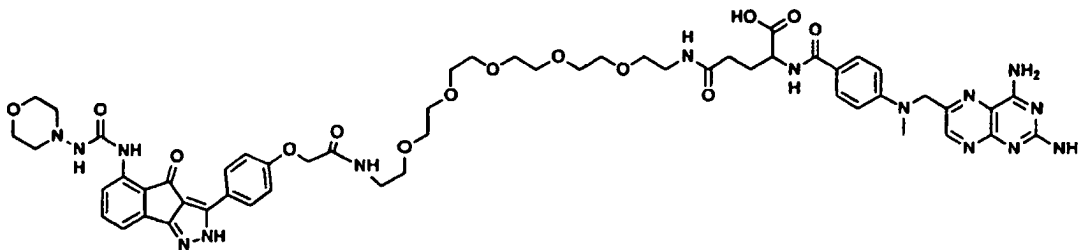
Estructura de GPC 286004 (R1 = metotrexato, Y = (CH₂-CH₂-O)₃, R2 es un inhibidor de CDK2 activo)

[0243]



[0244] ácido 2-[(4-{[(2,4-diaminopteridin-6-il)metil]metilamino}fenil)carbonilamino]-4-(N-{2-[2-(2-{2-[2-(4-{5-[(N-morfolin-4-ilcarbamoil)amino]-4-oxoindeno[3,2-c]pirazol-3-il}fenoxi)acetilamino]etoxi}etoxi]etil}carbamoil)butanoico.

Estructura de GPC 286026 (R1 = metotrexato, Y = (CH₂-CH₂-O)₅, R2 es un inhibidor de CDK2 activo) [0245]



[0246] ácido 2-[(4-[[[(2,4-diaminopteridin-6-il)metil]metilamino]fenil]carbonilamino]-4-{N-[2-(2-{2-[2-(2-{2-[4-{5-[(N-morfolin-4-ilcarbamoil)amino]-4-oxoindeno[3,2-c]pirazol-3-il}fenoxi)acetilamino]etoxi]etoxi]etoxi]etoxi]etoxi]etil]carbamoil}butanoico.

[0247] En una realización preferida, se emplea para la exploración más de una molécula pequeña híbrida, en la que R1 y/o R2 se unen por la misma secuencia enlazadora pero usando grupos de reacción diferentes, de un modo tal que pueda ajustarse la orientación relativa de R1 y R2. Esto es útil en la optimización de un ligando de compuesto eficaz puesto que ciertas orientaciones pueden superar o al menos aliviar impedimentos estéricos potenciales que sirven para debilitar la interacción entre el ligando y su compañero de unión proteico.

[0248] Las estructuras de las moléculas pequeñas híbridas mostradas anteriormente no deben entenderse de ningún modo como limitantes del alcance de la presente invención.

4. 2. Ligandos /proteínas de unión a ligando de alta afinidad

[0249] De acuerdo con la invención, deben estar presentes dos pares de interacciones de polipéptido/compuesto químico pequeño para que el sistema de triple híbrido active un sistema indicador. Un par de interacciones es entre un ligando conocido y su compañero de unión polipeptídico conocido. Esto sirve esencialmente como "adaptador" para crear una superficie de contacto de interacción R2::P2 y para proporcionar el segundo elemento necesario del sistema indicador, RS2. Por lo tanto, cuanto más fuerte sea la interacción P1::R1, mejor será el rendimiento global del sistema.

[0250] Existen al menos dos categorías de interacciones P1::R1 disponibles con este fin: interacciones covalentes y no covalentes. Las interacciones covalentes son casi siempre más fuertes. Por ejemplo, ciertas enzimas y sus inhibidores suicidas o sustratos suicidas pueden aprovecharse para constituir dichas parejas de interacción covalentes. Los inhibidores suicidas o sustratos suicidas se unen a sus enzimas posibles con alta especificidad y afinidad. Una vez unidos, se produce una reacción química que une físicamente el inhibidor/sustrato a la enzima, habitualmente en su sitio activo, inactivando de este modo irreversiblemente la enzima. Si se usa dicha enzima como P1 y su inhibidor/sustrato suicida se usa como R1 en el sistema de triple híbrido, puede establecerse un enlace covalente entre P1-R1. Por ejemplo, la beta-lactamasa puede unirse covalentemente a inhibidores suicidas tales como antibióticos beta-lactámicos. Sin embargo, sólo existen selecciones limitadas de estas parejas de enzima-sustrato/inhibidor, particularmente cuando el sustrato/inhibidor tiene que conectarse con otro compuesto pequeño R2 por medio de un enlazador conservando aún además la solubilidad y permeabilidad de membrana *in vivo*.

[0251] Por otro lado, las interacciones P1::R1 no covalentes son más versátiles. Hay muchas interacciones de ligando-receptor de alta afinidad conocidas que pueden emplearse en el sistema de triple híbrido. Por ejemplo, FK506 y FKBP (proteína de unión a FK506), FK506 y rapamicina, biotina y estreptavidina, DHFR y metotrexato (Mtx), receptor de glucocorticoides y dexametasona (Dex), etc., representan parejas de unión con afinidades lo bastante elevadas para ser potencialmente adecuados como parejas de unión de ligando-receptor. La interacción de DHFR-Mtx ofrece una afinidad pM y, por lo tanto, es mucho mejor que la interacción de FK506-FKBP.

[0252] Puede utilizarse cualquiera de varias parejas de ligando/proteína de unión a ligando conocidas en la técnica. Por ejemplo, puede emplearse la molécula esteroidea dexametasona que se une al receptor de glucocorticoides con gran afinidad. La dexametasona es modular por naturaleza; puede unirse covalentemente a otra molécula pequeña tal como la biotina sin perder su afinidad por el receptor de glucocorticoides. El uso de esteroides tales como dexametasona es ventajoso en el sentido de que estas moléculas presentan una alta permeabilidad de membrana y son de pequeño tamaño. El método de la invención puede utilizar otras moléculas esteroideas, así como pequeñas moléculas distintas de esteroides, como ligando R1. Otros ligandos tales como ciclosporina (PM 1200) también pueden usarse cuando la diana o receptor al que se une el ligando se ha identificado en la técnica. Como otro ejemplo, la pequeña molécula FK506 (PM 850) que se une a una proteína de unión a FK (FKBP) y derivados modificados de FK506 (es decir, compuestos modificados de "golpe") que se unen a proteínas de unión a FK modificadas (es decir, mutantes de FKBP que compensan dichas modificaciones de "golpe") también pueden adaptarse para usarse como ligandos/proteínas de unión a ligando de la invención (véase, por ejemplo, la Patente de Estados Unidos N° 6.054.436).

[0253] La Tabla 1 proporciona una lista de ligandos y parejas de unión a ligando que se conocen en la técnica y pueden adaptarse a las composiciones y métodos de la invención. Las parejas de ligando/proteína de unión a ligando particularmente preferidas tienen afinidades de unión fuertes como se refleja en constantes de disociación reducidas

(por ejemplo, metotrexato/DHFR a 52 pM; o dexametasona/receptor de glucocorticoides a 86 nM).

Tabla 1. Lista de algunos ligandos/proteínas de unión a ligando de alta afinidad

Ligando	Peso molecular (D)	Proteína de unión a ligando	Afinidad
Biotina	(244)	Avidina	80 fM
Ni	(59)	6X His	0,8 μ M
Rapamicina	(914)	FKB12	12 μ M
FK506	(804)	FKB12	12 μ M
Metotrexato	(454)	DHFR	52 pM
Tetraciclina	(444)	Tet-R	24 nM
Dexametasona	(392)	Receptor de glucocorticoides	86 nM
Glutación	(307)	Glutación-S-Transferasa	24 μ M
Maltosa	(342)	Proteína de unión a maltosa	40 nM
Novobiotina	(612)	GyrB	123 μ M

[0254] En general, prácticamente cualquier pareja de ligando/proteína de unión a ligando con una afinidad suficiente puede adaptarse a las composiciones y métodos de la invención. Las realizaciones particularmente preferidas utilizan proteínas de unión a ligando que se sabe que funcionan eficazmente intracelularmente. Por ejemplo, los receptores de esteroides aparecen intracelularmente y se unen con altas afinidades a sus hormonas esteroideas afines en condiciones fisiológicas intracelulares. Los ejemplos de dichos receptores de esteroides incluyen el receptor de estrógenos humano (por ejemplo, N° de Acceso GenBank NM_000125), que se encuentra en células animales sensibles a estrógenos, y la proteína receptora de glucocorticoides humana (por ejemplo, N° de Acceso GenBank NM_004491), que se encuentra en células sensibles a hormonas glucocorticoides. Otros esteroides con receptores adecuados para usar en la invención incluyen testosterona, progesterona y cortisona.

[0255] Debería entenderse que los ligandos mencionados anteriormente también incluir aquellos derivados y equivalentes que compartan una relación estructural estrecha con esos ligandos. Para ilustrarlo, el Mtx sólo usa su estructura de doble anillo de 2,4-diaminopteridina para unirse a DHFR. Por lo tanto, la 2,4-diaminopteridina se considerará un derivado de Mtx que también se incluye en el alcance de la invención. Un "derivado" generalmente comparte el resto eficaz con el compuesto original, pero también puede tener otros elementos estructurales no esenciales para una actividad dada.

[0256] Otros ligandos más preferidos para usar en la invención se conocen en la técnica y pueden adaptarse a los métodos y composiciones de la invención por expertos en la materia sin una experimentación innecesaria. Por ejemplo, otros ligandos preferidos que podrían adaptarse a la invención incluyen vitaminas liposolubles con receptores afines tales como vitamina D y sus diversas formas, tales como D₁, D₂ (9,10-secoergosta-5,7,10(19),22-tetraen-3-ol), D₃ (9,10-secoergosta-5,7,10(19)-trien-3-ol) y D₄ (9,10-secoergosta-5,7,10(19)-trien-3-ol). La vitamina D₃ se une con afinidad a la proteína receptora de vitamina D nuclear humana (por ejemplo, N° de Acceso GenBank NM_000376; véase también Haussler et al. (1995) Bone 17: 33S-38S) y su pareja de ligando/proteína de unión a ligando puede adaptarse a la invención. Otros ligandos más con proteínas de unión a ligando afines que pueden adaptarse a la invención incluyen hormona tiroidea y ácido retinoico. DeWolf y Brett ((2000) Pharmacol Rev. 52: 207-36) proporciona un resumen de muchas proteínas de unión a ligando útiles con ligandos afines que incluyen: proteínas de unión a biotina, proteína de unión a lípidos, proteínas de unión periplásmicas, lectinas, albúminas séricas, inmunoglobulinas, diversas enzimas inactivadas, proteínas de unión a feromonas de insectos, proteínas de unión a odorizantes, proteínas de unión a inmunosupresores, proteína de unión a fosfato y sulfato.

[0257] Además, también es posible usar en la presente invención como componente de la pareja de ligando/unión a ligando de alta afinidad descrita anteriormente esteroides, ácido retinoico, antibióticos beta-lactámicos, canabinoides, ácido nucleico, polipéptido, FK506, derivados de FK506, rapamicina, tetraciclina, metotrexato, 2,4-diaminopteridina, novobiocina, maltosa, glutación, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, ciclosporina y sus compañeros de unión naturales o sintetizados. En todos los compuestos mencionados anteriormente, debería entenderse que también pueden usarse compuestos básicamente equivalentes con sólo variaciones estructurales minoritarias.

[0258] Por otro lado, un segundo ligando especificado por el usuario tiene que estar ligado al ligando descrito anteriormente para formar un ligando compuesto. Al menos los siguientes grupos químicos y aquellos compuestos básicamente equivalentes con sólo variaciones estructurales minoritarias pueden usarse como dichos ligandos

especificados por el usuario: un péptido, un ácido nucleico, un carbohidrato, un polisacárido, un lípido, una prostaglandina, un haluro de acilo, un alcohol, un aldehído, un alcano, un alqueno, un alquino, un alquilo, un haluro de alquilo, un alcaloide, una amina, un hidrocarburo aromático, un éster de sulfonato, un carboxilato ácido, un haluro de arilo, un éster, un fenol, un éter, un nitrilo, un anhídrido del ácido carboxílico, una amida, una sal de amonio cuaternario, una imina, una enamina, un óxido de amina, una cianohidrina, un organocadmio, un aldol, un organometálico, un hidrocarburo aromático, un nucleósido, un nucleótido. Por ejemplo, en una publicación reciente (Patente de Estados Unidos N° 6.326.155), se describe un método que contribuye a seleccionar un ligando para una molécula diana dada.

5. Bibliotecas y métodos de exploración

5.1 Presentación abigarrada de péptidos

[0259] Un aspecto de la invención proporciona un método para identificar polipéptidos que se unen a una molécula pequeña/compuesto químico dado. Los polipéptidos se proporcionan habitualmente en forma de una biblioteca abigarrada, que puede contener diferente número de miembros, preferiblemente de 2 a 10 miembros, o de 10 a 500 miembros, de 500 a 10.000 miembros o más de 10.000 miembros. La biblioteca puede ser una biblioteca de ácidos nucleicos (ARNm, ADNc, ADN genómico, EST, YAC, clones p1, bibliotecas de BAC/PAC, etc.) que codifican polipéptidos. Dependiendo de las realizaciones específicas de las exploraciones usadas (por ejemplo, sistema híbrido basado en ubiquitina dividida o sistema híbrido de levadura basado en transcripción), la biblioteca de ácido nucleico se construye habitualmente en vectores adecuados para la realización seleccionada, usando procedimientos reconocidos en la técnica.

[0260] Las bibliotecas de péptidos abigarradas del presente método pueden generarse por cualquiera de varios métodos y, aunque no se limitan a, preferiblemente aprovechan tendencias recientes en la preparación de bibliotecas químicas. La biblioteca puede prepararse, por ejemplo, mediante estrategias sintéticas o biosintéticas. Como se usa en la presente memoria, el término "abigarrado" se refiere al hecho de que una población de péptidos está caracterizada por tener una secuencia peptídica que difiere de un miembro de la biblioteca al siguiente. Por ejemplo, en una biblioteca de péptidos dada de N aminoácidos de longitud, el número total de secuencias peptídicas diferentes en la biblioteca se proporciona por el producto de $(X_1 * X_2 * \dots * X_i)$, donde cada X_i representa el número de restos aminoácidos diferentes que aparecen en la posición X del péptido. En una realización preferida de la presente invención, la presentación de péptidos produce en conjunto una biblioteca de péptidos que incluye de al menos 96 a 10^7 péptidos diferentes, de modo que pueden ensayarse simultáneamente diversos péptidos para determinar la capacidad de interaccionar con la molécula pequeña/compuesto químico.

[0261] Las bibliotecas de polipéptidos pueden preexplorarse para determinar interacciones con la molécula pequeña/compuesto químico, por ejemplo, usando un método de presentación en fago. Las bibliotecas de péptidos son sistemas que presentan simultáneamente, en una forma que permite la interacción con una molécula diana, una colección muy diversa y numerosa de péptidos. Estos péptidos pueden presentarse en solución (Houghten (1992) *Biotechniques* 13: 412-421), o en perlas (Lam (1991) *Nature* 354: 82-84), microplacas (Fodor (1993) *Nature* 364: 555-556), bacterias (Ladner USSN 5.223.409), esporas (Ladner USSN 5.223.409), plásmidos (Cull et al. (1992) *Proc Natl Acad Sci USA* 89: 1865-1869) o en un fago (Scott y Smith (1990) *Science* 249: 386-390; Devlin (1990) *Science* 249: 404-406; Cwirla et al. (1990) *Proc. Natl. Acad. Sci.* 87: 6378-6382; Felici (1991) *J. Mol. Biol.* 222: 301-310; y Ladner USSN 5.223.409).

[0262] En una realización, la biblioteca de péptidos se deriva para expresar una biblioteca de péptidos combinatoria que no se basan en ninguna secuencia conocida, ni proceden de ADNc. Es decir, las secuencias de la biblioteca son en gran parte aleatorias. Será evidente que los péptidos de la biblioteca pueden variar en tamaño de dipéptidos a grandes proteínas.

[0263] En otra realización, la biblioteca de péptidos se deriva para expresar una biblioteca de péptidos combinatoria que se basan, al menos en parte, en una secuencia polipeptídica conocida o una porción de la misma (no una genoteca de ADNc). Es decir, las secuencias de la biblioteca son semialeatorias, obteniéndose por mutagénesis combinatoria de una secuencia o secuencias conocidas. Véase, por ejemplo, Ladner et al. Publicación PCT WO 90/02909; Garrard et al., Publicación PCT WO 92/09690; Marks et al. (1992) *J. Biol. Chem.* 267: 16007-16010; Griffiths et al. (1993) *EMBO J* 12:725-734; Clackson et al. (1991) *Nature* 352: 624-628; y Barbas et al. (1992) *PNAS* 89: 4457-4461. Por consiguiente, un polipéptido o polipéptidos que son ligandos conocidos para una molécula diana pueden mutarse por técnicas convencionales para obtener una biblioteca abigarrada de secuencias polipeptídicas que pueden explorarse adicionalmente para compañeros de unión, incluyendo agonistas y/o antagonistas.

[0264] En otra realización más, los polipéptidos combinatorios se producen a partir de una genoteca de ADNc, una genoteca de ADN genómico. La fuente de ADN puede ser de seres humanos, mamíferos no humanos, peces, anfibios, insectos, helmintos, levaduras, plantas o bacterias.

[0265] Dependiendo del tamaño, los péptidos combinatorios de la biblioteca pueden generarse como tales o pueden incorporarse en proteínas de fusión de mayor tamaño, tales como fusiones de biblioteca-sistema indicador. La proteína de fusión también puede proporcionar, por ejemplo, estabilidad frente a la degradación o desnaturalización, así como una señal de secreción si se secreta, o la función indicadora necesaria para las exploraciones. En una realización ejemplar, la biblioteca de polipéptidos se proporciona como parte de proteínas de fusión con tioredoxina (véanse, por ejemplo, las patentes de Estados Unidos 5.270.181 y 5.292.646; y la publicación PCT WO94/02502). El péptido

combinatorio puede unirse en el extremo terminal de la proteína tiorredoxina o, para bibliotecas de péptidos cortos, insertarse en el denominado bucle activo. En otra realización preferida, la biblioteca de proteínas de fusión puede proporcionarse como una fusión con el dominio Cub o Nux de las proteínas detectoras de ubiquitina dividida (véase a continuación). En otra realización preferida, la biblioteca de proteínas de fusión puede proporcionarse como una fusión con el dominio de unión a ADN o el dominio de activación de la transcripción basada en el sistema de triple híbrido de levadura.

[0266] En realizaciones preferidas, los péptidos combinatorios están en el intervalo de 3-1000 aminoácidos de longitud, más preferiblemente al menos 5-500 y, aún más preferiblemente, al menos 3-100, 5-50, 10, 13, 15, 20 ó 25 restos aminoácídicos de longitud. Preferiblemente, los polipéptidos de la biblioteca son de una longitud uniforme. Se entenderá que la longitud del péptido combinatorio no refleja ninguna secuencia extraña que pueda estar presente para facilitar la expresión, por ejemplo, tal como secuencias señal o porciones invariables de una proteína de fusión.

[0267] Independientemente de la naturaleza de las bibliotecas de péptidos, las mismas bibliotecas de péptidos también pueden proporcionarse como bibliotecas de ácido nucleico que codifican dichas bibliotecas de péptidos. Estas bibliotecas de ácido nucleico pueden proporcionarse en vectores adecuados para la expresión en diversos sistemas, incluyendo, pero sin limitación, sistemas de expresión de mamíferos, insectos, levaduras y bacterias. Un experto en la materia será capaz de determinar los vectores apropiados a usar para diversos sistemas de expresión.

5.1.1 Bibliotecas de péptidos biosintéticos

[0268] La utilización de sistemas biológicos para la generación de una diversidad de péptidos es ahora una técnica bien establecida que puede aprovecharse para generar las bibliotecas de péptidos del presente método. La fuente de diversidad es la síntesis química combinatoria de mezclas de oligonucleótidos. La síntesis de oligonucleótidos es una química bien caracterizada que permite un estrecho control de la composición de las mezclas creadas. Las secuencias de ADN degeneradas producidas se sitúan posteriormente en un contexto genético apropiado para su expresión como péptidos.

[0269] Hay dos formas principales en las que preparar la mezcla degenerada necesaria. En un método, los ADN se sintetizan de una base en una base. Cuando se desea una variación en una posición de base dictada por el código genético, se hace reaccionar una mezcla adecuada de nucleótidos con el ADN naciente, más que con el reactivo de nucleótidos puro de la síntesis de polinucleótidos convencional. El segundo método proporciona un control más exacto sobre la variación de aminoácidos. En primer lugar, se preparan reactivos de trinucleótidos, siendo cada trinucleótido un codón de uno (y solo uno) de los aminoácidos que se presentarán en la biblioteca de péptidos. Cuando se va a sintetizar un resto variable particular, se realiza una mezcla de los trinucleótidos apropiados y se hacen reaccionar con el ADN naciente. Una vez que se completa el ADN "degenerado" necesario, debe unirse con las secuencias de ADN necesarias para asegurar la expresión del péptido, como se analiza en más detalle a continuación, y la construcción de ADN completa debe introducirse en la célula.

[0270] Cualquiera que pueda ser el método para generar diversidad a nivel de codón, la síntesis química de una secuencia génica degenerada puede llevarse a cabo en un sintetizador de ADN automático, y los genes sintéticos pueden ligarse después en un gen o vector apropiado para su expresión. El propósito de un conjunto degenerado de genes es proporcionar, en una mezcla, todas las secuencias que codifican el conjunto deseado de secuencias peptídicas de ensayo potenciales. La síntesis de oligonucleótidos degenerados es bien conocida en la técnica (véase, por ejemplo, Narang, SA (1983) Tetrahedron 39: 3; Itakura et al. (1981) Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules, ed. AG Walton, Amsterdam: Elsevier págs. 273-289; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al. (1984) Science 198: 1056; Ike et al. (1983) Nucleic Acid Res. 11: 477. Dichas técnicas se han empleado en la evolución dirigida de otras proteínas (véase, por ejemplo, Scott et al. (1990) Science 249: 386-390; Roberts et al. (1992) PNAS 89: 2429-2433; Devlin et al. (1990) Science 249: 404-406; Cwirla et al. (1990) PNAS 87: 6378-6382; así como las patentes de Estados Unidos N° 5.223.409, 5.198.346 y 5.096.815).

[0271] Debido a que el número de péptidos diferentes que se pueden crear mediante esta estrategia de combinación puede ser enorme, y debido a que se espera que los péptidos con las características estructurales apropiadas para servir como ligandos para una proteína diana dada sean poco frecuentes en la población total de la biblioteca, es evidente la necesidad de métodos capaces de explorar convenientemente grandes cantidades de clones. Se han descrito en la técnica varias estrategias para seleccionar ligandos peptídicos de la biblioteca y son aplicables a ciertas realizaciones del presente método.

[0272] El número de péptidos posibles para una biblioteca dada puede superar en algunos casos 10^{12} . La obtención de muestras de tantas combinaciones como sea posible depende, en parte, de la capacidad para recuperar grandes cantidades de transformantes. Para fagos con formas tipo plásmido (como fagos filamentosos), la electrotransformación proporciona una eficacia comparable a la de la transfección de fagos con un empaquetamiento *in vitro*, además de una capacidad muy elevada de entrada de ADN. Esto permite que se usen grandes cantidades de ADN de vector para obtener un número muy grande de transformantes. El método descrito por Dower et al. (1988) Nucleic Acids Res., 16:6127-6145, por ejemplo, puede usarse para transformar recombinantes derivados de fd-tet a razón de aproximadamente 10^7 transformantes/ μ g de vector ligado en *E. coli* (tal como la cepa MC1061), y pueden construirse bibliotecas en fd-tet B1 de hasta aproximadamente 3×10^8 miembros o más. El aumento de la entrada de ADN y la realización de modificaciones en el protocolo de clonación dentro de la capacidad del experto en la materia puede

producir aumentos de más de aproximadamente 10 veces en la recuperación de transformantes, proporcionando bibliotecas de hasta 10^{10} o más recombinantes.

5.1.2 Bibliotecas de péptidos sintéticos

[0273] Al contrario que los métodos recombinantes, la síntesis química *in vitro* proporciona un método para generar bibliotecas de compuestos sin el uso de organismos vivos, que pueden explorarse para determinar su capacidad para unirse a una molécula de diana. Aunque se han usado métodos *in vitro* durante bastante tiempo en la industria farmacéutica para identificar fármacos potenciales, los métodos desarrollados recientemente se han centrado en generar y explorar rápida y eficazmente grandes cantidades de compuestos y son particularmente susceptibles de generar bibliotecas de péptidos para su uso en el presente método.

[0274] Unas características particularmente útiles de la biblioteca de péptidos sintéticos es que pueden usarse para suministrar bibliotecas de R2 que se acoplarán a R1-Y, para generar el ligando híbrido. Esto puede usarse para explorar para un polipéptido sintético que pueda unirse a un polipéptido especificado por el usuario. Por ejemplo, el polipéptido sintético puede ser un inhibidor peptídico potencial de un factor de transcripción o enzima especificada por el usuario, etc. Dichas exploraciones pueden ser una preexploración de un gran número de polipéptidos aleatorios en un entorno de alto rendimiento *in vitro*, de modo que puedan seleccionarse péptidos positivos primarios, y sus variantes codificadas por una biblioteca de ácidos nucleicos explorarse adicionalmente en una realización *in vivo*.

[0275] Otro uso para la biblioteca de péptidos sintéticos es generar bibliotecas de enlazadores peptídicos cortos que se insertarán entre los ligandos R1 y R2. Esto es particularmente útil puesto que puede generarse una secuencia enlazadora óptima para una pareja R1-R2 particular, de modo que el ligando híbrido final puede poseer las características estructurales y/o químicas óptimas, tales como solubilidad, permeabilidad de membrana, etc.

[0276] Ambos usos requieren acoplamiento de un polipéptido sintético, usando conocimientos bien conocidos en la técnica (tales como los descritos a continuación o en otra parte), a otra molécula (enlazador Y o ligandos R1 y R2) que puede ser de naturaleza peptídica o no peptídica.

[0277] Las diversas estrategias para preparación y análisis simultáneo de un gran número de péptidos sintéticos (en la presente memoria "síntesis de múltiples péptidos" o "MPS") dependen cada una del concepto fundamental de síntesis en un soporte sólido introducido por Merrifield in 1963 (Merrifield, R. B. (1963) J Am Chem Soc 85: 2149-2154; y referencias citadas en la sección I anterior). Generalmente, estas técnicas no dependen del grupo protector o de la química de activación empleada, aunque la mayoría de los trabajadores actualmente evitan la estrategia de tBoc/Bzl original de Merrifield a favor de una química de Fmoc/tBu más suave y agentes de acoplamiento basados en hidroxibenzotriazol eficaces. Se han usado con éxito muchos tipos de matrices sólidas en MPS y los rendimientos de péptidos individuales sintetizados varían ampliamente con la técnica adoptada (por ejemplo, de nanomoles a milimoles).

5.1.2.1 Síntesis multiclavija

[0278] Una forma que puede adoptar la biblioteca de péptidos del presente método es el formato de biblioteca multiclavija. En resumen, Geysen y colaboradores (Geysen et al. (1984) PNAS 81: 3998-4002) introdujeron un método para generar péptidos por síntesis en paralelo sobre clavijas de polietileno injertado con ácido poliacrílico dispuestas en el formato de placa de microtitulación. En los experimentos originales, aproximadamente 50 nmol de una sola secuencia peptídica se unieron covalentemente a la cabeza esférica de cada clavija, y las interacciones de cada péptido con receptor o anticuerpo podían determinarse en un ensayo de unión directa. La técnica de Geysen puede usarse para sintetizar y explorar miles de péptidos por semana usando el método de multiclavija, y los péptidos unidos pueden reutilizarse en muchos ensayos. En un trabajo posterior, el nivel de carga de péptido en clavijas individuales se ha aumentado hasta tanto como 2 μ mol/clavija injertando mayores cantidades de derivados de acrilato funcionalizados a cabezas de clavija desprendibles, y el tamaño de la biblioteca de péptidos se ha aumentado (Valerio et al. (1993) Int J Pept Protein Res 42: 1-9). También se han añadido restos enlazadores apropiados a las clavijas, de modo que los péptidos pueden escindirse de los soportes después de la síntesis para la valoración de la pureza y la evaluación en bioensayos funcionales o de unión competitiva (Bray et al. (1990) Tetrahedron Lett 31: 5811-5814; Valerio et al. (1991) Anal Biochem 197: 168-177; Bray et al. (1991) Tetrahedron Lett 32: 6163-6166).

[0279] Aplicaciones más recientes del método multiclavija de MPS se han aprovechado de la estrategia de enlazadores escindibles para preparar péptidos solubles (Maeji et al. (1990) J Immunol Methods 134: 23-33; Gammon et al. (1991) J Exp Med 173:609-617; Mutch et al. (1991) Pept Res 4:132-137).

5.1.2.2 Dividir-Acoplar-Recombinar

[0280] En otra realización más, una biblioteca de péptidos abigarrada puede proporcionarse en un conjunto de perlas que utilizan la estrategia de dividir-acoplar-recombinar (véase, por ejemplo, Houghten (1985) PNAS 82: 5131-5135; y las Patentes de Estados Unidos 4.631.211; 5.440.016; 5.480.971). En resumen, como implica el nombre, en cada etapa de síntesis en la que se introduce degeneración en la biblioteca, las perlas se dividen en tantos grupos separados como para que se correspondan con el número de restos aminoácidos diferentes a añadir en esa posición, los restos diferentes se acoplan en reacciones separadas y las perlas se recombinan en una combinación para la etapa siguiente.

[0281] En una realización, la estrategia de dividir-acoplar-recombinar puede llevarse a cabo usando el método de MPS denominado "bolsita de té" desarrollado por primera vez por Houghten, la síntesis peptídica se produce sobre resina que está sellada en el interior de bolsitas de polipropileno poroso (Houghten et al. (1986) PNAS 82: 5131-5135). Los aminoácidos se acoplan a las resinas poniendo las bolsas en soluciones de los monómeros activados individuales apropiados, mientras que todas las etapas comunes tales como el lavado de la resina y la desprotección de grupos amino se realizan simultáneamente en un recipiente de reacción. Al final de la síntesis, cada bolsita contiene una sola secuencia peptídica, y los péptidos pueden liberarse de las resinas usando un aparato de escisión múltiple (Houghten et al. (1986) Int J Pept Protein Res 27: 673-678). Esta técnica ofrece las ventajas de una flexibilidad de síntesis considerable y se ha automatizado parcialmente (Beck-Sickinger et al. (1991) Pept Res 4: 88-94). Además, pueden producirse péptidos solubles de más de 15 aminoácidos de longitud en cantidades suficientes (>0,5 mmol) para la purificación y caracterización completa si se desea.

[0282] La síntesis de múltiples péptidos usando la estrategia de bolsita de té es útil para la producción de una biblioteca de péptidos, aunque de tamaño limitado, para explorar el presente método, como se ilustra por su uso en una variedad de problemas de reconocimiento molecular que incluyen el análisis de epítomos de anticuerpos (Houghten et al. (1986) PNAS 82: 5131-5135), estudios de estructura-función de hormonas peptídicas (Beck-Sickinger et al. (1990) Int J Pept Protein Res 36: 522-530; Beck-Sickinger et al. (1990) Eur J Biochem 194: 449-456) y el mapeo conformacional de proteínas (Zimmerman et al. (1991) Eur J Biochem 200: 519-528).

[0283] Una síntesis ejemplar de un conjunto de péptidos mixtos que tienen cantidades equimolares de los veinte restos aminoácidos naturales es la siguiente. Se ponen alícuotas de cinco gramos (4,65 mmoles) de resina de clorhidrato de p-metilbenzidrilamina (MBHA) en veinte bolsitas de polipropileno poroso. Estas bolsitas se ponen en un recipiente común y se lavan con 1,0 litros de CH_2Cl_2 tres veces (tres minutos cada vez), después se lavan de nuevo tres veces (tres minutos cada vez) con 1,0 litros de DIEA/ CH_2Cl_2 al 5% (DIEA = diisopropililamina; CH_2Cl_2 = DCM). Después, las bolsitas se aclaran con DCM y se ponen en recipientes de reacción separados, conteniendo cada uno 50 ml (0,56 M) del t-Boc-aminoácido/DCM respectivo. Se añade N,N-diisopropilcarbodiimida (DIPCDI; 25 ml; 1,12 M) a cada recipiente como agente de acoplamiento. Los veinte derivados aminoácidos se acoplan por separado a la resina en DMF/DCM 50/50 (v/v). Después de una hora de agitación enérgica, se realiza un ensayo de ácido pícrico de Ginsen (Gisen (1972) Anal. Chem. Acta 58: 248-249) para determinar la finalización de la reacción de acoplamiento. Al confirmarse la finalización de la reacción, todos los paquetes de resina se lavan después con 1,5 litros de DMF y se lavan dos veces más con 1,5 litros de CH_2Cl_2 . Después de aclarar, las resinas se eliminan de sus paquetes separados y se mezclan en conjunto para formar una combinación en una bolsita común. La mezcla de resina resultante se seca después y se pesa, se divide de nuevo en 20 porciones iguales (alícuotas) y se ponen en 20 bolsitas de polipropileno adicionales (encerradas).

[0284] En un recipiente de reacción común, se llevan a cabo las etapas siguientes: (1) se lleva a cabo la desprotección en las alícuotas encerradas durante treinta minutos con 1,5 litros de TFA/DCM al 55%; y (2) se lleva a cabo la neutralización con tres lavados de 1,5 litros cada uno de DIEA/DCM al 5%. Cada bolsita se pone en una solución separada del derivado de t-Boc-aminoácido activado y la reacción de acoplamiento se lleva a cabo hasta su finalización como anteriormente. Todas las reacciones de acoplamiento se controlan usando el ensayo de ácido pícrico cuantitativo anterior.

[0285] A continuación, las bolsitas se abren y las resinas de dipéptidos protegidos con t-Boc resultantes se mezclan en conjunto para formar una combinación, se generan alícuotas a partir de la combinación, las alícuotas se encierran, se desprotegen y se llevan a cabo reacciones adicionales. Este procedimiento puede repetirse cualquier número de veces produciendo en cada etapa una representación equimolar del número deseado de restos aminoácidos en la cadena peptídica. Las etapas de procedimiento principales se denominan convenientemente como una síntesis mediante dividir-acoplar-recombinar.

[0286] Después de que se lleven a cabo un número deseado de dichos acoplamientos y mezclas, las bolsitas de polipropileno se mantienen separadas para proporcionar aquí los veinte conjuntos que tienen el resto amino-terminal como el único resto predeterminado, estando por ejemplo las posiciones 2-4 ocupadas por cantidades equimolares de los veinte restos. Para preparar conjuntos que tengan el único resto aminoácido predeterminado en otro que el extremo amino-terminal, el contenido de las bolsitas no se mezcla después de añadir un resto en la posición predeterminada deseada. En su lugar, el contenido de cada una de las veinte bolsitas se separa en 20 alícuotas, se desprotege y después se hace reaccionar por separado con los veinte derivados de aminoácidos. El contenido de cada conjunto de veinte bolsas producido de este modo se mezcla después de eso y se trata como se ha descrito anteriormente hasta que se consigue la longitud de oligopéptido deseada.

5.1.2.3 Síntesis de múltiples péptidos por acoplamiento de mezclas de aminoácidos

[0287] El acoplamiento simultáneo de mezclas de aminoácidos activados a un solo soporte de resina se ha usado como una estrategia de síntesis de múltiples péptidos en varias ocasiones (Geysen et al. (1986) Mol Immunol 23: 709-715; Tjoeng et al. (1990) Int J Pept Protein Res 35: 141-146; Rutter et al. (1991) Patente de Estados Unidos N° 5.010.175; Birkett et al. (1991) Anal Biochem 196: 137-143; Petithory et al. (1991) PNAS 88: 11510-11514) y puede tener aplicaciones en el presente método. Por ejemplo, de cuatro a siete análogos de los péptidos de magainina 2 y

angiotensinógeno se sintetizaron con éxito y se resolvieron en una purificación por HPLC después del acoplamiento de una mezcla de aminoácidos en una sola posición en cada secuencia (Tjoeng et al. (1990) *Int J Pept Protein Res* 35: 141-146). Esta estrategia se ha usado también para preparar mezclas de péptidos degenerados para definir la especificidad de sustrato de enzimas endoproteolíticas (Birkett et al. (1991) *Anal Biochem* 196: 137-143; Petithory et al. (1991) *PNAS* 88: 11510-11514). En estos experimentos, se sustituye una serie de aminoácidos en una sola posición dentro de la secuencia de sustrato. Después de la proteólisis, se usó degradación de Edman para cuantificar el rendimiento de cada componente aminoacídico en el producto de hidrólisis y, por lo tanto, evaluar los valores de K_{cat}/K_m relativos para cada sustrato en la mezcla.

[0288] Sin embargo, se señala que la simplicidad funcional de la síntesis de muchos péptidos por acoplamiento de mezclas de monómeros se compensa por la dificultad de controlar la composición de los productos. La distribución de producto refleja las constantes de velocidad individuales para las reacciones de acoplamiento de competición, añadiéndose derivados activados de restos impedidos estéricamente tales como valina o isoleucina a una velocidad significativamente más lenta que glicina o alanina, por ejemplo. La naturaleza del componente unido a resina de la reacción de acilación también influye en la velocidad de adición, y las constantes de velocidad relativas para la formación de 400 dipéptidos a partir de los 20 aminoácidos codificados genéticamente se han determinado por Rutter y Santi (Rutter et al. (1991) *Patente de Estados Unidos* N° 5.010.175). Estas velocidades de reacción pueden usarse para guiar la selección de concentraciones relativas de aminoácidos apropiadas en la mezcla para favorecer rendimientos de acoplamiento más estrechamente equimolares.

5.1.2.4 Síntesis de múltiples péptidos en soportes sólidos no tradicionales

[0289] La búsqueda de métodos innovadores de síntesis de múltiples péptidos ha conducido a la investigación de soportes poliméricos alternativos a la matriz de poliestireno-divinilbenceno originalmente popularizada por Merrifield. Se ha funcionalizado con éxito celulosa, en forma de discos de papel (Blankemeyer-Menge et al. (1988) *Tetrahedron Lett* 29:5871-5874; Frank et al. (1988) *Tetrahedron* 44: 6031-6040; Eichler et al. (1989) *Collect Czech Chem Commun* 54: 1746-1752; Frank, R. (1993) *Bioorg Med Chem Lett* 3: 425-430) o fragmentos de algodón (Eichler et al. (1991) *Pept Res* 4 : 296-307 ; Schmidt et al. (1993) *Bioorg Med Chem Lett* 3: 441-446) para la síntesis de péptidos. Las cargas típicas logradas con el papel de celulosa varían de 1 a 3 mmol/cm², y el análisis de HPLC del material escindido de estos soportes indica una calidad razonable para los péptidos sintetizados. Como alternativa, los péptidos pueden sintetizarse en láminas de celulosa mediante enlazadores no escindibles y después usarse en estudios de unión basados en ELISA (Frank, R. (1992) *Tetrahedron* 48: 9217-9232). La naturaleza polar porosa de este soporte puede ayudar a suprimir los efectos de unión de proteínas inespecíficas no deseados. Controlando el volumen de aminoácidos activados y otros reactivos aplicados puntualmente sobre el papel, el número de péptidos sintetizados en localizaciones separadas en el soporte puede variarse fácilmente. En una configuración conveniente, se realizan aplicaciones puntuales en un formato de placa de microtitulación de 8 x 12. Frank ha usado esta técnica para mapear los epítomos dominantes de un antisuero generado contra una proteína de citomegalovirus humano, siguiendo la estrategia de exploración de péptidos solapantes (Pepscan) de Geysen (Frank, R. (1992) *Tetrahedron* 48: 9217-9232). Otros soportes de tipo membrana que pueden usarse para múltiples síntesis en fase sólida incluyen películas de polietileno injertadas con poliestireno (Berg et al. (1989) *J Am Chem Soc* 111: 8024-8026).

5.1.2.5 Bibliotecas combinatorias por síntesis química en paralelo abordable espacialmente dirigida por luz

[0290] Un esquema de síntesis combinatoria en el que se proporciona la identidad de un compuesto por su localización en un sustrato de síntesis se denomina síntesis abordable espacialmente. En una realización, el proceso combinatorio se lleva a cabo controlando la adición de un reactivo químico a localizaciones específicas en un soporte sólido (Dower et al. (1991) *Annu Rep Med Chem* 26: 271-280; Fodor, S. P. A. (1991) *Science* 251: 767; Pirrung et al. (1992) *Patente de Estados Unidos* N° 5.143.854; Jacobs et al. (1994) *Trends Biotechnol* 12: 19-26). La técnica combina dos tecnologías bien desarrolladas: química de síntesis de péptidos en fase sólida y fotolitografía. Los altos rendimientos de acoplamiento de la química de Merrifield permiten una síntesis de péptidos eficaz, y la resolución espacial de la fotolitografía permite la miniaturización. La unión de estas dos tecnologías se realiza por medio del uso de grupos protectores de amino fotolábiles en el procedimiento sintético de Merrifield.

[0291] Los puntos clave de esta tecnología se ilustran en Gallop et al. (1994) *J Med Chem* 37: 1233-1251. Se prepara un sustrato de síntesis para el acoplamiento de aminoácidos a través de la unión covalente de enlazadores amino protegidos con nitroveratriloxicarbonilo (NVOC) fotolábiles. Se usa luz para activar selectivamente una región especificada del soporte de síntesis para el acoplamiento. La eliminación de los grupos protectores fotolábiles mediante luces (desprotección) da como resultado la activación de áreas seleccionadas. Después de la activación, el primero de un conjunto de aminoácidos, llevando cada uno un grupo protector fotolábil en el extremo amino-terminal, se expone a toda la superficie. El acoplamiento de aminoácidos se produce solamente en regiones que se abordaron con luz en la etapa precedente. La solución de aminoácidos se elimina y el sustrato se ilumina de nuevo a través de una segunda máscara que activa una región diferente para la reacción con un segundo componente básico protegido. El patrón de máscaras y la secuencia de reactivos definen los productos y sus localizaciones. Como este procedimiento utiliza técnicas de fotolitografía, el número de compuestos que pueden sintetizarse está limitado solamente por el número de sitios de síntesis que pueden abordarse con una resolución apropiada. La posición de cada compuesto se conoce con precisión; por tanto, sus interacciones con otras moléculas pueden evaluarse directamente. La proteína diana puede marcarse con un grupo indicador fluorescente para facilitar la identificación de interacciones específicas con miembros individuales de la matriz.

[0292] En una síntesis química dirigida por luz, los productos dependen del patrón de iluminación y del orden de adición de reactivos. Variando los patrones litográficos, pueden sintetizarse muchos conjuntos de péptidos de ensayo diferentes en el mismo número de etapas; esto conduce a la generación de muchas estrategias de enmascaramiento diferentes.

5.1.2. 6 Bibliotecas combinatorias codificadas

5 **[0293]** En otra realización más, el presente método utiliza una biblioteca de péptidos provista de un sistema de marcaje codificado. Una mejora reciente en la identificación de compuestos activos a partir de bibliotecas combinatorias emplea sistemas de indexación química usando etiquetas que codifican únicamente las etapas de reacción a las que se ha sometido una perla dada y, por inferencia, la estructura que lleva. Conceptualmente, esta estrategia imita las bibliotecas de presentación en fago anteriores, en las que la actividad procede de péptidos expresados pero las estructuras de los péptidos activos se deduce de la secuencia de ADN genómico correspondiente. La primera codificación de bibliotecas combinatorias sintéticas empleaba ADN como código. Se han descrito dos formas de codificación: codificación con biooligómeros secuenciables (por ejemplo, oligonucleótidos y péptidos) y codificación binaria con etiquetas no secuenciables.

5.1.2.6.1 Marcaje con biooligómeros secuenciables

15 **[0294]** El principio del uso de oligonucleótidos para codificar bibliotecas sintéticas combinatorias se describió en 1992 (Brenner et al. (1992) PNAS 89: 5381-5383) y apareció un ejemplo de dicha biblioteca al año siguiente (Needles et al. (1993) PNAS 90: 10700-10704). Una biblioteca combinatoria de nominalmente 77 péptidos (= 823.543) compuestos por todas las combinaciones de Arg, Gln, Phe, Lys, Val, D-Val y Thr (código de aminoácido de tres letras), cada uno de los cuales estaba codificado por un dinucleótido específico (TA, TC, CT, AT, TT, CA y AC, respectivamente) se preparó mediante una serie de rondas alternas de síntesis de péptidos y oligonucleótidos en soporte sólido. En este trabajo, la funcionalidad de unión a amina en la perla estaba específicamente diferenciada hacia la síntesis de péptidos u oligonucleótidos preincubando simultáneamente las perlas con reactivos que generan grupos OH protegidos para la síntesis de oligonucleótidos, y grupos NH₂ protegidos para la síntesis de péptidos (aquí, en una proporción de 1:20). Cuando estaban completas, las etiquetas consistían cada una en 69 unidades, 14 unidades de las cuales llevaban el código. La biblioteca unida a perlas se incubó con un anticuerpo marcado fluorescentemente y las perlas que contenían anticuerpo unido que presentaba una fluorescencia intensa se recogieron por separación de células activadas por fluorescencia (FACS). Las etiquetas de ADN se amplificaron por PCR y se secuenciaron, y los péptidos predichos se sintetizaron. Después de dichas técnicas, las bibliotecas de péptidos pueden derivarse para usarse en el presente método y explorarse usando el D-enantiómero de la proteína diana.

30 **[0295]** Se señala que una estrategia de alternativa útil para generar bibliotecas de péptidos sintéticos codificados por nucleótidos emplea un enlazador ramificado que contiene grupos OH y NH₂ protegidos selectivamente (Nielsen et al. (1993) J Am Chem Soc 115: 9812-9813; y Nielsen et al. (1994) Methods Companion Methods Enzymol 6: 361-371). Esta estrategia requiere que coexistan cantidades equimolares de péptido de ensayo y etiqueta, aunque esto puede ser una complicación potencial en la evaluación de la actividad biológica, especialmente con etiquetas basadas en ácidos nucleicos.

40 **[0296]** El uso de etiquetas de oligonucleótidos permite un análisis de etiquetas exquisitamente sensible. Aun así, el método requiere una selección cuidadosa de conjuntos ortogonales de grupos protectores necesarios para la cosíntesis alterna de la etiqueta y el miembro de la biblioteca. Además, la inestabilidad química de la etiqueta, particularmente los enlaces anoméricos de azúcares y fosfatos, puede limitar la elección de reactivos y condiciones que pueden emplearse para la síntesis en bibliotecas no oligoméricas. En realizaciones preferidas, las bibliotecas emplean enlazadores que permiten la separación selectiva del miembro de biblioteca peptídico de ensayo para su bioensayo, en parte (como se describe a continuación) porque los ensayos que emplean perlas limitan la elección de dianas, y en parte porque las etiquetas son potencialmente susceptibles a la biodegradación.

45 **[0297]** Los propios péptidos se han empujado como moléculas de marcaje para bibliotecas combinatorias. Se describen en la técnica dos estrategias ejemplares, empleando ambas enlazadores ramificados a fase sólida, tras lo cual se elaboran alternativamente cadenas de ligando y de codificación. En la primera estrategia (Kerr JM et al. (1993) J Am Chem Soc 115: 2529-2531), se consigue la ortogonalidad en la síntesis empleando una protección inestable a ácido para la cadena de codificación y una protección inestable a base para la cadena de ligando.

50 **[0298]** En una estrategia alternativa (Nikolaiev et al. (1993) Pept Res 6: 161-170), se emplean enlazadores ramificados de manera que la unidad de codificación y el péptido de ensayo se unen ambos al mismo grupo funcional en la resina. En una realización, puede ponerse un enlazador entre el punto de ramificación y la perla de modo que la escisión libere una molécula que contiene tanto código como ligando (Ptek et al. (1991) Tetrahedron Lett 32: 3891-3894). En otra realización, el enlazador puede situarse de modo que el péptido de ensayo pueda separarse selectivamente de la perla, dejando el código atrás. Esta última construcción es particularmente valiosa porque permite la exploración del péptido de ensayo sin una interferencia potencial, o biodegradación, de los grupos de codificación. Ejemplos en la técnica de escisión independiente y secuenciación de miembros de biblioteca de péptidos y sus etiquetas correspondientes han confirmado que las etiquetas pueden predecir con exactitud la estructura del péptido.

60 **[0299]** Se señala que las etiquetas peptídicas son más resistentes a la descomposición durante la síntesis de ligando de lo que lo son las etiquetas oligonucleotídicas, pero que deben emplearse en proporciones molares casi iguales a las del ligando en perlas de 130 mm típicas para que se secuencien con éxito. Como con la codificación con oligonucleótidos,

el uso de péptidos como etiquetas requiere químicas de protección/desprotección complejas.

5.1.2.6.2 *Marcaje no secuenciable: codificación binaria*

[0300] Una forma alternativa de codificación de la biblioteca de péptidos de ensayo emplea un conjunto de moléculas de marcaje electroforético no secuenciables que se usan como un código binario (Ohlmeyer et al (1993) PNAS 90: 10922-10926). Las etiquetas ejemplares son éteres de alquilo haloaromáticos que son detectables como sus éteres de tetrametilsililo a niveles inferiores a femtomolares por cromatografía de gases de captura de electrones (ECGC). Las variaciones en la longitud de la cadena de alquilo, así como la naturaleza y la posición de los sustituyentes de haluro aromático, permiten la síntesis de al menos 40 etiquetas de este tipo, que en principio pueden codificar 240 (por ejemplo, más de 1012) moléculas diferentes. En el informe original (Ohlmeyer et al., anteriormente), las etiquetas estaban unidas a aproximadamente el 1% de los grupos amina disponibles de una biblioteca de péptidos mediante un enlazador de O-nitrobencilo fotoescindible. Esta estrategia es conveniente cuando se preparan bibliotecas combinatorias de péptidos u otras moléculas que contienen amina. Sin embargo, se ha desarrollado un sistema más versátil que permite la codificación de esencialmente cualquier biblioteca combinatoria. En la presente memoria, el ligando se une al soporte sólido mediante el enlazador fotoescindible y la etiqueta se une a través de un enlazador de éter de catecol mediante inserción de carbeno en la matriz de perla (Nestler et al. (1994) J Org Chem 59:4723-4724). Esta estrategia de unión ortogonal permite la separación selectiva de miembros de la biblioteca para su bioensayo en solución y posterior descodificación por ECGC después de la separación oxidativa de los conjuntos de etiquetas.

[0301] La codificación binaria con etiquetas electroforéticas ha sido particularmente útil para definir interacciones selectivas de sustratos con receptores sintéticos (Borchardt et al. (1994) J Am Chem Soc 116: 373-374) y sistemas de modelo para comprender la unión y la catálisis de biomoléculas. Incluso usando un modelado molecular detallado, la identificación de las preferencias de selectividad para receptores sintéticos ha requerido la síntesis manual de docenas de sustratos potenciales. El uso de bibliotecas codificadas hace posible examinar rápidamente todos los miembros de un conjunto de unión potencial. El uso de bibliotecas de codificación binaria ha hecho que la determinación de las selectividades de unión sea tan fácil que se ha descrito la selectividad estructural para cuatro nuevos receptores tricíclicos y macrobíclicos sintéticos en una sola comunicación (Wennemers et al. (1995) J Org Chem J 60: 1108-1109, y Yoon et al. (1994) Tetrahedron Lett 35: 8557-8560), usando la biblioteca codificada mencionada anteriormente. Se esperaría una facilidad similar para definir la especificidad de interacción para muchas otras biomoléculas.

[0302] Aunque las varias bibliotecas unidas a amida en la técnica emplean una codificación binaria con las etiquetas electroforéticas unidas a grupos amina, la unión de estas etiquetas directamente a la matriz de perla proporciona una versatilidad mucho mayor en las estructuras que pueden prepararse en bibliotecas combinatorias codificadas. Unidas de este modo, las etiquetas y su enlazador son casi tan poco reactivas como la propia matriz de perla. Se han descrito dos bibliotecas combinatorias de codificación binaria en las que las etiquetas electroforéticas se unen directamente a la fase sólida (Ohlmeyer et al (1995) PNAS 92: 6027-6031) y proporcionan consejos para generar la presente biblioteca de péptidos. Ambas bibliotecas se construyeron usando una estrategia de unión ortogonal en la que el miembro de la biblioteca estaba unido al soporte sólido mediante un enlazador fotolábil y las etiquetas estaban unidas a través de un enlazador escindible solamente por oxidación enérgica. Debido a que los miembros de la biblioteca pueden fotoeluirse parcialmente de forma repetitiva del soporte sólido, los miembros de la biblioteca pueden utilizarse en múltiples ensayos. La fotoelución sucesiva también permite una estrategia de exploración iterativa de muy alto rendimiento: en primer lugar, se ponen múltiples perlas en placas de microtitulación de 96 pocillos; en segundo lugar, los ligandos se separan parcialmente y se transfieren a placas de ensayo; en tercer lugar, un bioensayo identifica los pocillos activos; en cuarto lugar, las perlas correspondientes se redisponen individualmente en nuevas placas de microtitulación; en quinto lugar, se identifican compuestos activos individuales; y en sexto lugar, se descodifican las estructuras.

[0303] La estrategia anterior se empleó en la exploración de la unión a anhidrasa carbónica (CA) e identificó compuestos que mostraban afinidades nanomolares por la CA. A diferencia del marcaje secuenciable, un gran número de estructuras pueden descodificarse rápidamente a partir de bibliotecas de codificación binaria (un solo aparato de ECGC puede descodificar 50 estructuras al día). Por lo tanto, pueden usarse bibliotecas de codificación binaria para el análisis rápido de relaciones de estructura-actividad y la optimización tanto de la potencia como de la selectividad de una serie activa. La síntesis y exploración de grandes bibliotecas de péptidos de codificación binaria imparciales para la identificación de candidatos, seguida de la preparación y análisis de bibliotecas centradas más pequeñas para la optimización de candidatos, ofrece una estrategia particularmente potente para el descubrimiento de fármacos usando el presente método.

5. 1. 3 *Bibliotecas de ácidos nucleicos*

[0304] En otra realización, la biblioteca está compuesta por una combinación abigarrada de ácidos nucleicos, por ejemplo, ADN o ARN mono- o bicatenario. Se conocen en la técnica una diversidad de procedimientos para generar bibliotecas de ácido nucleico explorables que pueden aprovecharse en la presente invención. Las bibliotecas que pueden usarse con la presente invención incluyen bibliotecas generadas a partir de: oligonucleótidos sintéticos, secuencias de ADNc, fragmentos de ADN genómico bacteriano y fragmentos de ADN genómico eucariota.

[0305] En particular, muchas de las técnicas descritas anteriormente para bibliotecas de péptido sintéticos pueden usarse para generar bibliotecas de ácidos nucleicos de una diversidad de formatos. Por ejemplo, pueden usarse técnicas de dividir-acoplar-recombinar junto con técnicas de síntesis de ácidos nucleicos convencionales para generar

bibliotecas de ácidos nucleicos inmovilizados en perlas.

[0306] En otra realización, pueden generarse bibliotecas de ácidos nucleicos en solución que dependen de técnicas de PCR para amplificar para secuenciar aquellas moléculas de ácido nucleico que se unen selectivamente a la diana de exploración. Mediante dichas técnicas, se han generado en solución bibliotecas que se aproximan a 10^{15} secuencias de nucleótidos diferentes (véase, por ejemplo, Bartel y Szostak (1993) Science 261: 1411-1418; Bock et al. (1992) Nature 355: 564; Ellington et al. (1992) Nature 355: 850-852; y Oliphant et al. (1989) Mol Cell Biol 9: 2944-2949).

[0307] De acuerdo con una realización del presente método, se emplea SELEX (evolución sistemática de ligandos por enriquecimiento exponencial) con la diana de exploración enantiomérica. Véase, por ejemplo, Tuerk et al. (1990) Science 249: 505-510 para una revisión del SELEX. En resumen, en la primera etapa de estos experimentos se crea una combinación de secuencias de ácido nucleico variantes, por ejemplo, como una biblioteca aleatoria o semialeatoria. En general, se proporciona una secuencia de cebador 3' y (opcionalmente) 5' invariable para usar con PCR de anclaje o para permitir la subclonación. La biblioteca de ácido nucleico se aplica para explorar una diana, y los ácidos nucleicos que se unen selectivamente (o actúan de otro modo sobre la diana) se aíslan de la combinación. Los aislados se amplifican por PCR y se subclonan, por ejemplo, en fagémidos. Después, los fagémidos se transfectan en células bacterianas y pueden obtenerse aislados individuales, y la secuencia del ácido nucleico clonado a partir de la combinación de exploración puede determinarse.

[0308] Cuando el ARN es el ligando de ensayo, la biblioteca de ARN puede sintetizarse directamente mediante química orgánica convencional, o puede proporcionarse por traducción *in vitro* como se describe por Tuerk et al., anteriormente. Así mismo, el ARN aislado por unión a la diana de exploración puede someterse a transcripción inversa y el ADNc resultante subclonarse y secuenciarse como anteriormente.

[0309] El aislamiento de ARNm para la síntesis de ADNc y el aislamiento de ADN genómico, de origen procariota o eucariota, son bien conocidos en la técnica de la biología molecular. Muchos manuales de laboratorio convencionales tales como *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989 o ediciones posteriores), o *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor press (1989 o ediciones posteriores) tienen descripciones detalladas de estos temas. Además, muchas compañías ofrecen kits comerciales diseñados específicamente para tales fines.

5. 2 Bibliotecas de moléculas pequeñas

[0310] Las tendencias recientes en la búsqueda de nuevos agentes farmacológicos se han centrado en la preparación de bibliotecas químicas. Se han descrito anteriormente bibliotecas de péptidos. Son bien conocidas en la técnica bibliotecas de ácidos nucleicos (incluyendo bibliotecas de ADNc, ADN genómico y EST). Se han descrito bibliotecas de sacáridos y su síntesis usando química combinatoria en el documento WO 98/16536 y en sus solicitudes relacionadas. Sin embargo, el campo de la química combinatoria también ha proporcionado grandes cantidades de bibliotecas de moléculas orgánicas pequeñas no poliméricas que pueden emplearse en el presente método.

[0311] Las bibliotecas combinatorias ejemplares incluyen benzodiazepinas, peptoides, biarilos e hidantoínas. En general, las mismas técnicas descritas anteriormente para los diversos formatos de bibliotecas de péptidos sintetizados químicamente pueden usarse también para generar y (opcionalmente) codificar bibliotecas no peptídicas sintéticas.

5. 3 Selección de compuestos de la biblioteca

[0312] Como con la diversidad contemplada para la biblioteca de compuestos y la forma en la que se proporciona la biblioteca de compuestos, el presente método está previsto para identificar ligandos híbridos con la fórmula general de R1-Y-R2, que interaccionen con una diana de exploración polipeptídica, o para identificar inhibidores o antagonistas de una interacción determinada. En la mayoría de las realizaciones, los programas de exploración ensayan bibliotecas de compuestos/ligandos híbridos adecuadas para un análisis de alto rendimiento para maximizar el número de compuestos estudiados en un periodo de tiempo dado. Sin embargo, como norma general, la porción de exploración del presente método implica poner en contacto la diana de exploración con la biblioteca de compuestos y aislar aquellos compuestos de la biblioteca que interaccionen con la diana de exploración o causen un efecto deseado. Dicha interacción entre el compuesto de ensayo/ligandos híbridos y la diana de exploración puede detectarse, por ejemplo, basándose en el cambio de estado de cualquiera de los sistemas indicadores adecuados, como se describen en la sección 3, o la modulación de una actividad enzimática/catalítica de la diana de exploración (por ejemplo, cuando se ensaya la unión de un ligando híbrido por su diana dimerizable potencial). La eficacia de los compuestos de ensayo puede evaluarse generando curvas de respuesta a la dosis a partir de datos obtenidos usando diversas concentraciones del compuesto de ensayo. Además, también puede realizarse un ensayo de control para proporcionar medidas basales para la comparación.

[0313] En una realización, la biblioteca de compuestos abigarrada se somete a enriquecimiento por afinidad para seleccionar compuestos que se unan a una diana de exploración preseleccionada. La expresión "separación por afinidad" o "enriquecimiento por afinidad" incluye, pero sin limitación (1) cromatografía de afinidad utilizando la inmovilización de dianas de exploración, (2) precipitación usando dianas de exploración, (3) separación de células activadas por fluorescencia a la que la biblioteca de compuestos es por tanto susceptible, (4) aglutinación y (5) transferencia de placas. En cada realización, la biblioteca de compuestos se separa en última instancia basándose en la capacidad de un compuesto particular para unirse a una diana de exploración de interés. Véase, por ejemplo, Patente de Estados Unidos Nº 5.223.409 de Ladner et al.; la Publicación Internacional Nº WO 92/18619 de Kang et al.; la

Publicación Internacional N° WO 91/17271 de Dower et al.; la Publicación Internacional WO 92/20791 de Winter et al.; la Publicación Internacional N° WO 92/15679 de Markland et al.; la Publicación Internacional N° WO 93/01288 de Breitling et al.; la Publicación Internacional N° WO 92/01047 de McCafferty et al., la Publicación Internacional N° WO 92/09690 de Garrard et al., y la Publicación Internacional N° WO 90/02809 de Ladner et al.

[0314] Será evidente que, además de utilizar la unión como criterio de separación, las bibliotecas de compuestos pueden fraccionarse basándose en otras actividades de la molécula diana, tales como la modulación de la actividad catalítica o ciertas propiedades bioquímicas.

[0315] En una realización, la unión entre un compuesto químico y un polipéptido diana puede medirse por la actividad del sistema indicador como se ha descrito anteriormente. Por ejemplo, si se usa un sistema indicador basado en ubiquitina para la detección, dependiendo de la identidad del resto Z (el primer aminoácido del resto indicador escindido), la detección podría ser la presencia de cierta actividad del resto indicador (si Z es un aminoácido estabilizante como metionina) o la ausencia de cierta actividad del resto indicador (si Z es un aminoácido distinto de metionina desestabilizante). La actividad a detectar podría ser actividad de transcripción, fluorescencia, actividad enzimática, o cualquier otra actividad biológica o bioquímica descrita anteriormente. Si se usa para la detección un sistema indicador basado en transcripción, la actividad de transcripción del resto indicador puede controlarse para explorar la unión del compuesto o polipéptido a su diana. Los expertos en la materia apreciarán y reconocerán fácilmente otros métodos apropiados adecuados para esas exploraciones.

6. Ácidos nucleicos

[0316] La invención proporciona ácidos nucleicos, incluyendo ciertos genes y homólogos de los mismos, y porciones de los mismos. Los ácidos nucleicos preferidos tienen una secuencia con una homología de al menos aproximadamente el 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80% y, más preferiblemente del 85%, y una homología de más preferiblemente el 90%, y más preferiblemente el 95%, y aún más preferiblemente de al menos el 99% con una secuencia de nucleótidos de un gen particular o complementaria de la misma del ácido nucleico. Se entiende que otros ácidos nucleicos equivalentes incluyen los que codifican polipéptidos que tienen funciones análogas a las descritas en la presente invención usando ejemplos ilustrativos. Los ácidos nucleicos con una identidad de al menos el 90%, más preferiblemente el 95% y más preferiblemente al menos aproximadamente el 98-99% con una secuencia de ácido nucleico representada en una de estas secuencias o una complementaria de la misma, se incluyen también por supuesto dentro del alcance de la invención.

[0317] La invención también se refiere a ácidos nucleicos aislados que comprenden una secuencia de nucleótidos que codifica ciertos polipéptidos, variantes y/o equivalentes de dichos ácidos nucleicos. Se entiende que el término equivalente incluye secuencias de nucleótidos que codifican polipéptidos funcionalmente equivalentes o péptidos funcionalmente equivalentes que tienen una actividad de una proteína tal como se describe en la presente memoria.

[0318] Las secuencias de nucleótidos equivalentes incluirán secuencias que difieren por sustitución, adición o delección de uno o más nucleótidos, tales como variantes alélicas; y por lo tanto incluirán secuencias que difieren de la secuencia de nucleótidos de la invención debido a la degeneración del código genético.

[0319] Independientemente de la especie, los ácidos nucleicos particularmente preferidos de la invención codifican polipéptidos que tienen una similitud o identidad de al menos el 60%, 65%, 70%, 72%, 74%, 76%, 78%, 80%, 90% o 95% con una secuencia de aminoácidos de la invención. Por ejemplo, dichos ácidos nucleicos pueden comprender aproximadamente 50, 60, 70, 80, 90 ó 100 pares de bases. También dentro del alcance de la invención hay moléculas de ácido nucleico para usar como sondas/cebadores o moléculas antisentido (es decir, moléculas de ácido nucleico no codificantes), que pueden comprender al menos aproximadamente 6, 12, 20, 30, 50, 60, 70, 80, 90 ó 100 pares de bases de longitud.

[0320] Otro aspecto de la invención proporciona un ácido nucleico que hibrida en condiciones rigurosas con un ácido nucleico de la invención. Las condiciones de rigurosidad apropiadas que promueven la hibridación de ADN, por ejemplo, cloruro sódico/citrato sódico (SSC) 6,0x a aproximadamente 45°C, seguido de un lavado de SSC 2,0x a 50°C, son conocidas para los expertos en la materia o pueden encontrarse en *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 o en *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor press (1989). Por ejemplo, la concentración salina en la etapa de lavado puede seleccionarse de una baja rigurosidad de SSC aproximadamente 2,0x a 50°C a una alta rigurosidad de SSC aproximadamente 0,2x a 50°C. Además, la temperatura en la etapa de lavado puede aumentarse de condiciones de baja rigurosidad a temperatura ambiente, a aproximadamente 22°C, a condiciones de alta rigurosidad a aproximadamente 65°C. Puede variarse tanto la temperatura como la concentración salina, o la temperatura y la concentración salina pueden mantenerse constantes mientras se cambia la otra variable.

[0321] También se incluyen en el alcance de la invención ácidos nucleicos que tienen una secuencia que difiere de las secuencias de nucleótidos proporcionadas por la invención, o complementarias de las mismas, debido a la degeneración del código genético. Dichos ácidos nucleicos codifican péptidos funcionalmente equivalentes pero difieren en secuencia de la secuencia mostrada en la lista de secuencias debido a la degeneración del código genético. Por ejemplo, varios aminoácidos están designados por más de un triplete. Los codones que determinan el mismo aminoácido, o sinónimos (por ejemplo, CAU y CAC codifican cada uno histidina) pueden dar como resultado mutaciones "silenciosas" que no afecten a la secuencia de aminoácidos de un polipéptido htrb. Sin embargo, se espera que existan

polimorfismos de secuencia de ADN que no conduzcan a cambios en las secuencias de aminoácidos de los presentes polipéptidos entre mamíferos. Un experto en la materia apreciará que estas variaciones en uno o más nucleótidos (por ejemplo, hasta aproximadamente el 3-5% de los nucleótidos) de los ácidos nucleicos que codifican polipéptidos pueden existir entre individuos de una especie dada debido a la variación alélica natural.

5 6.1. Sondas y cebadores

10 **[0322]** Las secuencias de nucleótidos determinadas a partir de la clonación de genes de organismos procariotas o eucariotas permitirán además la generación de sondas y cebadores diseñados para usarse en la identificación y/o clonación de otros homólogos de otras especies. Por ejemplo, la presente invención también proporciona una sonda/cebador que comprende un oligonucleótido sustancialmente purificado, comprendiendo dicho oligonucleótido una región de secuencia de nucleótidos que hibrida en condiciones rigurosas con al menos aproximadamente 12, preferiblemente 25, más preferiblemente 40, 50 ó 75 nucleótidos consecutivos de una secuencia con sentido o antisentido de la invención.

15 **[0323]** En realizaciones preferidas, los cebadores se diseñan para optimizar la especificidad y evitar estructuras secundarias que afecten a la eficacia del cebado. Los cebadores de PCR optimizados de la presente invención se diseñan de modo que los cebadores “cadena arriba” y “cadena abajo” tengan temperaturas de fusión aproximadamente equivalentes, tales como pueden estimarse usando las fórmulas: $T_m (^{\circ}\text{C}) = 81,5 - 16,6(\log[\text{Na}^+]) + 0,41(\%G+C) - 0,63(\%\text{formamida}) - (600/\text{longitud})$, para polinucleótidos largos; o $T_m (^{\circ}\text{C}) = 2(A + T) + 4(G + C)$, para polinucleótidos que comprenden menos de 20 bases. También pueden diseñarse cebadores optimizados usando diversos programas, tales como “Primer3”, proporcionado por el Whitehead Institute for Biomedical Research.

20 6.2. Vectores de la invención

25 **[0324]** La invención proporciona además ciertos plásmidos y vectores que codifican ciertos productos polipeptídicos *in vitro* o *in vivo*. La célula hospedadora puede ser cualquier célula procariota o eucariota. Por lo tanto, una secuencia de nucleótidos obtenida a partir de la clonación de un pre-ARNm de mamífero, que codifica toda o una porción seleccionada del pre-ARNm de longitud completa, puede usarse para producir una forma recombinante del pre-ARNm u otra secuencia de ARN de interés mediante procesos celulares eucariotas o microbianos. La ligación de la secuencia polinucleotídica en una construcción génica, tal como un vector de expresión, y su transformación o transfección en células hospedadoras eucariotas (de levadura, aves, insectos o mamíferos) o procariotas (bacterianas) son procedimientos convencionales bien conocidos en la técnica.

30 **[0325]** Los vectores que permiten la expresión de un ácido nucleico en una célula se denominan vectores de expresión. Típicamente, los vectores de expresión usados para expresar un sustrato de afinidad de ARN de la invención codifican una secuencia de ensamblaje de ribonucleoproteína y una secuencia de etiqueta de afinidad que contiene un ácido nucleico que codifica un sitio de unión de proteína de unión a ARN, unido operativamente a al menos una secuencia reguladora de la transcripción. Están reconocidas en la técnica secuencias reguladoras. Se describen secuencias reguladoras de la transcripción en Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990).

35 **[0326]** Los vectores adecuados para la expresión del sustrato de afinidad de ARN incluyen plásmidos de los tipos: plásmidos derivados de pBR322, plásmidos derivados de pEMBL, plásmidos derivados de pEX, plásmidos derivados de pBTac y plásmidos derivados de pUC para la expresión en células procariotas, tales como *E. coli*.

40 **[0327]** Existen varios vectores para la expresión de proteínas recombinantes en levaduras. Por ejemplo, YEP24, YIP5, YEP51, YBP52, pYES2 y YRP17 son vehículos de clonación y expresión útiles en la introducción de construcciones genéticas en *S. cerevisiae* (véase, por ejemplo, Broach et al. (1983) en Experimental Manipulation of Gene Expression, ed. M. Inouye Academic Press, pág. 83).

45 **[0328]** Estos vectores pueden replicarse en *E. coli* debido a la presencia del ori de pBR322, y en *S. cerevisiae* debido al determinante de la replicación del plásmido de 2 micrómetros de levadura. Además, pueden usarse marcadores de resistencia a fármacos tales como ampicilina.

50 **[0329]** Los vectores de expresión preferidos contienen ambas secuencias promotoras procariotas, tales como un promotor de T7 o un promotor de SP6, de modo que puedan generarse sustratos de afinidad de ARN sintéticos *in vitro* usando metodologías convencionales. Los diversos métodos empleados en la preparación de los plásmidos y la transformación de los organismos hospedadores son bien conocidos en la técnica. Para otros sistemas de expresión adecuados para células tanto procariotas como eucariotas, así como procedimientos recombinantes en general, véase Molecular Cloning A Laboratory Manual, 2ª Ed., ed. por Sambrook, Fritsch y Maniatis (Cold Spring Harbor Laboratory Press: 1989).

55 **[0330]** En algunos casos, puede ser deseable expresar un polipéptido recombinante usando un sistema de expresión de baculovirus. Los ejemplos de dichos ejemplos de expresión de baculovirus incluyen vectores derivados de pVL (tales como pVL1392, pVL1393 y pVL941), vectores derivados de pAcUW (tales como pAeUVf1) y vectores derivados de pBlueBac (tales como el pBlueBac III que contiene β -gal).

[0331] Cuando es deseable expresar sólo una porción de una proteína, tal como una forma que carece de una porción del extremo N-terminal, es decir, un mutante de truncamiento que carece del péptido señal, puede ser necesario añadir un codón de inicio (ATG) al fragmento oligonucleotídico que contiene la secuencia deseada a expresar. Es bien sabido en la técnica que una metionina en la posición N-terminal puede escindirse enzimáticamente mediante el uso de la enzima metionina aminopeptidasa (MAP). La MAP se ha clonado a partir de *E. coli* (Ben-Bassat et al. (1987) J. Bacteriol. 169: 751-757) y *Salmonella typhimurium* y su actividad *in vitro* se ha demostrado en proteínas recombinantes (Miller et al. (1987) PNAS 84: 2718-1722). Por lo tanto, la eliminación de una metionina N-terminal, si se desea, puede conseguirse *in vivo*, expresando polipéptidos en un hospedador que produce MAP (por ejemplo, *E. coli* ox CM89 o *S. cerevisiae*), o *in vitro*, mediante el uso de MAP purificada (por ejemplo, procedimiento de Miller et al., anteriormente).

[0332] Además, las construcciones génicas de la presente invención también pueden usarse como parte de un protocolo de terapia génica para administrar ácidos nucleicos que codifican una forma agonista o antagonista de uno de los presentes complejos de ribonucleoproteína. Por lo tanto, otro aspecto de la invención presenta vectores de expresión para la transfección *in vivo* o *in vitro* y la expresión de un polipéptido en tipos celulares particulares para reconstituir la función de, o como alternativa, suprimir la función de un complejo de ribonucleoproteína en un tejido. Por lo tanto, podría ser deseable, por ejemplo, cuando la forma de origen natural de la proteína está expresada erróneamente o la proteína natural está mutada y es menos activa.

7. Polipéptidos de la presente invención

[0333] La presente invención proporciona métodos para identificar polipéptidos que interaccionen con un ligando dado. Los polipéptidos identificados por medio de dichos métodos pueden producirse en gran cantidad usando cualquier método reconocido en la técnica, como un polipéptido purificado o como un polipéptido de fusión purificado con otros polipéptidos. Todas las formas de polipéptidos pueden formularse, con un excipiente farmacéutico aceptable, en una composición farmacéutica usando cualquier método reconocido en la técnica.

[0334] Dicho polipéptido purificado estará aislado de, o de otro modo sustancialmente libre de otras proteínas celulares. La expresión "sustancialmente libre de otras proteínas celulares" (también denominadas en la presente memoria "proteínas contaminantes") o "preparaciones sustancialmente puras o purificadas" se define que abarcan preparaciones de polipéptidos que tienen menos de aproximadamente el 20% (en peso seco) de proteína contaminante, y que tienen preferiblemente menos de aproximadamente el 5% de proteína contaminante. Pueden prepararse formas funcionales de los presentes polipéptidos por primera vez como preparaciones purificadas usando un gen clonado como se describe en la presente memoria.

[0335] Los presentes polipéptidos preferidos tienen una secuencia de aminoácidos que tiene una identidad u homología de al menos aproximadamente el 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 85%, 90% o 95% con una secuencia de aminoácidos. Los presentes polipéptidos aún más preferidos comprenden una secuencia de aminoácidos de al menos 10, 20, 30 ó 50 restos que tienen una homología o identidad de al menos aproximadamente el 70, 80, 90, 95, 97, 98 ó 99% con una secuencia de aminoácidos. Dichas proteínas pueden ser proteínas recombinantes y, por ejemplo, pueden producirse *in vitro* a partir de ácidos nucleicos que comprenden una secuencia de nucleótidos identificada por los métodos de la invención u homólogos de la misma. Por ejemplo, los polipéptidos recombinantes preferidos por la presente invención pueden estar codificados por un ácido nucleico que tenga una homología de al menos el 85%, y más preferiblemente una homología del 90%, y más preferiblemente una homología del 95% con una secuencia de nucleótidos identificada por los métodos de la invención. Los polipéptidos que están codificados por un ácido nucleico que tiene una homología de al menos aproximadamente el 98-99% con la secuencia identificada por los métodos de la invención también se incluyen en el alcance de la invención.

[0336] El alcance de la invención también incluye isoformas de los presentes polipéptidos codificados por variantes de corte y empalme. Dichas isoformas pueden tener actividades biológicas idénticas o diferentes. Dichas isoformas pueden surgir, por ejemplo, por corte y empalme alternativo de uno o más transcritos de genes.

[0337] Se incluyen en el alcance de la presente invención proteínas de longitud completa o fragmentos que corresponden a uno o más motivos y/o dominios particulares o a tamaños arbitrarios, por ejemplo, de al menos 5, 10, 20, 25, 50, 75 y 100 aminoácidos de longitud.

[0338] Por ejemplo, los polipéptidos aislados pueden estar codificados por toda o una porción de una secuencia de ácido nucleico. Pueden obtenerse porciones peptídica aisladas de proteínas por exploración de péptidos producidos recombinantemente a partir del fragmento correspondiente del ácido nucleico que codifica dichos péptidos. Además, los fragmentos pueden sintetizarse químicamente usando procedimientos conocidos en la técnica, tales como química de f-Moc o t-Boc en fase sólida de Merrifield convencional. Por ejemplo, un presente polipéptido puede dividirse arbitrariamente en fragmentos de longitud deseada sin solapamiento de los fragmentos, o dividirse preferiblemente en fragmentos solapantes de una longitud deseada. Los fragmentos pueden producirse (recombinantemente o por síntesis química) y ensayarse para identificar aquellos fragmentos de peptídico que pueden funcionar como agonistas o antagonistas de una proteína de tipo silvestre (por ejemplo, "auténtica").

[0339] Un polipéptido puede ser una forma unida a membrana o una forma soluble. Un polipéptido soluble preferido es un polipéptido que no contiene un dominio de secuencia señal hidrófobo. Dichas proteínas pueden crearse por

ingeniería genética mediante métodos conocidos en la técnica. La solubilidad de un polipéptido recombinante puede aumentarse por delección de dominios hidrófobos, tales como dominios transmembrana predichos, de la proteína de tipo silvestre.

[0340] En general, los polipéptidos que se menciona en la presente memoria que tienen una actividad (por ejemplo, que son "bioactivos") de una proteína se definen como polipéptidos que incluyen una secuencia de aminoácidos codificada por toda o una porción de las secuencias de ácido nucleico y que mimetizan o antagonizan toda o una parte de las actividades biológicas/bioquímicas de una proteína de origen natural. Los ejemplos de dicha actividad biológica incluyen una región de estructura conservada denominada el dominio conservado.

[0341] Otras actividades biológicas de las presentes proteínas serán razonablemente evidentes para los expertos en la materia. De acuerdo con la presente invención, un polipéptido tiene actividad biológica si es un agonista o antagonista específico de una forma de origen natural de una proteína.

[0342] Además de utilizar proteínas de fusión para aumentar la inmunogenicidad, se aprecia ampliamente que las proteínas de fusión también pueden facilitar la expresión de proteínas y, por consiguiente, pueden usarse en la expresión de los polipéptidos de la presente invención. Por ejemplo, pueden generarse polipéptidos como proteína con glutatión-S-transferasa (fusión con GST). Dichas proteínas de fusión con GST pueden permitir una fácil purificación del polipéptido, como por ejemplo, por aumento de matrices derivatizadas con glutatión (véase, por ejemplo, Current Protocols in Molecular Biology, eds, Ausubel et al. (N. Y.: John Wiley & Sons, 1991)). Además, puede usarse la fusión de polipéptidos con etiquetas epitópicas pequeñas, tales como las secuencias de etiquetas de FLAG o hemaglutinina, para simplificar la purificación inmunológica del polipéptido recombinante resultante o facilitar la detección inmunológica en una muestra celular o tisular. La fusión con la proteína verde fluorescente y versiones recombinantes de la misma que se conocen en la técnica y están disponibles en el mercado pueden usarse adicionalmente para localizar polipéptidos dentro de células y tejidos vivos.

[0343] Los presentes polipéptidos pueden producirse por cualquier método conocido en la técnica. Por ejemplo, una célula hospedadora transfectada con un vector de ácido nucleico que dirige la expresión de una secuencia de nucleótidos que codifica los presentes polipéptidos puede cultivarse en condiciones apropiadas para permitir que se produzca la expresión del péptido. Son bien conocidos en la técnica medios adecuados para el cultivo de células. El polipéptido recombinante puede aislarse del medio de cultivo celular, células hospedadoras o ambos usando procedimientos conocidos en la técnica para purificar proteínas incluyendo cromatografía de intercambio iónico, cromatografía de filtración en gel, ultrafiltración, electroforesis y purificación por inmunoafinidad con anticuerpos específicos para dicho péptido. En una realización preferida, el polipéptido recombinante es una proteína de fusión que contiene un dominio que facilita su purificación, tal como una proteína de fusión con GST.

[0344] Además, se apreciará en general que, en determinadas circunstancias, puede ser ventajoso proporcionar homólogos de uno de los presentes polipéptidos que funcionan en una capacidad limitada como uno de un agonista (mimético) o un antagonista para promover o inhibir sólo un subconjunto de las actividades biológicas de la forma de origen natural de la proteína. Por lo tanto, pueden generarse efectos biológicos específicos por tratamiento con un homólogo de función limitada, y con menores efectos secundarios respecto al tratamiento con agonistas o antagonistas que se dirigen a todas las actividades biológicas de formas de origen natural de proteínas.

[0345] Pueden generarse homólogos de cada una de las presentes proteínas por mutagénesis, tal como por mutación o mutaciones puntuales separadas, o por truncamiento. Por ejemplo, la mutación puede dar origen a homólogos que conserven sustancialmente la misma, o simplemente un subconjunto de las actividades biológicas del polipéptido del que proceden. Como alternativa, pueden generarse formas antagonistas de la proteína que son capaces de inhibir la función de la forma de origen natural de la proteína, tal como por unión competitiva a un receptor.

[0346] Los polipéptidos recombinantes de la presente invención también incluyen homólogos de las proteínas de tipo silvestre, tales como versiones de aquellas proteínas que son resistentes a la escisión proteolítica, como por ejemplo debido a mutaciones que alteran la ubiquitinación u otro direccionamiento enzimático asociado con la proteína.

[0347] Los polipéptidos también pueden modificarse químicamente para crear derivados por formación de conjugados covalentes o agregados con otros restos químicos, tales como grupos glicosilo, lípidos, fosfatos, grupos acetilo y similares. Pueden prepararse derivados covalentes de proteínas uniendo los restos químicos a grupos funcionales en las cadenas laterales de aminoácidos de la proteína o en el extremo N-terminal o el extremo C-terminal del polipéptido.

[0348] La modificación de la estructura de los presentes polipéptidos puede ser para fines tales como aumentar su eficacia terapéutica o profiláctica, su estabilidad (por ejemplo, vida útil ex vivo y resistencia a la degradación proteolítica) o modificaciones postraduccionales (por ejemplo, alterar el patrón de fosforilación de la proteína). Dichos péptidos modificados, cuando están diseñados para conservar al menos una actividad de la forma de origen natural de la proteína, o para producir antagonistas específicos de la misma, se consideran equivalentes funcionales de los polipéptidos descritos en más detalle en la presente memoria. Dichos péptidos modificados pueden producirse, por ejemplo, por sustitución, delección o adición de aminoácidos. La variante de sustitución puede ser un aminoácido conservado sustituido o un aminoácido no conservado sustituido.

[0349] Por ejemplo, es razonable esperar que una sustitución aislada de una leucina con una isoleucina o valina, un aspartato con un glutamato, una treonina con una serina, o una sustitución similar de un aminoácido con un aminoácido

estructuralmente relacionado (es decir, mutaciones isoestéricas y/o isoelectricas), no tenga un efecto importante sobre la actividad biológica de la molécula resultante. Las sustituciones conservativas son aquellas que tienen lugar dentro de una familia de aminoácidos que están relacionados en sus cadenas laterales. Los aminoácidos codificados genéticamente pueden dividirse en cuatro familias: (1) ácidos = aspartato, glutamato; (2) básicos = lisina, arginina, histidina; (3) no polares = alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano; y (4) polares sin carga = glicina, asparagina, glutamina, cisteína, serina, treonina, tirosina. De forma similar, el repertorio de aminoácidos puede agruparse como (1) ácidos = aspartato, glutamato; (2) básicos = lisina, arginina, histidina; (3) alifáticos = glicina, alanina, valina, leucina, isoleucina, serina, treonina, agrupándose opcionalmente por separado la serina y la treonina como hidroxilo-alifáticos; (4) aromáticos = fenilalanina, tirosina, triptófano; (5) amida = asparagina, glutamina; y (6) que contienen azufre = cisteína y metionina (véase, por ejemplo, Biochemistry, 2ª ed., Ed por L. Stryer, WFT Freeman and Co.: 1981). Puede determinarse fácilmente si un cambio en la secuencia de aminoácidos de un péptido da como resultado un homólogo funcional (por ejemplo, funcional en el sentido de que el polipéptido resultante mimetice o antagonice la forma de tipo silvestre) evaluando la capacidad del péptido variante para producir una respuesta en células de una forma similar a la proteína de tipo silvestre, o inhibir competitivamente dicha respuesta. Pueden ensayarse fácilmente de la misma forma polipéptidos en los que haya tenido lugar más de una sustitución.

[0350] Esta invención contempla además la generación de conjuntos de mutantes combinatorios de los presentes polipéptidos, así como mutantes de truncamiento, y es especialmente útil para identificar secuencias variantes potenciales (por ejemplo, homólogos). El propósito de explorar dichas bibliotecas combinatorias es general, por ejemplo, nuevos homólogos que pueden actuar como agonistas o antagonistas o, como alternativa, poseen nuevas actividades en conjunto. Por lo tanto, pueden generarse homólogos derivados combinatoriamente para que tengan una potencia aumentada respecto a una forma de origen natural de la proteína.

[0351] En una realización, la biblioteca abigarrada de variantes se genera por mutagénesis combinatoria de ácido nucleico y está codificada por una biblioteca de genes abigarrada. Por ejemplo, una mezcla de oligonucleótidos sintéticos puede ligarse enzimáticamente en secuencias de genes de modo que el conjunto degenerado de secuencias potenciales pueda expresarse como polipéptidos individuales o, como alternativa, como un conjunto de proteínas de fusión de mayor tamaño (por ejemplo, para presentación en fago) que contienen el conjunto de secuencias de la presente memoria.

[0352] Existen muchas formas por las que pueden generarse dichas bibliotecas de homólogos potenciales a partir de una secuencia oligonucleotídica degenerada. La síntesis química de una secuencia génica degenerada puede realizarse en un sintetizador de ADN automático, y los genes sintéticos ligarse entonces en todos los vectores de expresión apropiados. El propósito de un conjunto degenerado de genes es proporcionar, en una mezcla, todas las secuencias que codifican el conjunto deseado de secuencias potenciales. La síntesis de oligonucleótidos degenerados es bien conocida en la técnica (véase, por ejemplo, Narang, SA (1983) Tetrahedron 39: 3; Itakura et al. (1981) Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules, ed. AG Walton, Amsterdam: Elsevier págs. 273-289; Itakura et al. (1984) Annu. Rev. Biochem. 53: 323; Itakura et al. (1984) Science 198: 1056; Ike et al. (1983) Nucleic Acid Res. 11: 477. Dichas técnicas se han empleado en la evolución dirigida de otras proteínas (véase, por ejemplo, Scott et al. Science (1990) 249: 386-390; Roberts et al (1992) PNAS 89: 2429-2433; Devlin et al (1990) Science 249: 404-406; Cwirla et al (1990) PNAS 87: 6378-6382; así como las Patentes de Estados Unidos N° 5.223.409, 5.198.346 y 5.096.815).

[0353] Así mismo, puede proporcionarse una biblioteca de fragmentos de secuencia codificante para cualquier clon para generar una población abigarrada de fragmentos para la exploración y posterior selección de fragmentos bioactivos. Se conocen en la técnica una diversidad de procedimientos para generar dicha biblioteca, incluyendo la síntesis química. En una realización, una biblioteca de fragmentos de secuencia codificante puede generarse por (i) tratamiento de un fragmento de PCR bicatenario de una secuencia codificante con una nucleasa en condiciones en las que se produzca mellado sólo aproximadamente una vez por molécula; (ii) desnaturalización del ADN bicatenario; (iii) renaturalización del ADN para formar ADN bicatenario que puede incluir parejas con sentido/antisentido de productos mellados diferentes; (iv) eliminación de las porciones monocatenarias de dúplex vueltos a formar por tratamiento con S1 nucleasa; y (v) ligación de la biblioteca de fragmentos resultante en un vector de expresión. Mediante este método ejemplar, puede obtenerse una biblioteca de expresión que codifique fragmentos N-terminales, C-terminales e internos de diversos tamaños.

[0354] La invención también proporciona la reducción de las proteínas para generar miméticos, por ejemplo, agentes peptídicos o no peptídicos, tales como moléculas pequeñas, que son capaces de interrumpir la unión de un presente polipéptido con una molécula, por ejemplo, un péptido diana. Por lo tanto, dichas técnicas de mutagénesis que se han descrito anteriormente también son útiles para mapear los determinantes de las proteínas que participan en interacciones de proteína-proteína implicadas en, por ejemplo, la unión del presente polipéptido a un péptido diana. Para ilustrarlo, los restos críticos de un presente polipéptido que están implicados en el reconocimiento molecular de su receptor pueden determinarse y usarse para generar moléculas pequeñas o peptidomiméticos derivados que inhiban competitivamente la unión de la proteína auténtica con ese resto. Empleando, por ejemplo, mutagénesis por barrido para mapear los restos aminoácidos de las presentes proteínas que están implicadas en la unión a otras proteínas, pueden generarse compuestos peptidomiméticos que mimeticen aquellos restos de la proteína que facilitan la interacción. Dichos miméticos pueden usarse después para interferir con la función normal de una proteína. Por ejemplo, pueden generarse análogos peptídicos no hidrolizables de dichos restos usando benzodiazepina (por ejemplo, véase, Freidinger et al. en Peptides: Chemistry and Biology, G. R- Marshall ed., ESCOM Publisher: Leiden, Países

Bajos, 1988), azepina (por ejemplo, véase, Huffman et al. en Peptides: Chemistry and Biology G.R. Marshall ed., ESCOM. Publisher. Leiden, Países Bajos, 1988), anillos de gamma lactámicos sustituidos (Garvey et al. en Peptides: Chemistry and Biology, G. R. Marshall ed., ESCOM. Publisher: Leiden, Países Bajos, 1988), pseudopéptidos de cetometileno (Ewenson et al (1986) J Med Chem 29:295; y Ewenson et al en Peptides: Structure and Function (Proceedings of the American Peptide Symposium) Pierce Chemical Co, Rockland, IL, 1985), núcleos dipeptídicos de giros β (Nagai et al (1985) tetrahedron Lett 26: 647; y Sato et al (1986) 3 Chem Soc Perkin Trans 1:1231) y b-aminoalcoholes (Gordon et al (1985) Biochem Biophys Res Com: mun126:419; y Dann et al (1986) Biochem Biophys Res Commun 134: 71).

8. Kits

[0355] La invención proporciona además kits para crear ligandos híbridos que incluyen un ligando químico especificado por el usuario. El compuesto o agente envasarse en un recipiente adecuado. El kit puede comprender además instrucciones para usar el kit para aislar proteínas de unión para el ligando especificado por el usuario del ligando híbrido.

[0356] Por lo tanto, un aspecto de la invención proporciona un kit que comprende un polinucleótido que codifica al menos un dominio de unión a ligando y un dominio funcional heterólogo respecto al dominio de unión a ligando, que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional, que comprende además instrucciones 1) para sintetizar un ligando híbrido de estructura general R1-Y-R2 y 2) para ensayar la unión entre el ligando híbrido y el dominio de unión a ligando, en el que uno de R1 y R2 se une a o inhibe una quinasas.

[0357] Otro aspecto de la invención proporciona un kit que comprende un polinucleótido que codifica al menos un dominio de unión a ligando y un dominio funcional heterólogo respecto al dominio de unión a ligando, que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional, que comprende además instrucciones 1) para sintetizar un ligando híbrido de estructura general R1-Y-R2 y 2) para ensayar la unión entre el ligando híbrido y el dominio de unión a ligando, en el que Y es de la estructura general (CH₂-X-CH₂), donde X representa O, S, SO o SO₂ y n es un número entero de 2 a 25.

[0358] Otro aspecto de la invención proporciona un kit que comprende un polinucleótido que codifica al menos un dominio de unión a ligando y un dominio funcional heterólogo respecto al dominio de unión a ligando que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional, que comprende además instrucciones 1) para sintetizar un ligando híbrido de estructura general R1-Y-R2 y 2) para ensayar la unión entre el ligando híbrido y el dominio de unión a ligando, en el que el dominio funcional es Cub o Nux.

[0359] Otro aspecto de la invención proporciona un kit que comprende: 1) un compuesto de estructura general R1-Y-L, en el que Y es de la estructura general (CH₂-X-CH₂)_n y L es un grupo químico que se sustituye fácilmente por un grupo químico diferente y 2) instrucciones para usar el compuesto para la síntesis de un ligando híbrido R1-Y-R2, donde R1 es diferente de R2 y al menos uno de R1 y R2 no es un péptido.

9. Métodos de trabajo

[0360] Otros aspectos de la invención proporcionan ciertos métodos para trabajar. En particular, la puesta en práctica de los métodos de la invención puede identificar ciertos ligandos híbridos, inhibidores y polipéptidos. Esta etapa técnica, cuando se combina con una o más etapas adicionales proporciona nuevas estrategias para realizar trabajos farmacéuticos, agroquímicos, biotecnológicos o, preferiblemente, de las ciencias de la vida. Por ejemplo, dichas composiciones identificadas por el método de la invención pueden ensayarse para determinar su eficacia como agentes terapéuticos en una diversidad de modelos de enfermedad, y las composiciones terapéuticas potenciales ensayarse después para determinar su toxicidad y para la generación de otros perfiles de seguridad antes de la formulación, envasado y posterior comercialización de la formulación resultante para el tratamiento de una enfermedad. Como alternativa, los derechos a desarrollar y comercializar dichas formulaciones o a realizar dichas etapas pueden concederse a terceros para su consideración. En ciertos otros aspectos de la invención, los ligandos híbridos, inhibidores y polipéptidos identificados de este modo pueden tener utilidad en la forma de información que puede proporcionarse a un tercero para su consideración, de modo que se obtenga una comprensión mejorada de la función o de los efectos secundarios de dichos ligandos híbridos, inhibidores y polipéptidos en un contexto biológico o terapéutico.

[0361] A modo de ejemplo, un método preferiblemente particular de trabajar comprende:

- (i) la identificación de polipéptidos que se unen a un ligando híbrido de fórmula general R1-Y-R2, en la que Y es de la estructura general (CH₂-X-CH₂)_n, R1 es diferente de R2 y al menos uno de R1 y R2 no es un péptido, X = O, S, SO o SO₂, y en el que no se sabía previamente que dichos polipéptidos se unieran a dicho ligando híbrido, y
- (ii) proporcionar el acceso a los datos, ácidos nucleicos o polipéptidos obtenidos a partir de dicha identificación a otra parte para su consideración.

Ejemplos

[0362] La presente invención se ilustra adicionalmente mediante los ejemplos siguientes, que no deberían interpretarse como limitantes de ningún modo. Un experto en la materia, que haya leído la memoria descriptiva y los ejemplos de la presente memoria, apreciará fácilmente la posibilidad de numerosas modificaciones, sustituciones, combinaciones, permutaciones y mejoras de los métodos y composiciones de la invención como se describen en la presente memoria. Dichas modificaciones, sustituciones, combinaciones, permutaciones y mejoras se consideran parte de la presente invención.

[0363] La práctica de la presente invención empleará, a menos que se indique otra cosa, técnicas convencionales de química, biología celular, cultivos celulares, biología molecular, microbiología y ADN recombinante, que están dentro de la especialidad en la técnica. Dichas técnicas se explican con todo detalle en la bibliografía. Véase, por ejemplo, *Molecular Cloning: A Laboratory Manual*, 2ª Ed., ed. por Sambrook, Fritsch y Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volúmenes I y II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis et al.; Patente de Estados Unidos Nº 4.683.195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); el tratado, *Methods In Enzymology* (Academic Press, Inc., N. Y.); *Methods In Enzymology*, Vols. 154 y 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer y Walker, eds., Academic Press, Londres, 1987).

[0364] La técnica de ubiquitina dividida se usó para detectar interacciones de proteína *in vivo* o *in vitro*. Generalmente, es útil para todas las clases de interacciones de proteína-proteína, pero es particularmente útil en los casos en los que es problemático el ensayo de doble híbrido de levadura convencional, es decir, proteínas de membrana y citosólicas, activadores o represores de la transcripción, etc.

Ejemplo 1: Síntesis de compuesto

[0365] Lo siguiente es una descripción de la síntesis de los ligandos híbridos usados en la presente memoria. Sin embargo, esta descripción debe entenderse como ejemplar por naturaleza, y de ningún modo limitará el alcance de los compuestos de acuerdo con la invención inmediata. El experto en la materia será fácilmente capaz de prever otras rutas sintéticas para los compuestos que se proporcionan por la presente invención. Por ejemplo, sin limitación, los componentes básicos $H_2N-CH_2-(CH_2-O-CH_2-O)_n-CH_2-N_3$ con $n = 3, 6$ y 12 están disponibles de fuentes comerciales (Toronto Research Chemicals Inc., Toronto, CA; Fluka, Buchs, CH) y pueden emplearse para la síntesis de compuestos de la estructura general $R1-Y-R2$ con $Y = (-CH_2-O-CH_2)_n-$, por ejemplo, sin limitación, mediante una estrategia de síntesis como se usa a continuación en la síntesis de GPC 285937 siguiendo el Esquema 2 (véase la Figura 1B).

[0366] En los compuestos usados en la presente memoria, un resto de metotrexato se une mediante dos o más restos de polietilenglicol como enlazador a dexametasona (GPC 285937), o a compuestos que se sabe que se unen a o inhiben CDK. Estos inhibidores de CDK potenciales o conocidos (CDKi) pueden unirse a metotrexato mediante un enlazador en una orientación que conserve su actividad hacia la inhibición de CDK (GPC 285985, la CI_{50} para CDK2 es de aproximadamente 180 nM), o en una orientación que suprima esta actividad (GPC 285993, $CI_{50} > 10 \mu M$). Para una comparación con resultados previos usando metotrexato unido a otros compuestos en un ensayo de triple híbrido (Lin et al, J. Am. Chem. Soc. 2000, 122: 4247-8), se empleó un ligando híbrido de metotrexato-enlazador-dexametasona que usa un metadibenzotioéster como enlazador (Mtx-mdbt-Dex). Para el establecimiento del efecto de variar exclusivamente el enlazador, se sintetizaron dos ligandos híbridos en los que el metotrexato se unía a un compuesto con actividad inhibidora de CDK mediante un enlazador que contenía 3 (GPC 286004) o 5 (GPC 286026) unidades de polietilenglicol.

[0367] Excepto cuando se indique explícitamente, todos los reactivos químicos y disolventes usados están disponibles en el mercado de proveedores con los que el experto en la materia está bien familiarizado, por ejemplo, Sigma-Aldrich (St. Louis, MO, Estados Unidos) y sus filiales.

Síntesis de GPC 285937 siguiendo el Esquema 1 (véase la Figura 1A)

Síntesis de (2R)-4-[N-(2-{2-[2-(2-azidoetoxi)etoxi]etil})carbamoil]-2-[(fluoren-9-ilmetoxi)carbonilamino]butanoato de terc-butilo (3)

[0368] Se disolvió éster a-terc-butílico del ácido Fmoc-glutámico (2,15 g, 5,1 mmol) en 10 ml de dimetil formamida (DMF) y se añadió 1-amino-11-azido-3,6,9-trioxaundecano (1,0 g, 4,6 mmol) en 10 ml de DMF. A esta solución, se añadió O-benzotriazol-N,N,N',N'-tetrametil-uronio-hexafluorofosfato (HBTU) (2,3 g, 6 mmol) y diisopropiletilamina (DIEA) (1,75 ml, 10 mmol), y la reacción se agitó a temperatura ambiente durante 2 horas. La mezcla de reacción se diluyó con 100 ml de acetato de etilo y la fase orgánica se lavó con bicarbonato sódico saturado, ácido cítrico al 10% y salmuera, y después se secó sobre sulfato de magnesio y se concentró hasta un aceite marrón. El producto bruto (compuesto **3**) se purificó por cromatografía en sílice ultrarrápida (MeOH al 2% en EtOAc) para dar un aceite marrón claro, 2,3 g, 3,7 mmol, 80%.

Síntesis de (2R)-2-amino-4-[N-(2-{2-[2-(2-azidoetoxi)etoxi]etil})carbamoil]butanoato de terc-butilo (4).

[0369] El Compuesto **3** (2,7 g, 4,3 mmol) se disolvió en 30 ml de cloruro de metileno y se añadieron 30 ml de

dietilamina. La mezcla de reacción se agitó a temperatura ambiente durante 2 h y después se concentró hasta un aceite a presión reducida. El residuo se disolvió con éter dietílico y acetato de etilo (aproximadamente 50 ml c/u) y se extrajo con ácido cítrico al 10%. La fase acuosa se neutralizó hasta pH13 con NaOH 10N y se extrajo con acetato de etilo. La fase orgánica se lavó con salmuera, se secó sobre sulfato de magnesio y se concentró a presión reducida para dar 1,6 g de un aceite marrón, 4,0 mmol, 92% (compuesto 4).

Síntesis de (2R)-4-[N-(2-{2-[2-(2-azidoetoxi)etoxi]etoxi}etil)carbamoil]-2-[(4-[(2,4-diaminopteridin-6-il)metil] metilamino} fenil)carbonilamino]butanoato de terc-butilo (6)

[0370] El compuesto 4 (140 mg, 0,35 mmol) y ácido ptericoico (compuesto 5) se disolvieron juntos en 5 ml de DMF y se añadió hexafluorofosfato de benzotriazol-1-il-oxi-tris-pirrolidino-fosfonio (PyBop) (0,26 g, 0,50 mmol) como un sólido, seguido de DIEA (0,3 ml, 1,7 mmol). La mezcla de reacción se agitó a temperatura ambiente durante una noche, se diluyó con 30 ml de acetato de etilo y la fase orgánica se lavó con NaOH 1 N, salmuera, y después se secó sobre sulfato de magnesio y se concentró a presión reducida para dar un aceite marrón. El producto bruto se purificó por HPLC de fase inversa (C8) para dar 0,155 g de un aceite amarillo, con una pureza de aproximadamente el 70% (compuesto 6). El rendimiento era de 0,15 mmol, 43%.

Síntesis de (2R)-4-[N-(2-{2-[2-(2-aminoetoxi)etoxi]etoxi}etil)carbamoil]-2-[(4-[(2,4-diaminopteridin-6-il)metil] metilamino} fenil)carbonilamino]butanoato de terc-butilo (7)

[0371] El compuesto 6 (0,155 g, pureza del 70%, 0,15 mmol) se disolvió en 3 ml de tetrahidrofurano y se añadieron 200 ml de agua, seguidos de trifenilfosfina (130 mg, 0,5 mmol). La mezcla de reacción se agitó a temperatura ambiente durante 16 horas, se diluyó con 20 ml de éter dietílico y la fase orgánica se extrajo con ácido cítrico al 10%. La fase acuosa se neutralizó a pH 12 con NaOH 10N y se extrajo con acetato de etilo. La fase orgánica se lavó con salmuera, se secó sobre sulfato de magnesio y se concentró a presión reducida para dar un aceite. El producto bruto se purificó por HPLC de fase inversa (C8) para dar 16 mg de un aceite amarillo, 0,022 mmol, 15% (compuesto 7).

Síntesis de 4-((2,4-diamino-6-pteridinilmetil)metilamino)benzoil-L-Gln(11-(9-fluoro-11b, 17-dihidroxi-16a-metil-3-oxoandrosta-1, 4-dien-17b-carboxamido)-3,6,9-trioxoundecilo) (9, GPC285937)

[0372] Se combinó ácido 9-fluoro-11b,17-dihidroxi-16a-metil-3-oxoandrosta-1,4-dieno-17b-carboxílico (compuesto 8) (12 mg, 0,032 mmol) y compuesto 7 (15 mg, 0,021 mmol) en 0,5 ml de DMF y se añadió PyBop (20 mg, 0,038 mmol), seguido de 0,017 ml de DIEA (0,1 mmol). La mezcla de reacción se agitó a temperatura ambiente durante 16 horas y después se diluyó con 10 ml de acetato de etilo. La fase orgánica se lavó con NaOH 0,2 N y salmuera, y después se concentró a presión reducida para dar un aceite. Este aceite se disolvió en 2 ml de TFA:CH₂Cl₂ 1:1 y se dejó reposar durante 1 hora. El disolvente se retiró a presión reducida y el residuo se purificó por HPLC de fase inversa (C8) para dar 2,8 mg de producto, 0,0028 mmol, 13% (compuesto 9).

Síntesis de GPC 285937 siguiendo el Esquema 2 (Véase la Figura 1B)

Síntesis de (2S)-4-[N-(2-{2-[2-(2-azidoetoxi)etoxi]etoxi}etil)carbamoil]-2-[(4-[N-metil (fenilmetoxi) carbonilamino]fenil] carbonilamino]butanoato de terc-butilo (11)

[0373] El compuesto 4 (0,81 g, 2,0 mmol) y el ácido 4-carboxibencilmetilaminobenzoico (compuesto 10) (0,61 g, 2,1 mmol) se disolvieron en 10 ml de DMF. A esta solución, se añadió HBTU (1,0 g, 2,6 mmol) como sólido, seguido de DIEA (0,8 ml, 4,6 mmol). La mezcla de reacción se agitó durante una noche a temperatura ambiente, se diluyó con acetato de etilo y la fase orgánica se lavó con NaOH 0,5 N, salmuera, se secó sobre sulfato de magnesio y se concentró a presión reducida para dar un aceite marrón. El producto bruto se purificó por cromatografía en sílice ultrarrápida (MeOH al 5% en EtOAc) para dar un aceite marrón (1,03 g, 1,5 mmol, 77%, compuesto 11).

Síntesis de (2S)-4-[N-(2-{2-[2-(2-aminoetoxi)etoxi]etoxi}etil)carbamoil]-2-[(4-[N-metil (fenilmetoxi) carbonilamino]fenil] carbonilamino]butanoato de terc-butilo (12)

[0374] El compuesto 11 (1,0 g, 1,49 mmol) se disolvió en 50 ml de MeOH y se añadieron 130 mg de Pd/C al 10 %. La mezcla de reacción se agitó a 40 psi de hidrógeno durante 16 horas, el catalizador se eliminó por filtración y el filtrado se concentró a presión reducida para dar 0,75 g (1,47 mmol, 98%) de un aceite incoloro (compuesto 12).

Síntesis de 4-metilaminobenzoil-L-Gln(11-(9-fluoro-11b, 17-dihidroxi-16 a-metil-3-oxo androsta-1, 4-dien-17b-carboxamido)-3 6 9-trioxoundecil) terc-butil éster (13)

[0375] El compuesto 12 (0,75 g, 1,47 mmol) se disolvió en DMF con ácido 9-fluoro-11b,17-dihidroxi-16a-metil-3-oxoandrosta-1,4-dieno-17b-carboxílico (8) (0,60 g, 1,6 mmol) y a esta solución se añadió HBTU (0,75 g, 2 mmol), seguido de DIEA (0,35 ml, 2 mmol). La mezcla de reacción se agitó durante una noche a temperatura ambiente, se diluyó con acetato de etilo y la fase orgánica se lavó con bicarbonato sódico saturado, salmuera y se concentró a presión reducida para dar un aceite naranja. El producto bruto se purificó por cromatografía en sílice ultrarrápida (MeOH al 10% en EtOAc) para dar 0,54 g de una espuma blanca (0,62 mmol, 42%, compuesto 13).

Síntesis de bromhidrato de 2,4-diamino-6-(bromometil)pteridina (14)

[0376] La síntesis de bromhidrato de 2,4-diamino-6-(bromometil)pteridina (compuesto **14**) se llevó a cabo en dos etapas descritas individualmente en la bibliografía (Taghavi y Pfeleiderer, *Tetrahedron Lett.*, 1997, 38: 6835-36; Taylor y Portnoy, *J. Org. Chem.*, 1973, 38:806).

- 5 *Síntesis de 4-((2,4-diamino-6-pteridinilmetil)metilamino)benzoil-L-Gln(11-(9-fluoro-11b,17-dihidroxi-16a-metil-3-oxoandrosta-1,4-dien-17b-carboxamido)-3,6,9-trioxoundecil) terc-butil éster (15)*

[0377] El compuesto **13** (0,54g, 0,62 mmol) y 0,41 g de compuesto **14** (1,2 mmol) se combinaron en 8 ml de dimetilacetamida y se calentaron a 60°C durante 6 horas. Se añadió éter dietílico (100 ml) y se formó un precipitado. El sobrenadante se eliminó por decantación y el residuo se purificó por cromatografía en sílice (NH₄OH:MeOH:CH₂Cl₂ saturado 1:10:89) para dar 0,35 g de un sólido amarillo (0,33 mmol, 54%, compuesto **15**).

- 10 *Síntesis de 4-((2,4-diamino-6-pteridinilmetil)metilamino)benzoil-L-Gln(11-(9-fluoro-11b,17-dihidroxi-16a-metil-3-oxoandrosta-1,4-dien-17b-carboxamido)-3,6,9-trioxoundecilo) (9, GPC 285937)*

[0378] El compuesto **15** (0,35 g, 0,33 mmol) se disolvió en 20 ml (H₂O:Me₂S:CH₂Cl₂:TFA 1:1:8:10) y la reacción se agitó durante 1 hora a temperatura ambiente. El disolvente se eliminó a presión reducida y el residuo se disolvió en MeOH y se purificó por HPLC de fase inversa (C8). Las fracciones que contenían producto se concentraron hasta un volumen mínimo y después se liofilizaron para dar 0,30 g de un sólido amarillo (0,27 mmol, 83%).

Síntesis de GPC 285985 siguiendo el Esquema 3 (Véase la Figura 1C)

Síntesis de 2-metil-2-(4-[[3-(metiletil)-4-oxo-1-(2,4,6-triclorofenil)(5-hidropirazol[5,4-d]pirimidin-6-il)]metil} fenoxi) propanoato de etilo (17)

- 20 [0379] El compuesto **16** (2,5 g, 7,2 mmol) y 2-{4-[(etoxicarbonil)metil]fenoxi}-2-metilpropanoato de etilo (4,5 g, 15,3 mmol) se disolvieron en 15 ml de etanol y se añadieron 5,8 ml de una solución 2,66 M de etóxido sódico en etanol (15,3 mmol). La mezcla de reacción se calentó a reflujo durante 5 horas, se enfrió a temperatura ambiente y se dejó reposar durante una noche. La mezcla de reacción se diluyó después con acetato de etilo y se lavó con agua y salmuera, se secó sobre sulfato de magnesio, se filtró y se concentró hasta 1,6 g (2,8 mmol, 38%) de un sólido beige (compuesto **17**).

- 25 *Síntesis de ácido 2-metil-2-(4-[[3-(metiletil)-4-oxo-1-(2,4,6-triclorofenil)(5-hidropirazol[5,4-d]pirimidin-6-il)]metil} fenoxi) propanoico (18)*

[0380] El compuesto **16** (1,6 g, 2,8 mmol) se disolvió en 30 ml de dioxano, 10 ml de metanol y se trató con 5 ml (5 mmol) de NaOH 1 N. La reacción se agitó a temperatura ambiente durante una noche, después se diluyó con acetato de etilo y se lavó con HCl 1 N, y después con salmuera. La fase orgánica se secó sobre sulfato de magnesio, se filtró y se concentró hasta un sólido (1,4 g, 2,5 mmol, 91%, compuesto **18**).

- 30 *Síntesis de (2R)-2-[[4-(metilamino) fenil] carbonilamino]-4-(N-{2-[2-(2-[2-metil-2-(4-[[3-(metiletil)-4-oxo-1-(2,4,6-triclorofenil)(5-hidropirazol[5,4-d]pirimidin-6-il)]metil}fenoxi)propanoilamino]etoxi]etoxi]etoxi]etil}carbamoil)butanoato de terc-butilo (19)*

[0381] El compuesto **18** (0,70 g, 1,3 mmol) y el compuesto **12** (0,63 g, 1,2 mmol) se disolvieron en dimetil formamida y se añadió HBTU (0,75 g, 2 mmol), seguido de diisopropiletilamina (0,5 ml, 2,9 mmol). La mezcla de reacción se agitó a temperatura ambiente durante 3 días, se diluyó con acetato de etilo y después se lavó con NaOH 0,5 N y salmuera. La fase orgánica se secó sobre sulfato de magnesio, se filtró y se concentró hasta un aceite que se purificó por cromatografía en sílice ultrarrápida (MeOH/EtOAc del 5 al 10%) para dar 430 mg (0,41 mmol, 34%) de una espuma marrón (compuesto **19**).

- 40 *Síntesis de ácido (2R)-2-[[4-[[[(2,4-diaminopteridin-6-il)metil]metilamino]fenil]carbonilamino]-4-(N-{2-[2-(2-[2-metil-2-(4-[[3-(metiletil)-4-oxo-1-(2,4,6-triclorofenil)(5-hidropirazol[5,4-d]pirimidin-6-il)]metil}fenoxi)propanoilamino]etoxi]etoxi]etoxi]etil}carbamoil)butanoico (20, GPC 285985)*

[0382] El compuesto **19** (0,43 g, 0,41 mmol) se disolvió en 10 ml de dimetil acetamida y se añadieron 0,27 g de compuesto **14** (0,80 mmol) a la mezcla de reacción como un sólido. La mezcla de reacción se calentó a 60°C durante 5 horas, después se dejó enfriar a temperatura ambiente y se añadieron 100 ml de éter dietílico. El sobrenadante se eliminó por decantación, dejando un residuo marrón oscuro que se recogió en 10 ml de un cóctel de escisión (TFA:CH₂Cl₂:Me₂S:H₂O 10:10:1:1) y se agitó durante una hora. El disolvente se eliminó a presión reducida y el residuo se purificó por RPHPLC. Se combinaron las fracciones que contenían el producto, se concentraron hasta un volumen pequeño y se liofilizaron para dar un sólido amarillo (101 mg, 0,086 mmol, 21%, compuesto **20**).

- 50 *Síntesis de GPC 286004 y GPC 286026 siguiendo los Esquemas 4 y 5 (Véanse las Figuras 1D y 1E)*

Síntesis de 2-{4-[[4-nitro-1,3-dioxo-2-hidrociclopenta[3,4-a]bencen-2-il]carbonil]fenoxi}acetato de etilo (21)

[0383] El 2-[4-(4,4,4-trifluoro-3-oxobutanoil)fenoxi]acetato de etilo (31,9 g, 0,1 mol) se combinó con 19,3 g de anhídrido

3-nitroftálico (0,1 mol) y se añadieron 57 ml de anhídrido acético (0,6 mol). La suspensión semifundida se agitó a 0°C y se añadieron 28 ml (0,2 mol) de trietilamina. La mezcla de reacción se volvió homogénea y roja y se agitó a temperatura ambiente durante una noche, momento en el que se añadieron 600 ml de HCl 1 N. La suspensión pegajosa resultante se agitó durante 2 horas y el precipitado se volvió un sólido granular que se retiró por filtración, se resuspendió en 200 ml de etanol, se calentó a reflujo y después se enfrió hasta 0°C. Se obtuvo por filtración un sólido amarillo, se lavó con etanol (3 x 40 ml) y se secó hasta 12,7 g, 32 mmol, rendimiento del 32% (compuesto **21**).

Síntesis de 2-[4-[(4-amino-1,3-dioxo-2-hidrociclopenta[3,4-a]bencen-2-il)carbonil]fenoxi]acetato de etilo (22)

[0384] El compuesto **21** (12,7 g, 32 mmol) se disolvió parcialmente en 600 ml de acetato de etilo y se añadieron 1,5 g de Pd/C al 10%. La reacción se agitó en un globo de H₂ durante una noche. El globo se recargó con H₂ y se agitó durante 24 horas más. La reacción se filtró a través de celite con la ayuda de THF y CH₂Cl₂ para disolver el producto, y el filtrado se concentró hasta 10,7 g (29,1 mmol, 91%) de sólido (compuesto **22**).

Síntesis de 2-[4-[(4-[(morfolin-4-ilamino) carbonilamino]-1, 3-dioxo-2-hidrociclopenta[3,4-a]bencen-2-il)carbonil]fenoxi]acetato de etilo (23)

[0385] El compuesto **22** (6,4 g, 17,4 mmol) se combinó en acetonitrilo con morfolin-4-carboxilato de 4-nitrofenilo (que contenía 1 eq. de impureza de cloruro de trietilamonio) (8,0 g, 19,8 mmol) y se añadió dimetilaminopiridina (0,20 g, 1,6 mmol). La suspensión se calentó a reflujo durante 3 horas, se enfrió hasta 0°C y se obtuvo por filtración un sólido amarillo. Este sólido se lavó con un mínimo de acetonitrilo frío y se secó hasta 6,7 g, 13,5 mmol, 78% (compuesto **23**).

Síntesis de ácido 2-[4-[(4-[(morfolin-4-ilamino)carbonilamino]-1,3-dioxo-2-hidrociclopenta[3,4-a]bencen-2-il)carbonil]fenoxi]acético (24)

[0386] El compuesto **23** (6,7 g, 13,5 mmol) se disolvió en 200 ml de dioxano y se añadieron 20 ml de NaOH 1 N (20 mmol). La mezcla de reacción se agitó durante una hora. La suspensión blanca se diluyó con 1 l de acetato de etilo y se lavó con HCl 1 N y salmuera. La fase orgánica se secó sobre sulfato de magnesio, se filtró y se concentró hasta un sólido amarillo (6,3 g, 13,5 mmol, 100%, compuesto **24**).

Síntesis de ácido 2-[4-[5-[(morfolin-4-ilamino)carbonilamino]-4-oxoindeno[3,2-c]pirazol-3-il]fenoxi]acético (25)

[0387] El compuesto **24** (6,5 g, 13,5 mmol) se disolvió en 200 ml de THF, 100 ml de DMSO y se trató con 4 g (80 mmol) de hidrazina hidrato y 190 mg de ácido p-toluenosulfónico hidrato (1 mmol). La mezcla de reacción se calentó a 60°C durante 5 horas, se dejó enfriar a temperatura ambiente y se añadieron 600 ml de Et₂O. La suspensión resultante se filtró después, el precipitado se lavó con HCl 1 N y se secó al vacío para dar 4,0 g (8,6 mmol, 64%) de sólido amarillo (compuesto **25**).

Síntesis de (2S)-4-(N-[2-[2-(2-[2-(2-aminoetoxi) etoxi] etoxi]etoxi) etoxi] etil] carbamoil)-2-[[4-(metilamino)fenil]carbonilamino]butanoato de terc-butilo (26)

[0388] El compuesto **26** se sintetizó mediante un procedimiento análogo al empleado para el compuesto **12**, pero usando 1-amino-17-azido-3,6,9,12,15-pentaoxaheptadecano en lugar de 1-amino-11-azido-3,6,9-trioxaundecano en la primera etapa de síntesis.

Síntesis de (2S)-2-[[4-(metilamino)fenil]carbonilamino]-4-(N-[2-[2-(2-[2-(4-{5-[(N-morfolin-4-ilcarbamoil)amino]-4-oxoindeno[3,2-c]pirazol-3-il]fenoxi]acetilamino]etoxi]etoxi]etoxi]etoxi]etil]carbamoil)butanoato de terc-butilo (27)

[0389] El compuesto **12** (0,71 g, 1,4 mmol) y el compuesto **25** (0,57 g, 1,2 mmol) se disolvieron en 10 ml de DMF y se añadió HBTU (0,8 g, 2,1 mmol) como un sólido, seguido de DIEA (0,52 ml, 3 mmol). La mezcla de reacción se agitó a temperatura ambiente durante 3 días, se diluyó con EtOAc y la fase orgánica se lavó con NaHCO₃ saturado. La fase acuosa se volvió a extraer con EtOAc dos veces y las fases orgánicas combinadas se secaron sobre MgSO₄, se filtraron y se concentraron hasta un aceite. Este aceite se purificó mediante cromatografía en sílice ultrarrápida (MeOH/EtOAc del 2 al 5%) para dar un aceite naranja (0,50 g, 0,52 mmol, 44%, compuesto **27**).

Síntesis de (2S)-2-[[4-(metilamino)fenil]carbonilamino]-4-(N-[2-[2-(2-[2-(2-[2-(4-{5-[(N-morfolin-4-ilcarbamoil)amino]-4-oxoindeno[3,2-c]pirazol-3-il]fenoxi]acetilamino]etoxi]etoxi]etoxi]etoxi]etil]carbamoil)butanoato de terc-butilo (28)

[0390] El compuesto **25** (0,60 g, 1 mmol) y el compuesto **26** (0,46 g, 1 mmol) se disolvieron en 10 ml de DMF y se añadió HBTU (0,7 g, 1,8 mmol) como un sólido, seguido de DIEA (1,0 mm, 5,7 mmol). La mezcla de reacción se agitó a temperatura ambiente durante una noche, se diluyó con EtOAc y la fase orgánica se lavó con NaOH 0,5N, salmuera, se secó sobre MgSO₄, se filtró y se concentró hasta un aceite. Este aceite se purificó por cromatografía en sílice ultrarrápida (MeOH/EtOAc del 10 al 20%) para dar una espuma amarilla (0,65 g, 0,62 mmol, 62%, compuesto **28**).

Síntesis de (2S)-2-[[4-[(2,4-diaminopteridin-6-il)metil]metilamino]fenil]carbonilamino]-4-(N-[2-[2-(2-[2-(4-{5-(metoxi)carbonilamino]-4-oxoindeno[3,2-c]pirazol-3-il]fenoxi]acetilamino]etoxi]etoxi]etoxi]etil]carbamoil)butanoato de terc-butilo (29)

[0391] El compuesto **27** (0,50 g, 0,52 mmol) se disolvió en dimetilacetamida y se añadieron 0,33 g de compuesto **14** (1,0 mmol) a la mezcla de reacción como un sólido. La mezcla de reacción se calentó a 60°C durante 6 horas, después se dejó enfriar a temperatura ambiente y se añadieron 80 ml de éter dietílico. El sobrenadante se eliminó por decantación,

dejando un residuo marrón oscuro, que se purificó por cromatografía en sílice ultrarrápida (MeOH/CH₂Cl₂ del 5 al 10%, después MeOH/CH₂Cl₂ del 5 al 10% con NH₄OH al 1%) para dar 0,33 g (0,29 mmol, 56%) de un sólido amarillo (compuesto **29**).

Síntesis de (2S)-2-[(4-[(2,4-diaminopteridin-6-il)metil]metilamino)fenil]carbonilamino-4-{N-[2-(2-[2-(2-[2-(4-[5-[(morfolin-4-ilamino)carbonilamino]-4-oxoindeno[3,2-c]pirazol-3-il]fenoxi)acetilamino]etoxi]etoxi]etoxi]etil]carbamoil}butanoato de terc-butilo (30)

[0392] El compuesto **28** (0,65 g, 0,62 mmol) se disolvió en dimetilacetamida y se añadieron 0,4 g de compuesto **14** (1,2 mmol) a la mezcla de reacción como un sólido. La mezcla de reacción se calentó a 60°C durante 6 horas, después se dejó enfriar a temperatura ambiente y se añadieron 80 ml de éter dietílico, y se dejó reposar durante 3 días. El sobrenadante se eliminó por decantación, dejando un residuo marrón oscuro que se purificó por cromatografía en sílice ultrarrápida (MeOH/CH₂Cl₂ del 5 al 10%, después MeOH/CH₂Cl₂ del 5 al 10% con NH₄OH al 1%) para dar 0,45 g (0,37 mmol, 60%) de un sólido amarillo (compuesto **30**).

Síntesis de ácido (2S)-2-[(4-[(2,4-diaminopteridin-6-il)metil]metilamino)fenil]carbonil-amino-4-{N-[2-(2-[2-(2-[2-(4-[5-(metoxi-carbonil-amino)-4-oxoindeno[3,2-c]pirazol-3-il]fenoxi)acetilamino]etoxi]etoxi]etoxi]etil]carbamoil}butanoico (31, GPC 286004)

[0393] El compuesto **29** (0,33 g, 0,29 mmol) se trató con 20 ml de un cóctel de escisión (TFA:CH₂Cl₂:Me₂S:H₂O 10:10:1:1). Después de una hora, se eliminó el disolvente y el residuo se purificó por RPHPLC. Se combinaron las fracciones que contenían el producto, se concentraron hasta un volumen pequeño y se liofilizaron para dar un sólido amarillo (0,19 g, 0,18 mmol, 61% compuesto **31**).

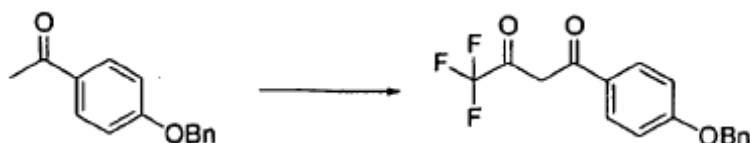
Síntesis de ácido (2S)-2-[(4-[(2,4-diaminopteridin-6-il)metil]metilamino)fenil]carbonil-amino-4-{N-[2-(2-[2-(2-[2-(4-[5-[(morfolin-4-ilamino)carbonilamino]-4-oxoindeno[3,2-c]pirazol-3-il]fenoxi)acetilamino]etoxi]etoxi]etoxi]etil]carbamoil}butanoico (32, GPC 286026)

[0394] El compuesto **30** (0,45 g, 0,37 mmol) se trató con 20 ml de un cóctel de escisión (TFA:CH₂Cl₂:Me₂S:H₂O 10:10:1:1). Después de una hora, el disolvente se eliminó y el residuo se purificó por RPHPLC. Las fracciones que contenían el producto se combinaron, se concentraron hasta un volumen pequeño y se liofilizaron para dar un sólido amarillo (0,23 g, 0,18 mmol, 49% compuesto **32**).

Síntesis de GPC 285993 siguiendo el Esquema 6 (véase la Figura 1F)

Síntesis de 1-(4-benciloxi-fenil)-4,4,4-trifluoro-butano-1,3-diona

[0395]

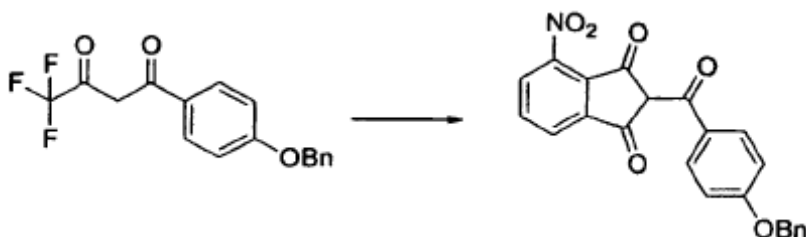


[0396] Se recogieron 45,2 g de 1-(4-benciloxi-fenil)etanona (200 mmol) en THF (250 ml) y se trataron con CF₃CO₂Et (30 ml, 250 mmol). La solución se enfrió hasta 0°C y se trató con una solución de NaOEt 2,66 M (94 ml, 250 mmol) durante 1 h. El baño en hielo se retiró y la solución se agitó a temperatura ambiente durante 4 h. La reacción se vertió en HCl 1 N (1.000 ml) y se extrajo con EtOAc (1500 ml).

[0397] La fase orgánica se lavó con salmuera, se secó y se evaporó para dar 64,2 g de 1-(4-benciloxi-fenil)-4,4,4-trifluoro-butano-1,3-diona (200 mmol, rendimiento del 100%).

Síntesis de 4-nitro-2-[(4-hidroxifenil)carbonil]-2-hidrociclopenta[1,2-a] benceno-1,3-diona (33)

[0398]



[0399] Se suspendieron 64 g de 1-(4-benciloxi-fenil)-4,4,4-trifluoro-butano-1,3-diona (200 mmol) en Ac₂O (114 ml, 1,2 mmol) y se trataron con anhídrido 3-nitroftálico (28,6 g, 200 mmol). La suspensión se enfrió hasta 0°C y se trató lentamente con Et₃N (56 ml, 400 mmol). La reacción se agitó a temperatura ambiente durante 16 h, después se vertió

en hielo/HCl 3 N (500 ml) y se agitó enérgicamente durante 1 h. El precipitado se filtró y se lavó con agua. El precipitado se suspendió en etanol en ebullición (450 ml) durante 10 min, después se enfrió hasta 0°C durante 2 h y se filtró. El sólido se lavó con etanol frío y se secó al vacío para dar 34 g (72 mmol, rendimiento del 36%, compuesto **33**).

*Síntesis de 4-amino-2-[(4-hidroxifenil)carbonil]-2-hidrociclopenta[1,2-a]benceno-1,3-diona (**34**)*

- 5 **[0400]** El compuesto **33** (32,1 g, 67,6 mmol) se disolvió en 1500 ml de EtOAc y se añadieron 3,2 g de Pd/C al 10%. La mezcla de reacción se agitó en una atmósfera (globo) de H₂ durante 3 días. Se añadió metanol para contribuir a la disolución y la mezcla de reacción se filtró a través de celite. El filtrado se concentró hasta 19 g (67 mmol, 100%) de un sólido naranja (compuesto **34**).

10 *Síntesis de N-{2-[(4-hidroxifenil)carbonil]-1,3-dioxo(2-hidrociclopenta[2,1b]bencen-4-il)}(morfolin-4-ilamino)carboxamida (**35**)*

[0401] El compuesto **34** (10,0 g, 35,3 mmol) se disolvió en acetonitrilo con morfolin-4-carboxilato de 4-nitrofenilo (que contenía 1 eq. de impureza de cloruro de trietilamonio) (13,0 g, 32,1 mmol) y se añadió dimetilaminopiridina (0,60 g, 5,4 mmol). La mezcla de reacción se calentó a reflujo durante 3 horas, se enfrió a temperatura ambiente y se obtuvo por filtración un sólido verde claro, y se secó hasta 7,5 g (18,3 mmol, 57%, compuesto **35**).

15 *Síntesis de N-[3-(4-hidroxifenil)-4-oxoindeno[3,2-c]pirazol-5-il](morfolin-4-ilamino)carboxamida (**36**)*

[0402] El compuesto **35** (7,5 g, 18,3 mmol) se suspendió en 200 ml de THF y se añadió hidrazina hidrato (4,5 g, 90 mmol), seguido de ácido p-toluenosulfónico hidrato (340 mg, 1,8 mmol). La mezcla de reacción se calentó a reflujo durante una noche (solución homogénea), se dejó enfriar a temperatura ambiente y se formó un precipitado, que se filtró para dar 1,2 g de producto. El filtrado se concentró hasta un sólido, se suspendió en EtOAc y se filtró. Este sólido se purificó por cromatografía en sílice ultrarrápida (MeOH/EtOAc del 5 al 10%) para dar 2,2 g más de producto. El rendimiento combinado era de 3,3 g, 8,4 mmol, 46% (compuesto **36**).

*Síntesis de 2-{3-(4-hidroxifenil)-5-[(morfolin-4-ilamino)carbonilamino]-4-oxoindeno[3,2-c]pirazol-2-il}acetato de etilo (**37**)*

- 25 **[0403]** El compuesto **36** (2,2 g, 5,6 mmol) se disolvió en 50 ml de acetona, 10 ml de THF y 10 ml de DMF y se añadió Cs₂CO₃ (1,8 g, 5,6 mmol), seguido de bromoacetato de etilo (0,93 g, 5,6 mmol). La mezcla de reacción se agitó durante 2 horas, se diluyó con acetato de etilo y la fase orgánica se lavó con HCl 1 N, salmuera, se secó sobre MgSO₄, se filtró y se concentró hasta un sólido amarillo. El sólido se purificó por cromatografía en sílice ultrarrápida (MeOH/CH₂Cl₂ del 2 al 3 al 4%) para dar 1,2 g (2,4 mmol, 44%) de un sólido amarillo (compuesto **37**).

*Síntesis de ácido 2-{3-(4-hidroxifenil)-5-[(morfolin-4-ilamino)carbonilamino]-4-oxoindeno[3,2-c]pirazol-2-il}acético (**38**)*

- 30 **[0404]** El compuesto **37** (1,2 g, 2,4 mmol) se disolvió en 60 ml de dioxano:etanol:DMSO 3:2:1 y se añadieron 12 ml de NaOH 0,5 N y la reacción se volvió roja. La mezcla de reacción se agitó a temperatura ambiente durante una hora, se diluyó con EtOAc y se lavó con HCl 1 N. La fase acosa se volvió a extraer una vez con acetato de etilo y las fases orgánicas combinadas se secaron sobre MgSO₄ y se concentraron hasta un sólido naranja. El sólido se trituró con 10 ml de MeOH/100 ml de Et₂O, se filtró y se secó hasta un sólido (1,1 g, 2,4 mmol, 100% compuesto **38**).

35 *Síntesis de (2S)-4-{N-[2-(2-{2-(2-{3-(4-hidroxifenil)-5-[(N-morfolin-4-ilcarbamoil)amino]-4-oxoindeno[3,2-c]pirazol-2-il}acetilamino)etoxi}etoxi)etil]carbamoil}-2-{[4-(metilamino)fenil]carbonilamino}butanoato de terc-butilo (**39**)*

[0405] El compuesto **38** (0,52 g, 1,1 mmol) y el compuesto **12** (0,55 g, 1,1 mmol) se disolvieron en DMF y se añadió HBTU (0,8 g, 2,1 mmol) como un sólido, seguido de DIEA (0,52 ml, 3 mmol). La mezcla de reacción se agitó a temperatura ambiente durante una noche, se diluyó con EtOAc y la fase orgánica se lavó con NaHCO₃ saturada, salmuera, se secó sobre MgSO₄, se filtró y se concentró hasta un aceite. Este aceite se purificó por cromatografía en sílice ultrarrápida (MeOH/CH₂Cl₂ del 1 al 2 al 3 al 4 al 5%) para dar una espuma amarilla (0,45 g, 0,47 mmol, 43%, compuesto **39**).

*Síntesis de (2S)-2-{[4-{[(2,4-diaminopteridin-6-il)metil]metilamino}fenil]carbonilamino}-4-{N-[2-(2-{2-(2-{3-(4-hidroxifenil)-5-[(N-morfolin-4-ilcarbamoil)amino]-4-oxoindeno[3,2-c]pirazol-2-il}acetilamino)etoxi}etoxi)etil]carbamoil}butanoato de terc-butilo (**40**)*

- 45 **[0406]** El compuesto **39** (0,45 g, 0,47 mmol) se disolvió en 8 ml de dimetilacetamida y se añadieron 0,2 g de compuesto **14** (0,60 mmol) a la mezcla de reacción como un sólido. La mezcla de reacción se calentó a 60°C durante 6 horas, después se dejó enfriar a temperatura ambiente y se añadió éter dietílico. El sobrenadante se eliminó por decantación, dejando un residuo marrón oscuro, que se purificó por cromatografía en sílice ultrarrápida (MeOH/CH₂Cl₂ del 5 al 10%, después MeOH/CH₂Cl₂ del 5 al 10% con NH₄OH al 1%) para dar 0,32 g (0,27 mmol, 56%) de sólido amarillo (compuesto **40**).

*Síntesis de ácido (2S)-2-{[4-{[(2,4-diaminopteridin-6-il)metil]metilamino}fenil]carbonilamino}-4-{N-[2-(2-{2-(2-{3-(4-hidroxifenil)-5-[(N-morfolin-4-ilcarbamoil)amino]-4-oxoindeno[3,2-c]pirazol-2-il}acetilamino)etoxi}etoxi)etil]carbamoil}butanoico (**41**, GPC 285993)*

- 55 **[0407]** El compuesto **40** (0,30 g, 0,27 mmol) se trató con 20 ml de un cóctel de escisión (TFA:CH₂Cl₂:Me₂S:H₂O 10:10:1:1). Después de una hora, se eliminó el disolvente y el residuo se purificó por RPHPLC. Las fracciones que

contenían el producto se combinaron, se concentraron hasta un volumen pequeño y se liofilizaron para dar un sólido amarillo (78 mg, 0,073 mmol, 27%, compuesto **41**).

Ejemplo 2: Medición de afinidades de ligandos híbridos para proteínas de unión seleccionadas

[0408] Para demostrar la caracterización de la afinidad entre ligandos híbridos y proteínas a las que se unen, se analizó la unión de GPC 285985 a sus compañeros de unión esperados DHFR y CDK2/E (complejo de quinasa dependiente de ciclina 2/ciclina E). El análisis se realizó en un BIACORE 2000 SPR-Biosensor (Biacore, Uppsala, Suecia) a 22°C usando un tampón de procesamiento que contenía HEPES 20 mM (pH 7,4), NaCl 150 mM, DTT 1 mM y Tween20 al 0,005% (Protein Grade, Calbiochem). El vector pQE40 (Quiagen, Hilden, Alemania), que comprendía el gen que codifica DHFR fusionado a una etiqueta His₆, se usó para transformar *E. coli* y la proteína de fusión His₆-DHFR se purificó siguiendo los protocolos de los fabricantes. La His₆-DHFR se acopló posteriormente a pH 4,6 a la superficie de dextrano de una microplaca detectora CM5 (Biacore, Uppsala, Suecia; Research Grade) de acuerdo con las instrucciones de los fabricantes. La densidad de carga alcanzó 1100 UR (Unidades de Resonancia). Se dejó que una solución 10 µM de GPC 285985 pasara sobre la superficie de la microplaca cargada con DHFR durante 5 minutos a un caudal de 30 µl/min, seguido de 5 minutos de tampón de procesamiento al mismo caudal. Se obtuvo y se almacenó un perfil para la adsorción y desorción de GPC 285985 sobre DHFR. Se evaluó la unión inespecífica de GPC 285985 usando una superficie de CM5 con grupos COOH desactivados. El sensograma resultante (no se muestra) demostró la unión específica y de alta afinidad del ligando híbrido a la superficie recubierta con DHFR.

[0409] Para caracterizar la unión de GPC 285985 a otras proteínas, la superficie de CM5-DHFR se cargó primero con GPC 285985 pasando una solución 10 µM de GPC 285985 sobre la superficie de la microplaca durante 5 minutos a un caudal de 10 µl/min. Después, se diluyó complejo de CDK2/E, por ejemplo purificado a partir de células infectadas con baculovirus que expresaban CDK2 y Ciclina E (Sarcevic et al., J. Biol. Chem., 1997 272: 33327- 37) en tampón de procesamiento para obtener ocho concentraciones de proteína diferentes que variaban de 6 nM a 750 nM, que después se dejó que pasaran cada una sobre la superficie detectora consecutivamente, durante 5 min cada una, seguido de 5 min de tampón de procesamiento al mismo caudal. La asociación y disociación del complejo de CDK2/E sobre superficie de microplaca cargada con CM5-DHFR::GPC 285985 se midió a un caudal de 30 µl/min. Después de cada experimento de asociación/disociación, la microplaca se regeneró para eliminar la proteína unida mediante dos inyecciones consecutivas de clorhidrato de guanidinio 3 M (20 s, 30 µl/min) antes de que se cargara la siguiente muestra. Se evaluó la unión inespecífica usando una superficie de CM5 cargada con DHFR solamente.

[0410] Los datos se analizaron usando el software Bioevaluation versión 3.1 (Biacore AB, Uppsala, SE). Las curvas se normalizaron al inicio de la inyección, y se restó la unión inespecífica a la superficie de control cargada con DHFR y la desviación de la línea de fondo resultante de la desorción de GPC 285985 de la CM5-DHFR durante el procesamiento de 10 min. Las velocidades de asociación y disociación se determinaron por separado o en conjunto usando un modelo de unión de Langmuir 1:1, como se proporciona por el software Bioevaluation 3.1. Las afinidades (K_D) se calcularon usando la ecuación:

$$K_D = k_{dis}/k_{as}$$

[0411] Este experimento de asociación/disociación dio una K_D de 8,0 nM para la unión de GPC 285985 a CDK2, confirmando la alta especificidad del ligando híbrido GPC 285985 por CDK2. La Figura 2 muestra como ejemplo los resultados de un experimento de asociación/disociación análogo obtenidos para la unión de CDK4/D1 a la microplaca cargada con CM5-DHFR::GPC 285985. La K_D para la unión de GPC 285985 al complejo de CDK4/D1 se calculó a partir de estos datos como de 920 nM. Esto confirma los resultados esperados de la fuerte unión de GPC 285985 a DHFR y CDK2, pero una débil unión a la quinasa estrechamente relacionada CDK4. El complejo de CDK4/CiclinaD1 se purificó, por ejemplo, a partir de células infectadas con baculovirus que lo expresaban (Konstantinidis et al., J. Biol. Chem., 1998, 273: 26506-15).

Ejemplo 3: Construcción de construcciones genéticas y cepas de levadura para un experimento de triple híbrido de levadura que emplea un sistema de interacción basado en transcripción

[0412] Se demostró un experimento de triple híbrido de levadura que emplea un sistema de interacción basado en transcripción utilizando una cepa de levadura que comprendía tres construcciones genéticas: una primera construcción que codifica una proteína de fusión que comprende un dominio de unión a ADN (BD) y una primera proteína o péptido (P1) capaz de unirse específicamente al primer ligando R1 del ligando híbrido previsto R1-Y-R2; una segunda construcción que codifica una proteína de fusión que comprende un dominio de activación de la transcripción (AD) y una segunda proteína o péptido, o una biblioteca de segundas proteínas o péptidos, (P2), capaz de o que se sospecha que se une al segundo ligando R2 de dicho ligando híbrido previsto; una tercera construcción que comprende un gen indicador bajo el control transcripcional de un promotor que comprende la secuencia genética a la que es capaz de unirse el BD, en la que el AD debe ser capaz de iniciar la transcripción del gen indicador cuando se pone en la proximidad espacial del promotor mediante una interacción del ligando híbrido de formación de un puente entre la proteína de fusión que comprende el BD y la proteína de fusión que comprende el AD.

[0413] Se construyeron dos plásmidos: el primer plásmido, que contenía un fragmento que codifica el dominio de unión a LexA bacteriano para su expresión como una fusión con una primera proteína; el segundo plásmido, que contenía un fragmento que codifica el dominio de activación de la transcripción GAL4 de levadura para su expresión como una

- fusión con una segunda proteína. Estos plásmidos se transformaron en células de levadura deficientes en el locus *HIS3* endógeno, pero que comprendían una construcción genética que combinaba un gen *his3* recombinante con un promotor que contenía la secuencia de unión a LexA. Puesto que se seleccionó el metotrexato como el primer ligando R1 en las presentes investigaciones, la secuencia que codifica el BD de LexA se fusionó al gen que codifica la dihidrofolato reductasa de *E. coli* (*folA*). La secuencia que codifica el dominio de activación de la transcripción GAL4 se fusionó al gen que codifica el receptor de glucocorticoides de rata de unión a dexametasona *gr2*, los genes para *cdk2* (*hcdk2*) o *cdk4* (*hcdk4*) humanos o una biblioteca de genes a partir de una genoteca de ADNc de cerebro humano, dependiendo de la elección de R2.
- [0414] Se seleccionó la cepa de levadura L40 (Invitrogen; MATa, *his3-Δ200*, *trp1-901*, *leu2-3,112*, *ade2*, *LYS2::(lexAop)4-HIS3*, *URA3::(lexAop)8-LacZ*, *gal80*) para los experimentos en levaduras descritos en la presente memoria. Sin embargo, pueden seleccionarse otras cepas de levadura adecuadas, o incluso otros tipos celulares, tales como bacterias, células de insecto, células vegetales o células de mamífero para los métodos de la invención, con tal de que las células comprendan un sistema indicador que permita una lectura detectable que depende de la formación de un complejo trimérico del ligando híbrido junto con la primera y segunda proteínas de fusión.
- [0415] Para el plásmido de fusión con dominio de unión de ADN, la secuencia codificante de *folA* (dihidrofolato reductasa, DHFR) de *E. coli* se amplificó por PCR a partir de una biblioteca genómica (Clonotech, N° Cat. XL4001AB) usando los cebadores
- 5'-GGGGTCGACATGATCAGTCTGATTGCGGCGTTAGCG-3', y
- 5'-GGGGGCGGCCGCTTACCGCCGCTCCAGAATCTCAAAG-3'.
- [0416] El producto de PCR se dirigió con Sall y NotI y el fragmento de 479 pb resultante se subclonó en pBTM118c, que contenía *TRP1* como marcador de selección en levadura (véase Wanker et al., documento WO 99/31509), dando como resultado la construcción pBTM118c-DHFR.
- [0417] Para el plásmido de fusión del dominio de activación que comprende el receptor de glucocorticoides de rata, se amplificó por PCR un fragmento génico que codifica los aminoácidos 524-795 del receptor de glucocorticoides de rata a partir de una genoteca de ADNc de cerebro de rata (Life Technologies, N° Cat.: 10653-012) usando los cebadores:
- 5'-GGGGTCGACATGGGTGGTGGTGGTGGTGGTGCAGGAGTCTCACAAGAC-3', y
- 5'-GGGGGCGGCCGCTTTTTGATGAAACAGAAG-3'.
- [0418] El producto de PCR se dirigió con Sall y NotI, y el fragmento de 813 pb resultante se subclonó en pGAD426c, que contenía LEU2 como marcador de selección en levadura (Wanker et al., documento WO 99/31509). Posteriormente, los aminoácidos F620 y C656 de GR2 se sustituyeron con Ser y Gly, respectivamente, para aumentar la afinidad de GR2 por dexametasona (Chakraborti et al., 1991, J. Biol. Chem., 266: 22075-22078), usando una reacción de PCR de mutagénesis dirigida. La mutagénesis se realizó empleando el "kit de mutagénesis dirigida de QuickChange" (Stratagene, Amsterdam, Países Bajos) de acuerdo con los protocolos de los fabricantes. La presencia de estas mutaciones se confirmó por secuenciación. La construcción resultante se denominó pGAD426c-GR2.
- [0419] Para la fusión de dominio de activación que comprendía *hcdk2*, el ADNc que codifica la hCDK2 se amplificó por PCR a partir de la genoteca de ADNc de placenta humana MATCHMAKER (Clontech, N° Cat. HL4025AH, Heidelberg, Alemania) usando los cebadores
- 5'-GGGTTCGACGCATGGAGAACTTCC-3' y
- 5'-GGGCGGCCGCTCAGAGTCGAAG-3'.
- [0420] De forma similar, se amplificó por PCR el ADNc de *hcdk4* usando los cebadores:
- 5'-GGGTTCGACGCATGGCTACCTCTCG-3' y
- 5'-GGGCGGCCGCTCAGGCTGTATTCAGC-3'.
- [0421] Después de la digestión de los productos de PCR con Sall y NotI, los fragmentos de 894 pb (CDK2) y 909 pb (CDK4) resultantes se subclonaron individualmente en pGAD426c y las secuencias de los clones se verificaron por secuenciación de ADN. Las construcciones resultantes se denominaron pGAD426c-hCDK2 y pGAD426c-hCDK4, respectivamente.
- [0422] Una genoteca de ADNc de cerebro fetal humano fusionados al gen que codifica el dominio de activación de GAL4, clonados en el vector pACT2 (Clontech, N° Cat.: HY4004AH; véase la Figura 17), que lleva LEU2 como marcador de selección en levadura, se usó según se adquiere para los experimentos de selección de clones en levaduras que se describen en el Ejemplo 10.

Ejemplo 4: Ensayo de crecimiento en halo

[0423] Se realizó un ensayo de crecimiento en halo para ensayar la capacidad de dimerización de ligandos híbridos de

la invención. La Figura 4 a. muestra un crecimiento en halo en una placa de petri en la que se ha aplicado puntualmente GPC 285937. La dimerización de las proteínas de fusión de LexA-dominio de unión a ADN (LexA-BD)-DHFR y GAL4-dominio de activación de la transcripción (GAL4-AD)-GR2 en presencia de GPC 285937 en la cepa de levadura L40 causaba la transcripción del gen indicador *His3*. Esta expresión transcripcional de *HIS3* permitía a las células de levadura superar la ausencia de histidina en el medio, conduciendo a un crecimiento celular en el área hasta el que había difundido suficiente GPC 285937 desde el centro de la placa. Por el contrario, no aparece ningún crecimiento visible en el disco de control en el que se ha aplicado puntualmente sólo DMSO, mostrado en la Figura 4 b.

[0424] Para realizar el ensayo de halo, los plásmidos pGAD426c-GR2 y pBTM118c-DHFR se usaron para cotransformar la cepa de levadura L40 usando métodos de levadura convencionales (Burke et al., *Methods in yeast genetics: A Cold Spring Harbor Laboratory course manual*; Cold Spring Harbor Laboratory Press, 2000). Los transformantes que recibieron ambos plásmidos se seleccionaron en medios que carecían de trp y leu. Después, se inocularon colonias individuales y se incubaron en medio SD líquido durante 24 h. Los cultivos se diluyeron a una densidad de 10^6 células/ml y se sembraron 100 μ l en una placa de petri de 10 cm que contenía medio SD que carecía de trp, leu e his. Se aplicó puntualmente 1 μ l de una solución 1 mM de GPC 285937 disuelto en DMSO o 1 μ l de DMSO como control en el centro de cada placa de petri. El crecimiento de células de levadura se determinó después de 2 días de cultivo a 30°C.

Ejemplo 5: El ensayo de crecimiento por detección de la fluorescencia

[0425] Para demostrar la idoneidad del ensayo de crecimiento por detección de la fluorescencia que emplea la PreSens Precision Sensing GmbH (Regensburg, Alemania) OxoPlate, se realizó un experimento análogo al Ejemplo 4. Se transformaron células de levadura con el plásmido que codifica la proteína de fusión de DHFR-dominio de unión a ADN de LexA y el plásmido que codifica hCDK2 o hCDK4 fusionado al dominio de activación de GAL4. Las células de la cepa resultante se sembraron en pocillos de una OxoPlate y se expusieron a una de cuatro condiciones: 1) medio SD que carecía de leu y trp (control positivo); 2) medio SD que carecía de leu, trp e his (control negativo); 3) medio SD que carecía de leu, trp e his y complementado con una variedad de concentraciones (de 1 mM a 4 μ M) de GPC 285985, un compuesto que se sabe que se une fuertemente a DHFR y hCDK2, pero sólo débilmente a hCDK4; 4) medio SD que carecía de leu, trp e his y complementado con GPC 285993 1 mM, un compuesto que se sabe que se une fuertemente a DHFR, pero no a hCDK2 o hCDK4 (control de selectividad de compuesto).

[0426] Los resultados obtenidos en este experimento se representan en la Figura 8 y, como se esperaba, no se observó consumo de oxígeno debido al crecimiento de células en los controles negativos o los controles de selectividad de compuesto. Por el contrario, se observó crecimiento en los controles positivos y en las células transformadas con la construcción que codifica la proteína de fusión de hCDK2 a todas las concentraciones de GPC 285985, aunque el inicio del crecimiento se retrasaba ligeramente a las menores concentraciones de GPC 285985. Las células transformadas con la construcción que codifica la proteína de fusión de hCDK4 crecían sólo cuando se exponían a una alta concentración (1 mM) de GPC 285985, confirmando adicionalmente la especificidad de la unión de este compuesto de ligando híbrido a hCDK2.

[0427] El ensayo fluorescente se realizó de la forma siguiente: En primer lugar, se cotransformaron células de la cepa de levadura L40 con pBTM118c-DHFR y uno de pGAD426c-hCDK2 o pGAD426c-hCDK4 usando técnicas convencionales (Burke et al., *Methods in yeast genetics: A Cold Spring Harbor Laboratory course manual*; Cold Spring Harbor Laboratory Press, 2000). Se seleccionaron transformantes que contenían ambos plásmidos en medios SD que carecían de trp y leu y se inocularon colonias individuales en medio SD líquido, y se incubaron durante 48 h a 30°C. En segundo lugar, las células se precipitaron y se lavaron con agua estéril 3 veces, el número de células se ajustó a una densidad de 10^8 células/ml y se transfirieron 50 μ l a cada pocillo de una OxoPlate F96 (PreSens Precision Sensing GmbH, Regensburg). Se añadieron 150 μ l de una solución que representaba una de cuatro condiciones: 1) medio SD que carecía de leu, trp e his (pocillos A1-F1, control negativo); 2) SD -leu, -trp (pocillos A2-F2, control positivo), 3) medio SD que carecía de leu, trp e his complementado con el compuesto GPC285985 a concentraciones de 1 mM, 0,5 mM, 0,25 mM, 125 μ M, 63 μ M, 31 μ M, 16 μ M, 8 μ M o 4 μ M (pocillos A3-F11); 4) medio SD que carecía de leu, trp e his complementado con 1 mM del compuesto de control GPC285993 (A12-F12, control de selectividad de compuesto). En tercer lugar, se controló el consumo de oxígeno de células de levadura en cultivo en función de la relación de emisiones fluorescentes de un primer colorante fluorescente que era inactivable por oxígeno (emisión a 590 nm) y un segundo colorante no inactivable por oxígeno (emisión a 640 nm). Esta relación de fluorescencia se controló durante 18 horas en intervalos de 20 min a 30°C usando un contador HTS multimarcador Perkin Elmer Wallac Victor2 V 1420 (Perkin Elmer, Wellesley, MA, Estados Unidos) con un ajuste de la excitación de 540 nm y un ajuste de la emisión de 590/640 nm (modo de cinética doble).

Ejemplo 6: Ensayo de compuestos de ligando híbridos para determinar efectos no relacionados con la dimerización

[0428] Los efectos de compuestos de ligando híbridos independientes de su acción de dimerización sobre las células usadas para un ensayo pueden invalidar los resultados de los ensayos que emplean estos compuestos. Dichos efectos pueden ser, por ejemplo, toxicidad o estimulación del crecimiento mediante vías distintas de la ausencia de, o producción inducida de, leucina, triptófano y/o histidina en los ensayos descritos anteriormente. Por lo tanto, el efecto *in vivo* de los ligandos híbridos se determinó en un ensayo de crecimiento en halo como se describe en el Ejemplo 4, pero

usando pGAD426c vacío (es decir, que no contenía el gen *gr* subclonado y, por lo tanto, que carecía de un segundo ligando P2 para unirse a R2) en lugar de pGAD426c-GR. Se usó 1 μ l de cada una de una serie de diluciones de los ligandos híbridos (de 10 mM a 1 μ M en DMSO) para aplicar puntualmente en el centro de placas de petri preparadas para contener medio que carecía de trp y leu, o trp, leu e his, y sembradas con células de levadura L40 que contenían los plásmidos pGAD426c y pBTM118c-DHFR. Se controló el crecimiento después de dos días de incubación a 30°C. Se espera que las células crezcan independientemente de la concentración del compuesto del ligando híbrido en medios que carecen sólo de trp y leu, aunque no debería aparecer crecimiento en medios que carecen de trp, leu e his. El comportamiento esperado se observó con todos los compuestos de ligando híbridos usados en la presente memoria a todas las concentraciones ensayadas.

Ejemplo 7: Funcionalidad mejorada de los ligandos híbridos de dimerización de la presente invención sobre el estado de la técnica.

[0429] Para comparar el Mtx-mdbt-Dex (Lin et al., J. Am. Chem. Soc. 2000, 122: 4247-8) con el Mtx-(etilenglicol)₃-Dex (GPC 285937) en un ensayo de triple híbrido de levadura, se prepararon primero diluciones de ambos compuestos en medio SD líquido que carecía de his, trp y leu, en un intervalo de concentraciones de 1 mM a 1 μ M, añadiendo la cantidad apropiada de compuesto disuelto en DMSO al medio. En segundo lugar, se transformaron células de levadura L40 con los plásmidos pBTM118c-DHFR y pGAD426c-GR2 y se inocularon en los medios que contenían los compuestos en diferentes cantidades a una densidad de 0,1 DO₅₉₅. Se controló el crecimiento durante 48 horas midiendo la DO₅₉₅ en un contador HTS multimarcador Perkin Elmer Wallac Victor2 V 1420 (Perkin Elmer, Wellesley, MA, Estados Unidos). Parecía que la cepa de levadura crecía en una ventana de entre 25 y 400 μ M, mostrando un crecimiento óptimo a GPC 285937 100 μ M (no se muestran los datos). Sin embargo, a estas concentraciones, el Mtx-mdbt-Dex mostraba una gran precipitación en el medio (véase la Figura 5). Esta precipitación puede causar que el compuesto sea menos biodisponible y, por lo tanto, que se altere el crecimiento de células de levadura en presencia de este compuesto.

[0430] Las ventajas funcionales de un ligando híbrido de la invención, Mtx-(etilenglicol)₃-Dex (GPC 285937) sobre el compuesto de la técnica anterior Mtx-mdbt-Dex se mostraron adicionalmente en un ensayo de halo de la forma siguiente. En primer lugar, se transformó la cepa de levadura L40 con los plásmidos pBTM118c-DHFR y pGAD426c-GR2 y los transformantes que contenían ambos plásmidos se seleccionaron en medios que carecían de trp y leu. En segundo lugar, se inocularon colonias individuales en medio SD líquido y se incubaron durante 24 horas. Los cultivos celulares se diluyeron hasta una densidad de 10⁶ células/ml y se sembraron 100 μ l en una placa de petri de 10 cm que contenía medio SD que carecía de trp, leu e his. En tercer lugar, se aplicó puntualmente 1 μ l de una solución 1 mM de GPC 285937 (tres unidades de etilenglicol como enlazador) o Mtx-mdbt-Dex (metadibenzotioéster como enlazador) disuelto en DMSO en el centro de cada placa de petri. El crecimiento de células de levadura se determinó después de 2 días de crecimiento a 30°C.

[0431] La Figura 6 a. muestra el halo de crecimiento que se desarrollaba alrededor del punto de aplicación de GPC 285937, mientras que la Figura 6 b presenta el mismo resultado para Mtx-mdbt-Dex. El halo de crecimiento de células de levadura que recibían Mtx-mdbt-Dex era mucho menor que el del ligando híbrido de la invención, demostrando además la superioridad del último.

[0432] Un ligando híbrido de la invención también mostraba una mejora significativa sobre el ligando híbrido de la técnica anterior en condiciones apropiadas para la exploración de una biblioteca de células de levadura. La cepa de levadura L40 se cotransformó con los plásmidos pBTM118c-DHFR y pGAD426c-GR2. Los transformantes que contenían ambos plásmidos se seleccionaron en medios que carecían de trp y leu, y se inocularon colonias individuales en medio SD líquido y se incubaron durante 24 h. Estos cultivos celulares se diluyeron hasta una densidad de 10⁴ células/ml y se sembraron 2 x 10⁴ células en placas de 22 x 22 cm que contenían medio de agar sintético de levadura que carecía de his, trp y leu, pero que contenía GPC 285937 o Mtx-mdbt-Dex 200 μ M. Se controló el crecimiento de colonias individuales después de 48 h a 30°C. Las colonias que crecían en medios SD con Mtx-mdbt-Dex eran difícilmente detectables, mientras que los clones crecían visiblemente mejor en medios que contenían GPC 285937, un ligando híbrido de la invención (Figura 7).

Ejemplo 8: Ventajas de diferentes realizaciones de los ligandos híbridos de dimerización de la presente invención

[0433] Para ciertas moléculas pequeñas, propiedades físico-químicas particulares tales como la solubilidad pueden requerir una elección particular de enlazador a usar para generar ligandos híbridos particularmente ventajosos de la estructura general R1-Y-R2. Por ejemplo, la biodisponibilidad y, por lo tanto, la actividad biológica pueden aumentarse por adición de repeticiones (-CH₂-X-CH₂) adicionales al enlazador Y. Este era el fundamento detrás de la síntesis de los ligandos híbridos GPC 286004, que comprende un enlazador de (etilenglicol)₃, y GPC 286026, que comprende un enlazador de (etilenglicol)₅. El plásmido pGAD426c-hCDK2 se cotransformó con pBTM118c-DHFR en la cepa de levadura L40. Se seleccionaron los transformantes que contenían ambos plásmidos en medios que carecían de trp y leu, y se inocularon colonias individuales en medio SD líquido y se incubaron durante 24 h. Estos cultivos se diluyeron 1:10 y se aplicaron puntualmente 20 μ l del cultivo diluido por duplicado en una placa de petri de 10 cm que contenía medio SD que carecía de trp, leu e his. Se aplicó puntualmente 1 μ l de una solución 1 mM de GPC 286004 o GPC 286026 disuelto en DMSO en el centro de cada aplicación puntual. El crecimiento de células de levadura se determinó

después de 3 días de cultivo a 30°C. Los resultados de este ensayo de halo muestran que después de 3 días en medio que carecía de leu, trp e his, sólo se observaba crecimiento en halo en presencia de GPC 286026 (cinco unidades de etilenglicol como enlazador; Figura 16 b.), pero no en presencia de GPC 286004 (tres unidades de etilenglicol como enlazador; Figura 16 a.). Esto demostraba la idoneidad superior del grupo enlazador de (etilenglicol)₅ sobre el grupo enlazador de (etilenglicol)₃ cuando unen estos dos compuestos particulares para formar un ligando híbrido.

Ejemplo 9: Métodos de ensayo de un polipéptido para determinar su unión a un ligando especificado por el usuario: un sistema de ensayo de triple híbrido basado en un sistema indicador que usa la activación de la transcripción

[0434] En ciertas realizaciones, los métodos de la invención se usan para ensayar polipéptidos para determinar su capacidad para unirse a un ligando especificado por el usuario. Para demostrar este concepto, se diseñó primero un experimento de triple híbrido que usaba un compuesto molecular pequeño para distinguir entre dos polipéptidos. Se sabe que el primer polipéptido se une con gran afinidad al compuesto molecular pequeño, mientras que se sabe que el segundo polipéptido se une al compuesto molecular pequeño sólo débilmente. Con este fin, dicho compuesto molecular pequeño se integró en un ligando híbrido de la invención y se usó en una exploración de triple híbrido con un sistema de interacción basado en transcripción.

[0435] Se desarrolló un resto de hidropirazol-pirimidina por GPC como inhibidor selectivo de hCDK2. Se une con gran afinidad a hCDK2, pero sólo débilmente a hCDK4, como puede determinarse por ejemplo usando un método análogo al Ejemplo 4. Cuando se une mediante un enlazador (-CH₂-O-CH₂)₃ a metotrexato (GPC 285985), se esperaría que el ligando híbrido resultante se uniera a y formara un puente con una combinación de proteínas de fusión de BD-DHFR y hCDK2-AD y, por consiguiente, activase un gen indicador controlado por *lexA*. Sin embargo, el mismo ligando híbrido no debería ser capaz de unirse a y formar un puente con la combinación de proteínas de fusión de BD-DHFR y hCDK4-AD cuando se usase a concentraciones de trabajo. Para ensayar esta hipótesis, las células de la cepa de levadura L40 se cotransfectaron con pBTM118c-DHFR y pGAD426c-hCDK2 o pGAD426c-hCDK4 según fuera apropiado. Los transformantes que recibieron ambos plásmidos se seleccionaron en medios que carecían de trp y leu, y se inocularon colonias individuales en medio SD líquido y se incubaron durante 24 horas. Estos dos cultivos de cepas de levadura se diluyeron hasta una densidad de 10⁶ células/ml y se sembraron 100 µl de cada cultivo diluido en una placa de petri de 10 cm que contenía medio SD que carecía de trp, leu, y también en una placa de petri de 10 cm que contenía medio SD que carecía de trp, leu e his. Se aplicó puntualmente 1 µl de una solución 1 mM de GPC 285985 disuelto en DMSO o 1 µl de DMSO como control en el centro de cada placa de petri. El crecimiento de las células de levadura se determinó después de 2 días de cultivo a 30°C (Figura 10), cuando se observó crecimiento en medio que carecía de leu, trp e his solamente para células que contenían pGAD426c-hCDK2. Después de 6 días, las células que contenían pGAD426c-hCDK2 habían cubierto completamente la placa de petri, mientras que se observó un crecimiento muy escaso en células que contenían pGAD426c-hCDK4 (Figura 11). Esto concuerda con las afinidades relativas de GPC 285985 por hCDK2 y hCDK4, y demuestra un método de ensayo de la capacidad de un polipéptido para unirse a un ligando especificado por el usuario.

Ejemplo 10: Métodos de identificación de un polipéptido que se une a un ligando especificado por el usuario: un sistema de ensayo de triple híbrido basado en un sistema de interacción basado en la transcripción

[0436] Para demostrar la idoneidad de ciertos métodos de la invención para la identificación de polipéptidos que se unen a un ligando especificado por el usuario a partir de grandes colecciones de polipéptidos candidatos, se llevó a cabo una exploración genética usando tres moléculas híbridas: en primer lugar, GPC 285985, un ligando híbrido de la invención; en segundo lugar, una proteína de fusión de BD-DHFR capaz de unirse al resto de metotrexato en GPC 285985 y de unirse al promotor de *lexA*; en tercer lugar, una genoteca de ADNc de cerebro fetal humano fusionados al GAL4-AD. Como control negativo, se usó un ligando híbrido alternativo que comprendía una molécula pequeña unida a metotrexato mediante un enlazador (-CH₂-O-CH₂)₃, de modo que fuera incapaz de unirse a hCDK2 (GPC 285985), para confirmar el crecimiento específico de compuesto.

[0437] La exploración de triple híbrido de la invención se realizó de la forma siguiente. En primer lugar, se transformaron células de la cepa de levadura L40 con pBTM118c-DHFR, y los transformantes que recibieron el plásmido se seleccionaron en medio sintético que carecía de triptófano. En segundo lugar, se volvieron a cultivar colonias individuales en medios líquidos, se volvieron competentes y las células L40 que contenían pBTM118c-DHFR se transformaron con una genoteca de ADNc de cerebro fetal humano clonada en el vector pACT2 (Clontech, N° Cat.: HY4004AH). Se seleccionaron 1 x 10⁷ colonias individuales en 60 placas de agar SD de 22 x 22 cm que carecían de trp y leu. Después de tres días de cultivo a 30°C, las colonias se retiraron por lavado de las placas, se mezclaron y se congelaron en alícuotas pequeñas. Se sembraron 2 x 10⁶ células en cada una de las 18 placas de SD que contenían medios que carecían de trp, leu e his, pero que contenían 20 µM de GPC 285985, y se incubaron durante 2-5 días. Aparecieron un total de 2811 colonias y se escogieron en placas de microtitulación de 384 pocillos que contenían medio SD que carecía de trp y leu. Todos los clones se ensayaron en un ensayo de halo de alto rendimiento contra GPC 285985 disuelto en DMSO como promotor del crecimiento, o GPC 285993 disuelto en DMSO, o DMSO puro (LTH) como control negativo. Este ensayo de halo era análogo al descrito en el Ejemplo 4, excepto por que se ensayaron múltiples ensayos diferentes (entre 10 y 1000) únicos o repetidos en bandejas de agar de 22 x 22 cm que contenían medios de cultivo apropiados. Se depositaron cepas de levadura de ensayo y de control, o ligandos híbridos/compuestos de ensayo y control en el agar en un patrón regular (entre 3 y 50 mm de separación) usando un robot de pipeteo de

laboratorio convencional (Multiprobe II, Packard, Estados Unidos). La Figura 12 muestra un ejemplo del análisis realizado. Los clones que eran capaces de crecer tras la aplicación puntual con GPC 285993 o DMSO solamente se desecharon. Aproximadamente 10^2 clones mostraron crecimiento solamente tras la aplicación puntual de GPC 285985. Estos clones se recuperaron y se identificaron por secuenciación de ADN y comprendían clones de ADNc que representaban genes *hcdk2*, así como otros genes.

[0438] Para validar la especificidad de compuesto de la interacción entre genes aislados en la exploración anterior, los genes se volvieron a clonar y se repitió el ensayo de halo. Se aisló un gen desconocido (denominado GPC-761) cuatro veces en la exploración descrita anteriormente. Uno de los plásmidos aislados que codifica este gen en el vector pACT2 se cotransformó con pBTM118c-DHFR en la cepa de levadura L40 y se realizó un ensayo de halo contra GPC 285985 o GPC 285993 (disuelto en DMSO) o 1 μ l de DMSO como control. La Figura 13 demuestra el crecimiento específico de compuesto del clon que contiene GPC-761. También se observaron resultados equivalentes para dichos ensayos de validación realizados usando los genes *hcdk2* identificados a partir de la exploración anterior.

[0439] La sustitución del nitrógeno en la posición 2 del grupo 4-oxoindeno [3,2-c]pirazol como en GPC 285993 había demostrado suprimir toda actividad hacia CDK2 en esta clase de sustancia (no se muestran los datos). La unión de GPC-761 a GPC 285985, pero no al equivalente sustituido en n GPC 285993, es similar en características a la unión de CDK2 a estos compuestos. Esto demuestra que los métodos proporcionados en la presente memoria son capaces de identificar un polipéptido que se une a un ligando especificado por el usuario a partir de una gran combinación de polipéptidos sin un conocimiento previo del polipéptido.

Ejemplo 11: Un ensayo de triple híbrido usando células de mamífero

[0440] Las células de mamífero pueden poseer distintas ventajas para realizar el ensayo de triple híbrido. Pueden presentar una mejor captación de compuesto y pueden permitir la detección de interacciones que no se observarían en células hospedadoras heterólogas debido a su capacidad para proporcionar la maquinaria/entorno para un plegamiento correcto y/o modificaciones postraduccionales que pueden ser necesarias para ciertas interacciones.

[0441] Para ensayar el rendimiento de los ligandos híbridos de dimerización y los métodos de la invención en células de mamífero, se ensayó la activación de un gen indicador CAT usando el sistema de mamífero Macthmaker (Clontech, N° Cat.: K1602-1). Con este fin, se clonó la DHFR en el vector pM (Clontech) y el GR2 en el vector pVP16 (Clontech) usando métodos análogos a los descritos en el Ejemplo 3; los vectores resultantes se denominan pM-DHFR y pVP16-GR2. Se transfectaron células HeLa convencionales con pM3-VP16 y pG5CAT (control positivo) o pM-DHFR, pVP16-GR2 y pG5CAT. 24 horas después de la transfección, el medio se cambió por medio al que se añadieron 100 μ l/100 ml de medio de una solución 100 μ M de GPC 285937 en DMSO (Figura 14A, B) o medio que contenía la misma cantidad de DMSO (Figura 14C). 24 horas después, se visualizó la actividad de CAT usando el equipo de tinción de CAT (Roche, N° Cat.: 1836358). Se observó claramente un precipitado coloreado en el control positivo (Figura 14A) y en las células que expresaban las fusiones de DHFR y GR2 incubadas con GPC 285937 (Figura 14B), pero no se observó precipitado coloreado en el control de DMSO (Figura 14C).

[0442] Esto muestra que los métodos de la invención pueden transferirse a un sistema celular distinto de levadura.

Ejemplo 12: Métodos de identificación de un ligando para un péptido especificado por el usuario: un sistema de ensayo de triple híbrido basado en un sistema de interacción basado en la transcripción

[0443] En ciertas aplicaciones, es ventajoso tener métodos a mano que puedan identificar una molécula pequeña, a partir de una combinación o biblioteca de moléculas pequeñas, que sea capaz de unirse a cierto primer polipéptido P1 de interés. Con este fin, puede prepararse una biblioteca de moléculas pequeñas R1 por métodos bien establecidos, por ejemplo, química combinatoria, u otros métodos conocidos por el experto en la materia y, posteriormente, acoplarse a un segundo ligando R2 que se sabe que se une a un segundo polipéptido P2 mediante un enlazador $(-\text{CH}_2-\text{X}-\text{CH}_2)_n$ para formar una biblioteca de compuestos de ligando híbridos $\text{R1}(-\text{CH}_2-\text{X}-\text{CH}_2)_n\text{-R2}$. Como alternativa, puede prepararse *de novo* una biblioteca de compuestos de ligando híbridos $\text{R1}(-\text{CH}_2-\text{X}-\text{CH}_2)_n\text{-R2}$ usando etapas tales como las proporcionadas en los esquemas 1-4 en la Figura 1. Sin embargo, esto no pretende limitar el alcance de la invención a dichos esquemas. Más bien, el experto en la materia escogerá, dependiendo de la aplicación deseada, de la gran variedad de reacciones químicas conocidas aquellas mejor adaptadas para generar la biblioteca que se ajuste a sus necesidades.

[0444] Por ejemplo, sin limitación, si se selecciona que R2 sea metotrexato, la biblioteca de compuestos de ligando híbridos puede usarse en la exploración siguiente: la secuencia codificante de P1 se amplifica a partir de una biblioteca adecuada o muestra que se sabe que contiene esta secuencia usando cebadores seleccionados para que sean específicos para P1, se digiere y se subclona en el vector pGAD426c, para dar pGAD426c-P1. Las células de la cepa de levadura L40 se cotransfectan con pBTM118c-DHFR y pGAD426c-P1. Los transformantes que reciben el plásmido se seleccionan en medio sintético que carece de triptófano y leucina, y las colonias individuales se vuelven a cultivar en medio líquido. Se preparan placas de microtitulación para que contengan miembros individuales o combinados de la biblioteca de compuestos de ligando híbridos a una concentración apropiada (que puede estar entre 10 mM y 0,1 nM) en medio SD que carece de leu, trp e his. Se inoculan en cada pocillo aproximadamente 1×10^4 , preferiblemente 1×10^5 , más preferiblemente 1×10^6 , o más preferiblemente 1×10^7 células cotransformadas con pGAD426c-P1 y pBTM118c-DHFR, como se han preparado anteriormente, y se incuban durante aproximadamente 1 a 3 días con las

soluciones que contienen los ligandos híbridos.

[0445] Se registra el crecimiento celular en los pocillos después de este periodo de cultivo. Los compuestos de ligando híbridos que se sabe que están presentes en aquellos pocillos en los que se detecta crecimiento pueden volver a ensayarse posteriormente en un ensayo de halo de validación como se ha descrito anteriormente en el Ejemplo 4. En el caso de combinaciones de ligandos híbridos, las combinaciones pueden fraccionarse por metodologías convencionales y los ligandos híbridos individuales ensayarse en ensayos de halo y, posteriormente, identificarse por metodologías convencionales. Cuando puede determinarse un crecimiento específico de ligando híbrido, el compuesto unido a metotrexato para formar este ligando híbrido se selecciona como capaz de unirse a P1.

Ejemplo 13: Métodos de identificación de un polipéptido que se une a un ligando especificado por el usuario: un sistema de ensayo de triple híbrido basado en la técnica detectora de proteína ubiquitina dividida

[0446] La técnica detectora de proteína ubiquitina dividida se ha usado para detectar interacciones de proteína *in vivo* o *in vitro*. Generalmente, es útil para todas las clases de interacciones de proteína-proteína, pero es particularmente útil en los casos en los que es problemático el ensayo de doble híbrido de levadura convencional, es decir, cuando están implicadas proteínas de membrana, activadores o represores de la transcripción, etc. Pueden tomarse detalles adicionales de esta técnica, por ejemplo, de los documentos US 5.585.245, US 5.503.977 o Johnsson & Varshavsky (1997) en: *The Yeast Two-Hybrid System (Advances in Molecular Biology)*, Ed. Paul L. Bartel y Stanley Fields, Oxford University Press, págs. 316-332. En la presente memoria, se muestra cómo el principio de detector de ubiquitina dividida puede emplearse igualmente en un experimento de triple híbrido para investigar interacciones entre proteínas y moléculas pequeñas.

Construcción de vectores para un sistema de ensayo de triple híbrido basado en detector de proteína ubiquitina dividida

[0447] Se seleccionó la cepa de levadura JD53 (Dohmen et al., JBC, 1995, 270: 18099-109) para los experimentos que implicaban GFP como indicador y detección en transferencias de Western, y se usó la cepa de levadura L40 en los experimentos en los que se usaba la transcripción inducida por PLV de *HIS3* como lectura.

[0448] El plásmido pS_{DHFR}-Cub-PLV que codifica una proteína de fusión (Figura 9) que comprende Sec62, que facilita el anclaje a membrana, DHFR (dihidrofolato reductasa), Cub, la parte C-terminal de ubiquitina y PLV (factor de transcripción quimérico: proteínaA::lexA::VP16) se construye de la forma siguiente. En primer lugar, se amplifica por PCR un fragmento de *folA'* (DHFR) de *E. coli* a partir de una genoteca de ADN genómico de *E. coli* (Clontech, N° Cat. XL4001AB), usando los cebadores:

5'-GGGGGTCGACATGATCAGTCTGATTGCGGCGTTAGCG-3', y,

5'-GGGGGCGGCCGCTTACCGC82CGCTCCAGAATCTCAAAG-3'.

[0449] En segundo lugar, el producto de PCR se digiere entonces con Sall y NotI y se subclona en el vector Cub-PLV (Stagljar et al. (1998) Proc. Natl. Acad. Sci. U.S.A., 95: 5187-92) de modo que Cub está cadena abajo del DHFR insertado y cabeza arriba del PLV indicador, al tiempo que las tres proteínas están en fase de lectura, produciendo el plásmido pDHFR-Cub-PLV. En tercer lugar, el gen que codifica el anclaje a membrana Sec62 se inserta cadena arriba de DHFR después de la amplificación por PCR del gen usando cebadores con sitios de restricción Sall flanqueantes. Los cebadores de pCR apropiados para la amplificación de Sec62 a partir de ADN genómico de levadura (*S. cerevisiae*) son:

5'-GATCGTCGACATGGTAGCCGAGCAAACACAGGAG-3' y

5'-GATCGTCGAC GTTTTGTTCGGCTTTTTCATTGATG-3'.

[0450] Tras la escisión de la proteína de fusión después del resto Cub, el PLV se liberará de la fusión y su localización anclada a membrana, y se transfiere al núcleo, donde activa la transcripción de genes bajo el control de un promotor que comprende sitios de unión a LexA.

[0451] Para construir el plásmido pDHFR-Cub-GFP, el resto PLV en pDHFR-Cub-PLV se sustituye con un casete de GFP a partir de pCK GFP-S65C usando sitios de restricción compatibles que flanquean ambos casetes (Reichel, et al., PNAS, 1996, 93: 5888-93). Se construye un plásmido indicador alternativo, pDHFR-Cub-R-GFP, de modo que se clona una secuencia líder de 20 aminoácidos que contiene lisina entre Cub y GFP, de modo que el primer aminoácido del fragmento de líder-GFP producido después de la escisión del enlace peptídico de Cub-R es un resto de arginina.

[0452] El plásmido pNubl-hCDK2 se construye digiriendo el fragmento de PCR de *hcdk2* producido en el Ejemplo 3 con enzimas de restricción apropiadas y subclonando el producto en el plásmido pNubl (Laser et al., PNAS, 2000, 97: 13732-7).

[0453] Para construir una biblioteca de plásmidos que codifican la mitad N-terminal de ubiquitina fusionada a una biblioteca de polipéptidos, se genera una genoteca de ADNc a partir de ARN poli A+ aislado de cerebro fetal humano (hFB) (Clontech, N° CAT. 6525-1), usando esencialmente un protocolo y reactivos suministrados por Invitrogen (Life Technologies, Superscript, N° CAT. 18248-013) pero empleando cebadores oligo-dT para la síntesis de primera cadena de la forma siguiente:

TT1-A: 5' TTT TGT ACA TCT AGA TCG CGA GCG GCC GCC CTT TTT TTT TTT TV-3'

siendo V A, G, o C a proporción molar equivalente. Los fragmentos de ADNc resultantes se subclonaron en el plásmido pNubl como fragmentos de restricción con Sall/NotI (pADNX-N_{ub}IBC; Laser et al., PNAS, 2000, 97: 13732-7) para producir una biblioteca de plásmidos denominados en la presente memoria pNubl-hFB.

5 Cuantificación del grado de escisión de DHFR-Cub-GFP

10 [0454] El plásmido de "cebo-Cub-indicador" pDHFR-Cub-GFP (1 µg) se cotransforma con pNubl-hCDK2 en la cepa de levadura JD53 (Dohmen et al., JBC, 1995, 270: 18099-109) por técnicas convencionales (Burke et al., Methods in yeast genetics: A Cold Spring Harbor Laboratory course manual; Cold Spring Harbor Laboratory Press, 2000). Los cotransformantes que contienen ambos plásmidos se seleccionan en medio que carece de leu y trp. Se vuelven a cultivar colonias individuales en medios líquidos y se inoculan 1×10^4 , preferiblemente 1×10^5 , más preferiblemente 1×10^6 , o más preferiblemente 1×10^7 células en pocillos individuales de placas de microtitulación que contienen medio SD que carece de trp y leu pero que contienen el ligando híbrido de dimerización GPC 285985 a aproximadamente 50 µM en DMSO o con DMSO como control. Después de 1 a 3 días de incubación a 30°C, la escisión del resto indicador GFP de Cub se detecta mediante análisis de transferencia de Western usando anticuerpos específicos de GFP (Clontech, N° Cat.: 8369-1) y se observa solamente para células de los pocillos que contienen GPC 285985. La detección del resto GFP escindido (de aproximadamente 29kDa) es indicativa de la interacción del ligando híbrido y las proteínas de fusión.

20 [0455] La repetición del experimento anterior, pero usando el pDHFR-Cub-R-GFP en lugar de pDHFR-Cub-GFP, demuestra pérdida de actividad de GFP a través de la degradación de la regla del N-terminal después de su escisión de Cub provocada por la formación de un complejo trimérico de las proteínas de fusión de DHFR-Cub-R-GFP y Nubl-hCDK2 unidas por el ligando híbrido. La intensidad fluorescente de GFP en aquellas células de levadura expuestas al ligando híbrido GPC 285985 se reduce en comparación con aquellas células expuestas sólo a DMSO. La intensidad fluorescente se mide usando un lector de placas de microtitulación convencional (Victor V, Perkin Elmer) o un dispositivo de exploración/separación de células por fluorescencia (FACS), por ejemplo, de Cytomation o Beckton Coulter.

Cuantificación del grado de escisión de Sec62-DHFR-Cub-PLV por exploración para un marcador auxotrófico

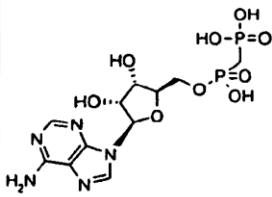
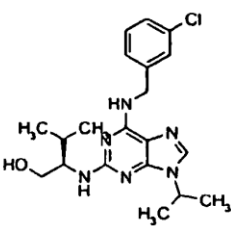
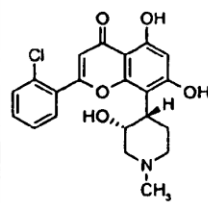
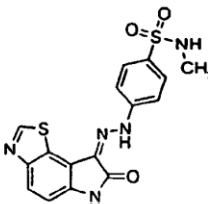
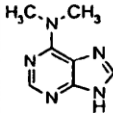
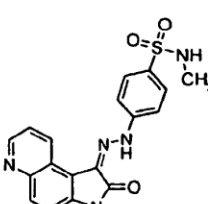
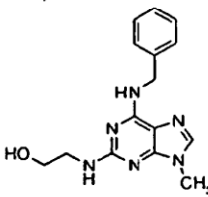
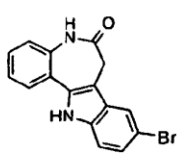
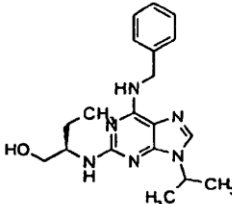
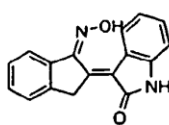
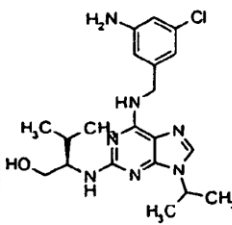
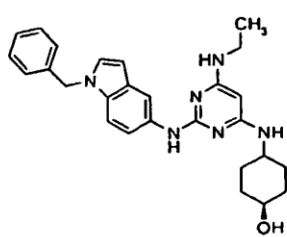
25 [0456] El resto PLV, cuando se sintetiza como una fusión de Sec62-DHFR-Cub-PLV del plásmido pS_{DHFR}-Cub-PLV, se une a la membrana del RE fuera del núcleo y, por lo tanto, no está disponible para la activación de la transcripción de genes indicadores. Sólo tras la escisión de la proteína de fusión después del resto Cub se liberará PLV, que sirve como factor de transcripción para activar genes indicadores bajo el control del promotor que alberga sitios de unión a *lexA* en el interior del núcleo (Stagljar et al. (1998) Proc. Natl. Acad. Sci. U.S.A., 95: 5187-92).

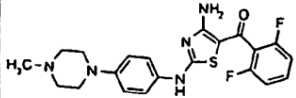
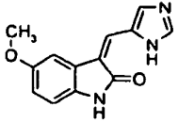
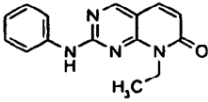
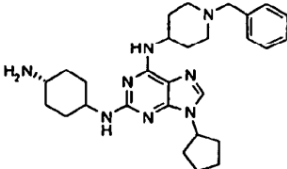
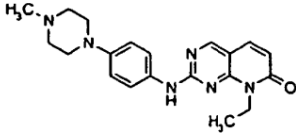
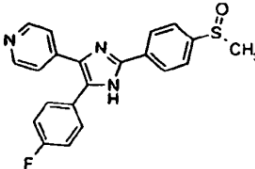
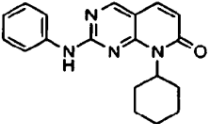
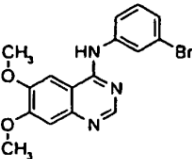
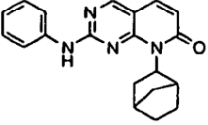
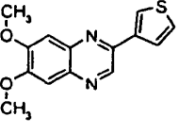
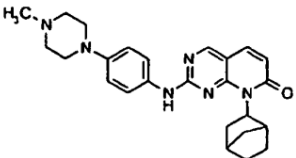
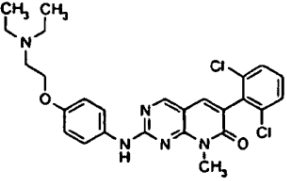
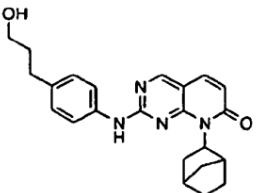
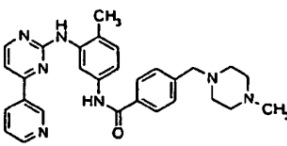
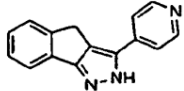
30 [0457] El plásmido de "cebo-Cub-indicador" pS_{DHFR}-Cub-PLV (1 µg) se cotransforma con la biblioteca de plásmidos pNub-hFB (5 µg) en la cepa de levadura L40 por técnicas convencionales. Después, los transformantes se siembran en placas de SD de 22 x 22 preparadas con medio que carece de leu y trp. Después de 3 días de incubación a 30°C, los cotransformantes se retiran por lavado de las placas, se mezclan y se congelan como alícuotas pequeñas. Se siembran 2×10^6 células en placas de SD que carecen de trp, leu e his, pero que contienen GPC 285985 50 µM, y se incuban durante 2-5 días. Sólo las células que contienen ambos plásmidos y que presentan un gen de HIS3 activo (imidazol-glicerol-fosfato-deshidratasa) pueden sobrevivir (positivas en la primera exploración). La activación del gen de HIS3 depende de la interacción entre pNub-hFB, GPC 285985 y pS_{DHFR}-Cub-PLV, que desencadena la escisión mediada por UBP del indicador PLV de la proteína de fusión de cebo. El indicador PLV liberado se trasladará después al núcleo, donde se inicia la transcripción del gen indicador (HIS3), conduciendo al crecimiento en medio SD que carece de histidina.

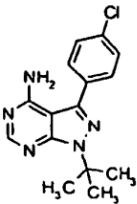
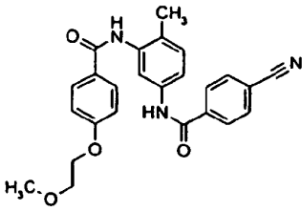
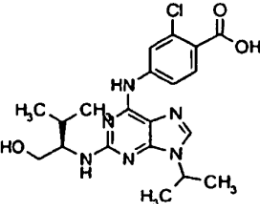
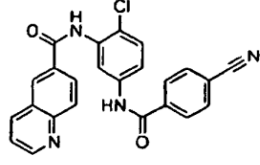
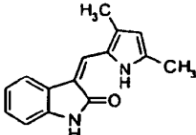
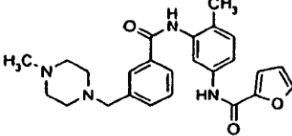
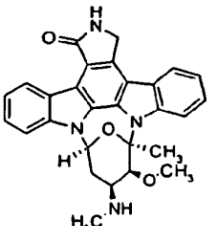
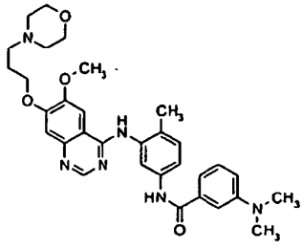
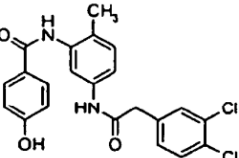
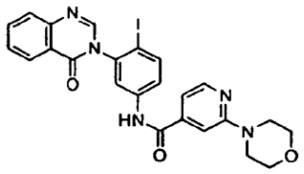
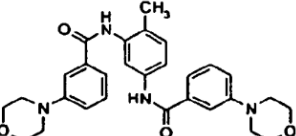
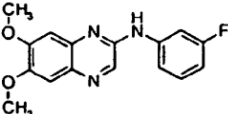
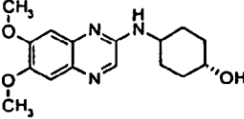
[0458] Los clones positivos en la primera exploración se escogen y se ensayan en un ensayo de halo de alto rendimiento análogo al descrito en el Ejemplo 10. Los clones positivos de esta exploración se identifican por secuenciación de ADN, e incluyen clones que contienen genes que expresan CDK2 y otros genes.

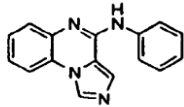
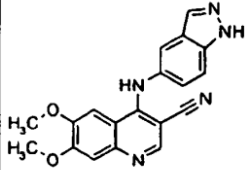
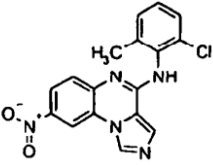
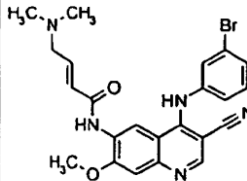
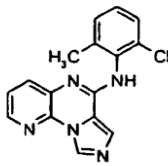
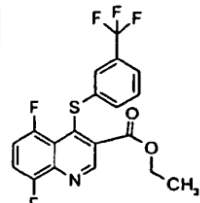
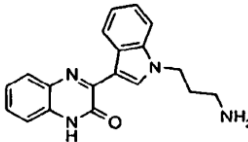
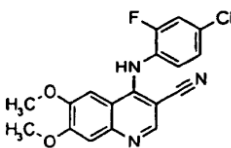
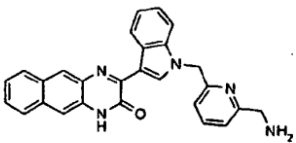
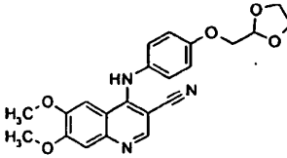
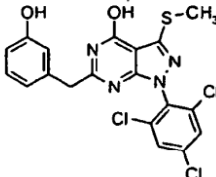
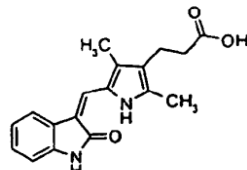
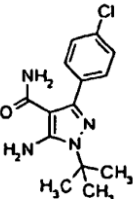
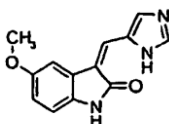
Equivalentes

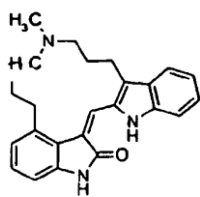
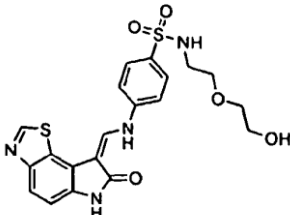
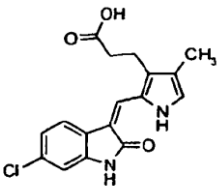
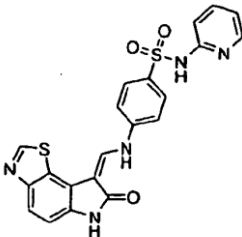
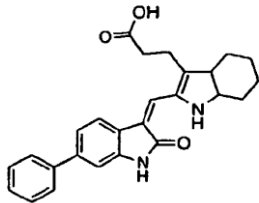
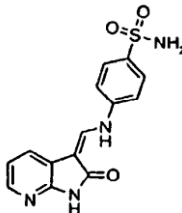
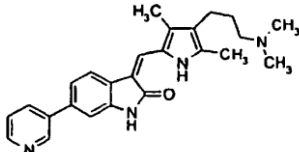
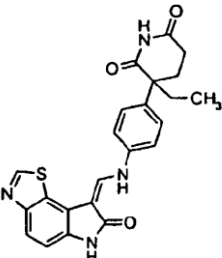
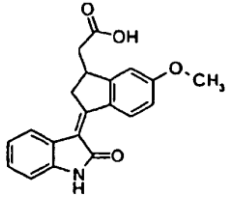
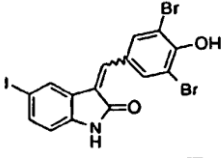
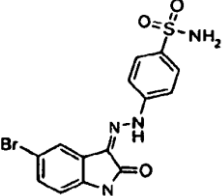
45 [0459] Los expertos en la materia reconocerán, o serán capaces de determinar usando no más que una experimentación de rutina, numerosos equivalentes a los procedimientos específicos descritos en la presente memoria. Dichos equivalentes se consideran dentro del alcance de esta invención y están cubiertos por las reivindicaciones siguientes.

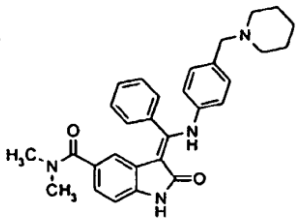
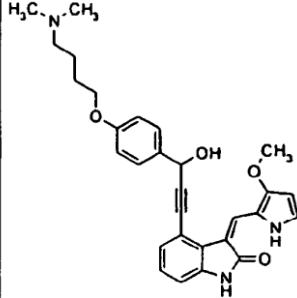
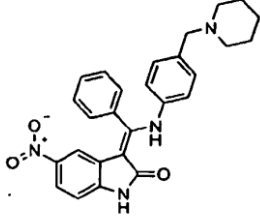
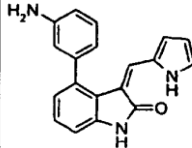
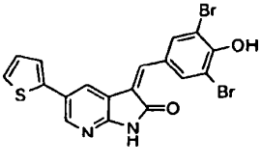
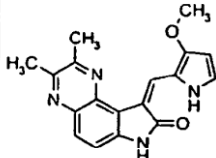
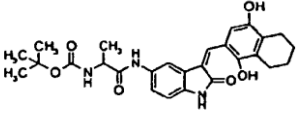
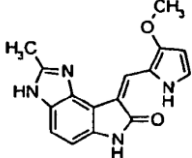
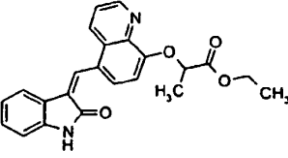
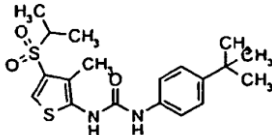
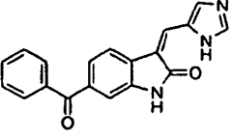
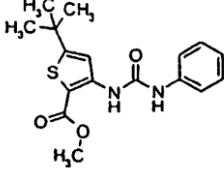
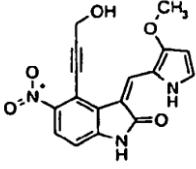
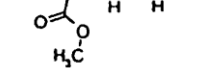
Estructura	Referencia	Estructura	Referencia
	Science 1997, 278(5336): 286-90		
			
			
			Cancer Res. 1999, 59, 2566
			
			

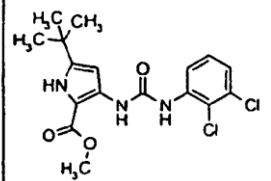
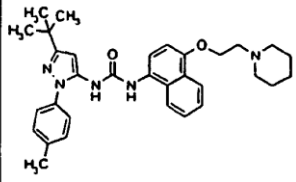
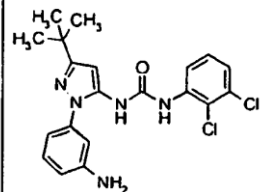
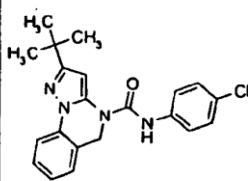
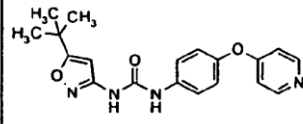
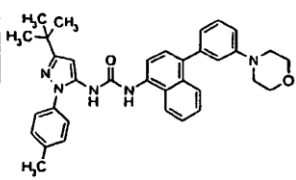
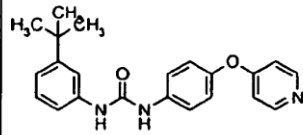
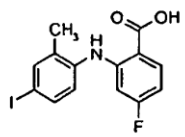
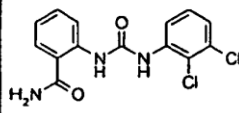
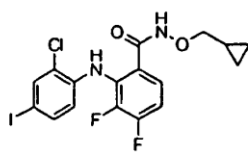
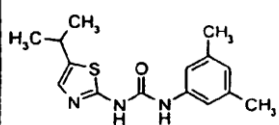
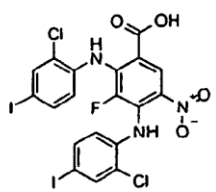
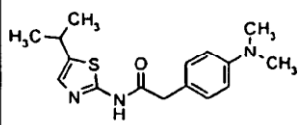
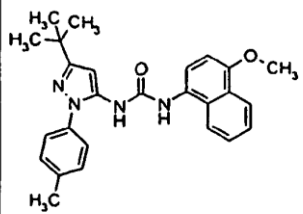
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			J. Med. Chem. 1998, 41, 3276
			
	Bioorg. Med. Chem. Lett. 2002 , 12, 221-224		

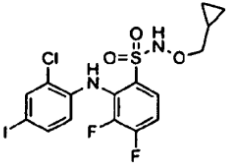
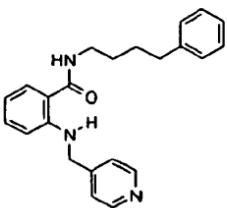
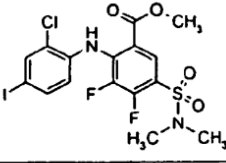
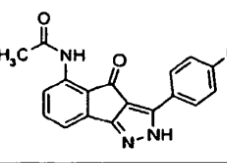
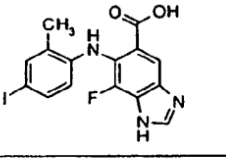
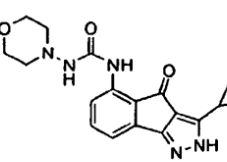
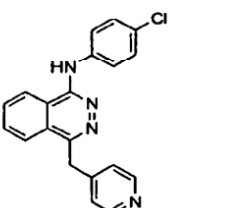
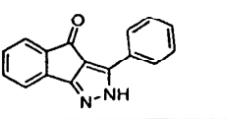
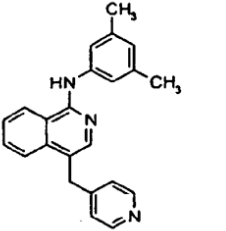
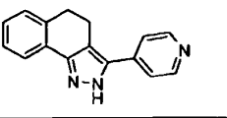
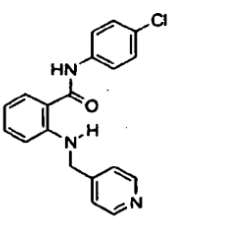
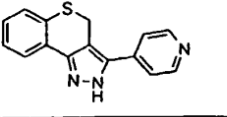
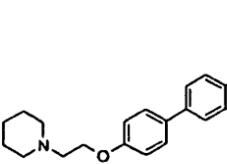
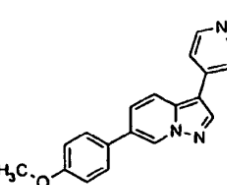
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			
			

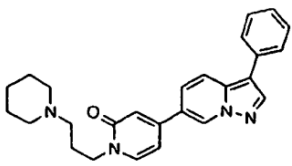
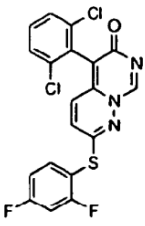
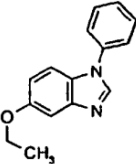
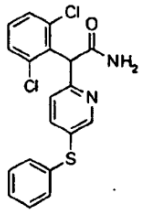
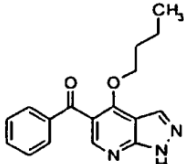
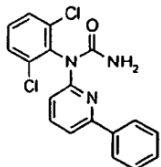
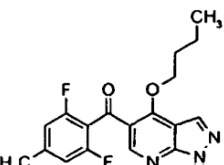
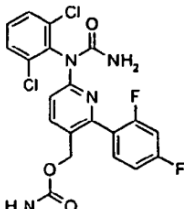
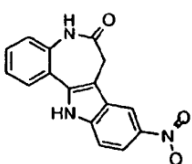
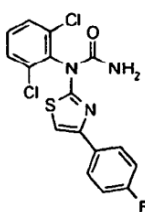
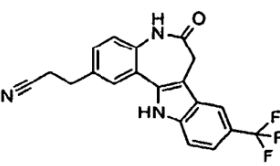
Estructura	Referencia	Estructura	Referencia
			Bior.Med.Chem.Lett. 2001, 11. 1401-1405
			
			
			
			
			
			

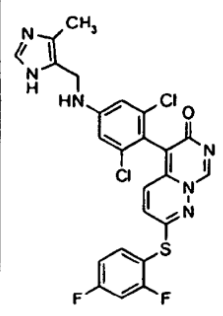
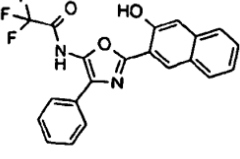
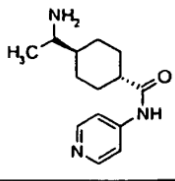
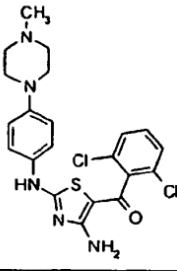
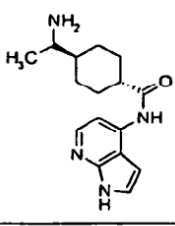
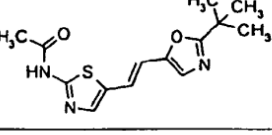
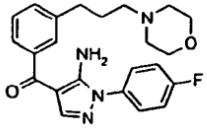
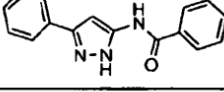
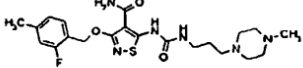
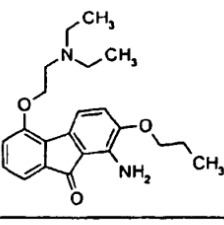
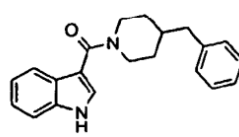
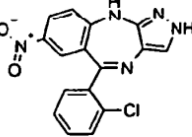
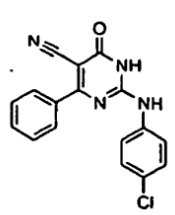
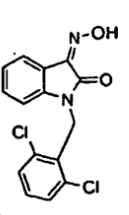
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			Bioorg. Med. Chem. Lett. 2000, 10, 223
			

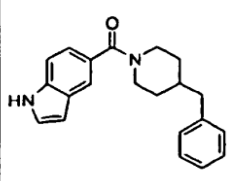
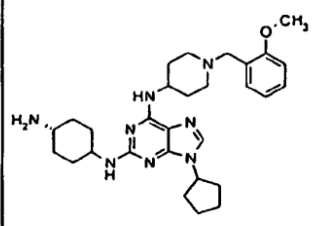
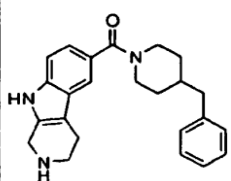
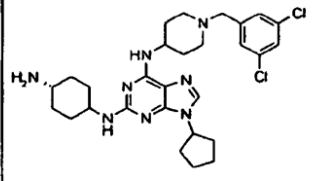
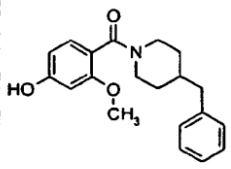
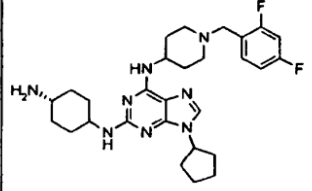
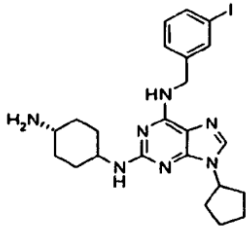
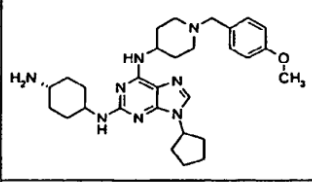
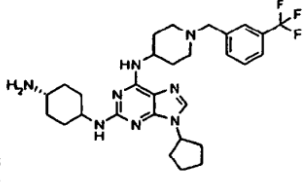
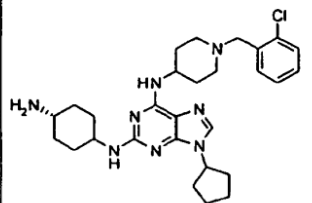
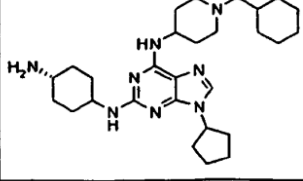
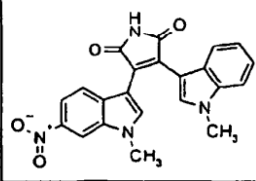
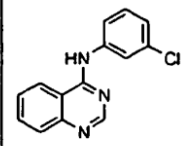
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			
			Bioorg. Med. Chem. Lett. 2001, 11, 9-12

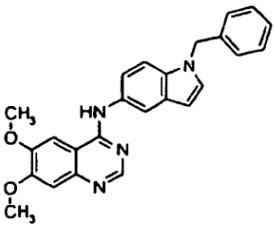
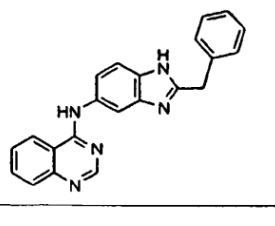
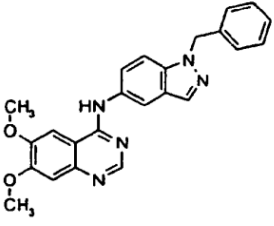
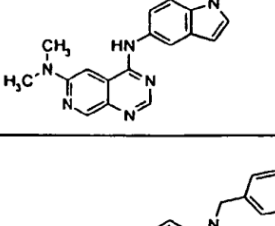
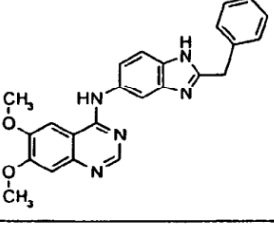
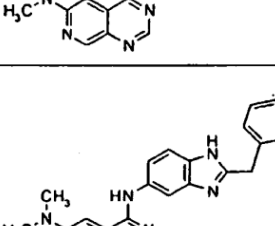
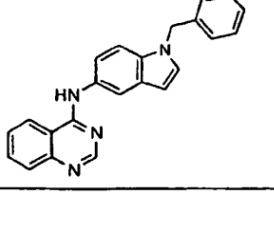
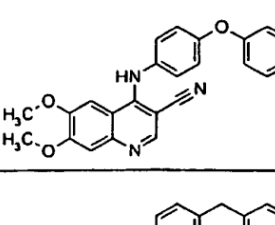
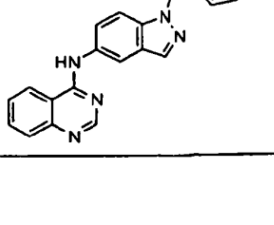
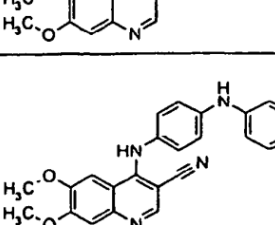
Estructura	Referencia	Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		
	Bioorg. Med. Chem. Lett. 2000, 10, 2051-2054		
			
			
			
			
			
			

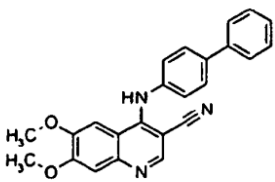
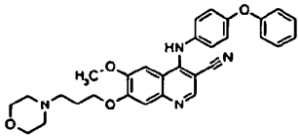
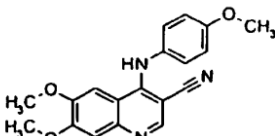
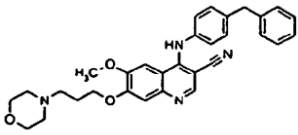
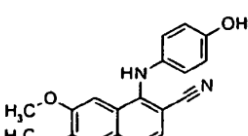
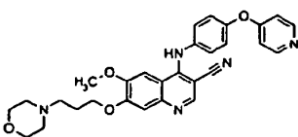
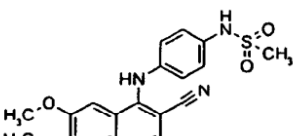
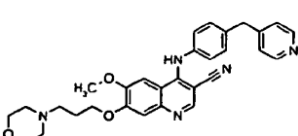
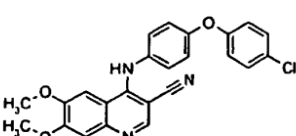
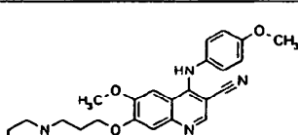
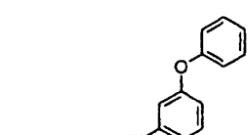
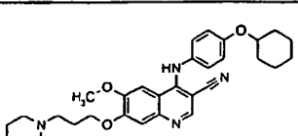
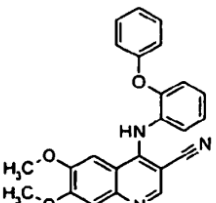
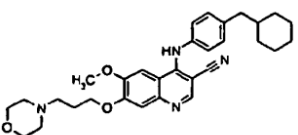
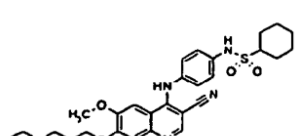
Estructura	Referencia	Estructura	Referencia
			
			
			
			WO 9917770
			
	WO 01/081311		
			
			

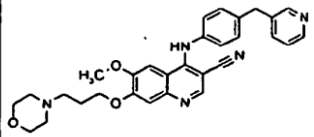
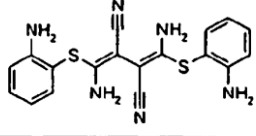
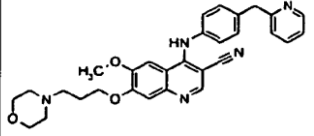
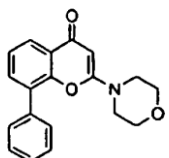
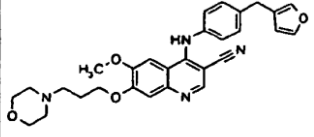
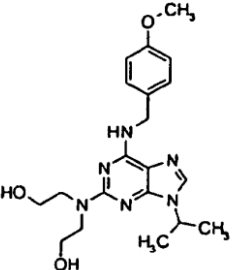
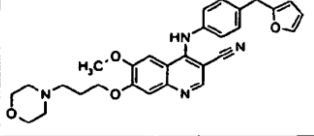
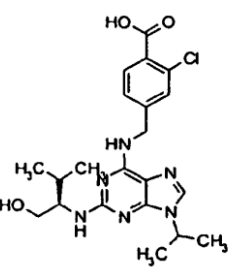
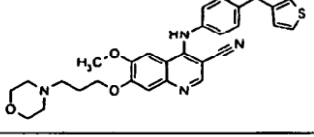
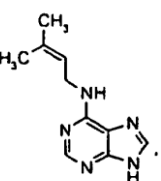
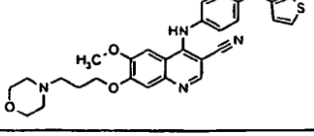
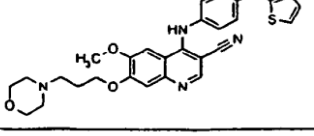
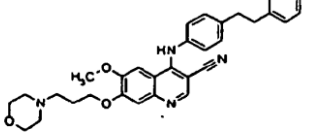
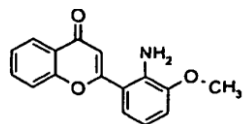
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
	Bioorg. Med. Chem. Lett., 2000, 10, 567-569		

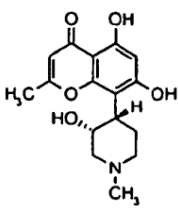
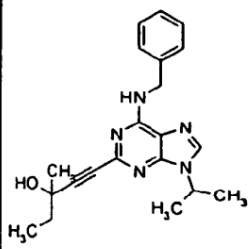
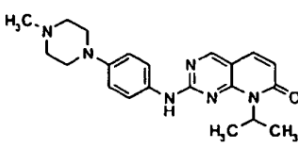
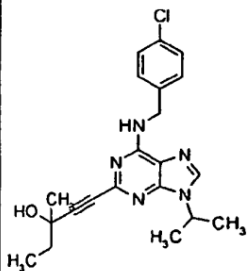
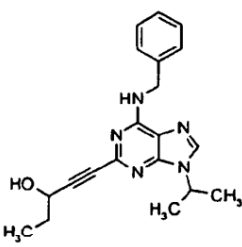
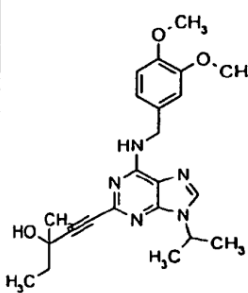
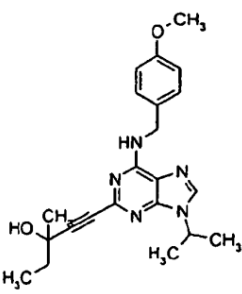
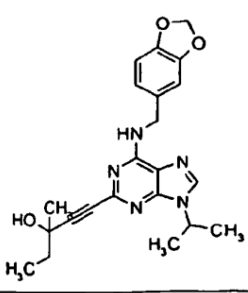
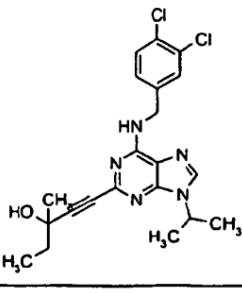
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			
			

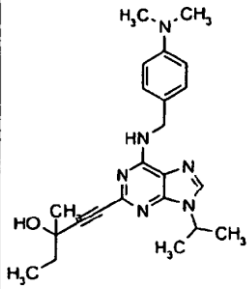
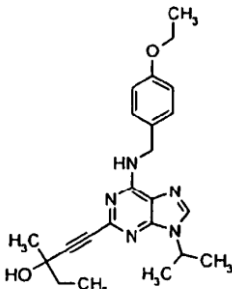
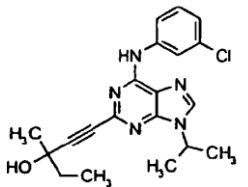
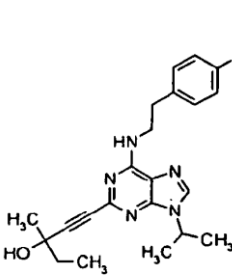
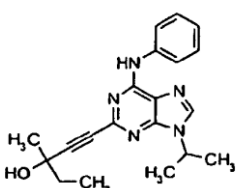
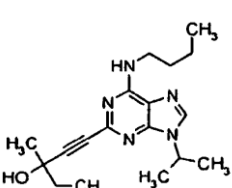
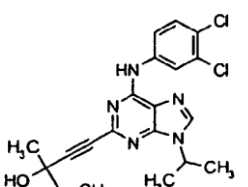
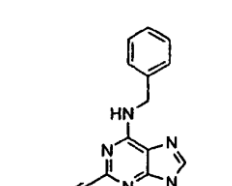
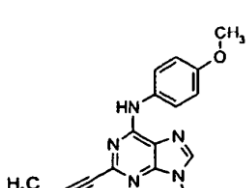
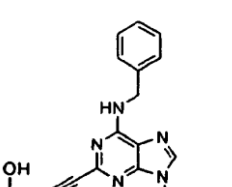
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			
			

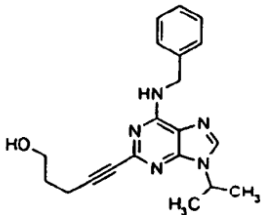
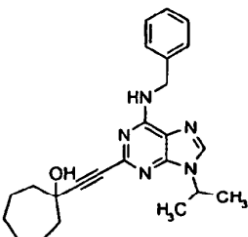
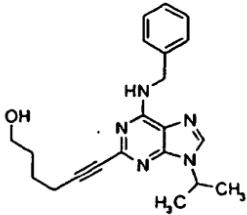
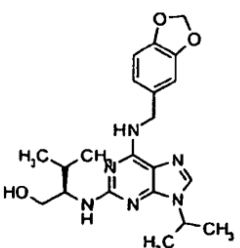
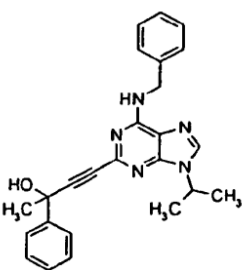
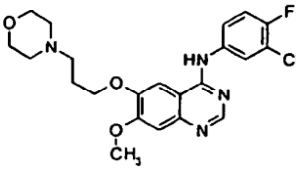
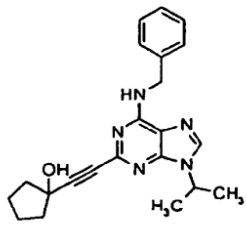
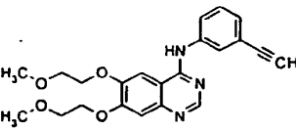
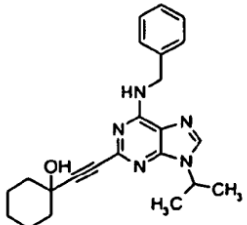
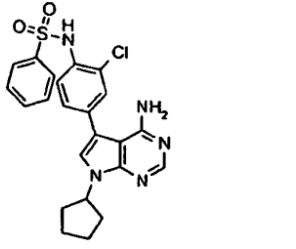
Estructura	Referencia	Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405		
	Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405		Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405
	Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405		Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405
	Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405		Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405
	Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405		
			

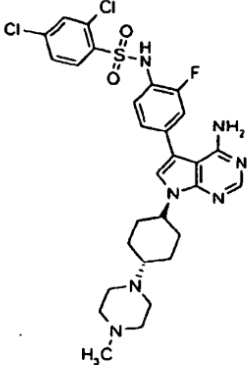
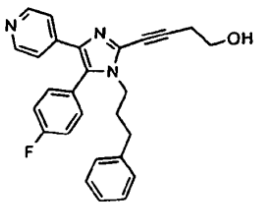
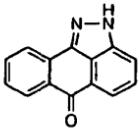
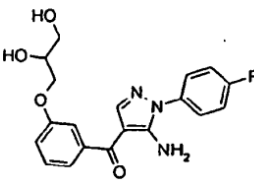
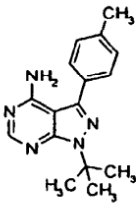
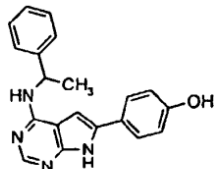
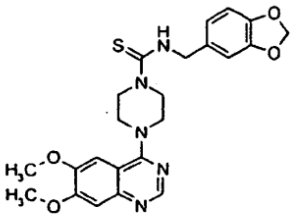
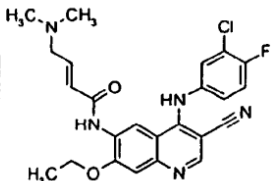
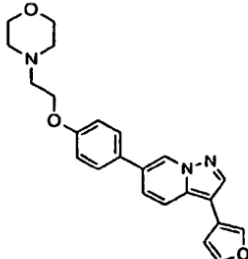
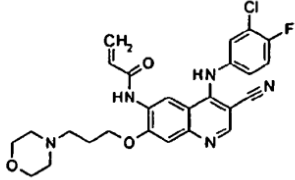
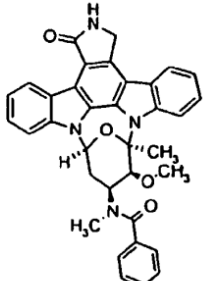
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			
			
			

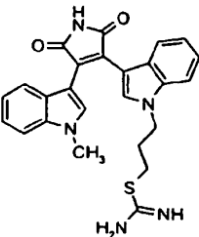
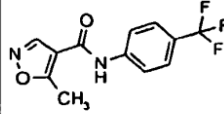
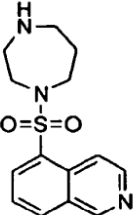
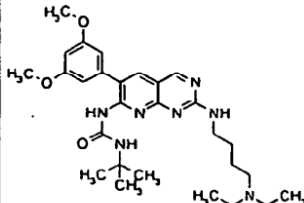
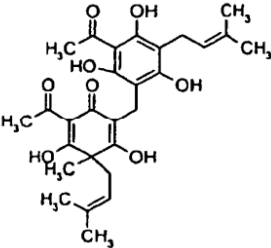
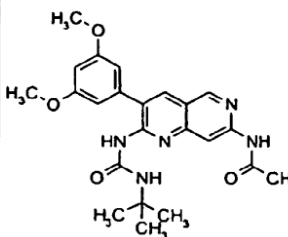
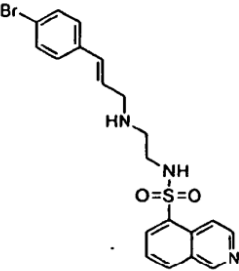
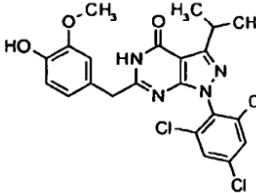
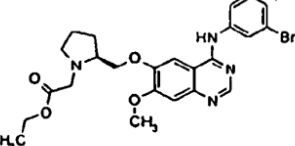
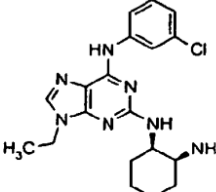
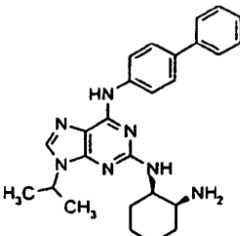
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			
			
			
			

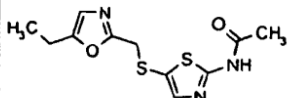
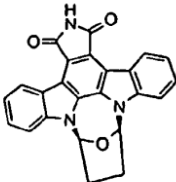
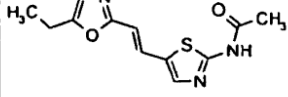
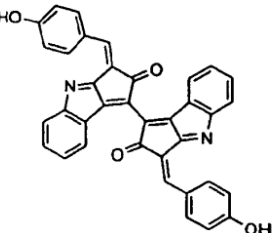
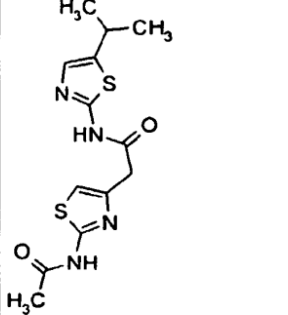
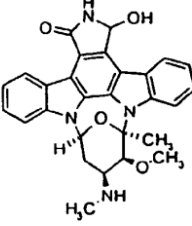
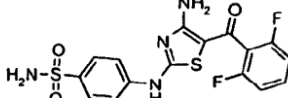
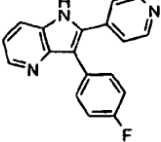
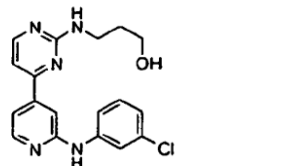
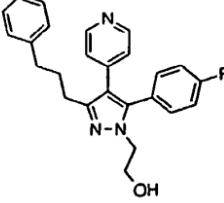
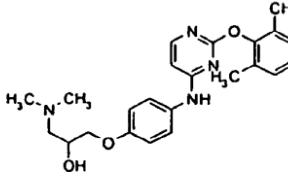
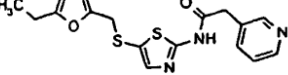
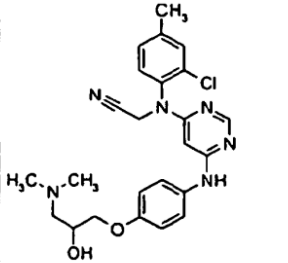
Estructura	Referencia	Estructura	Referencia
			
			
	J. Med. Chem. 2000, 43, 4089-4108		
			
			

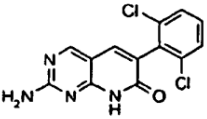
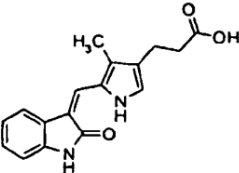
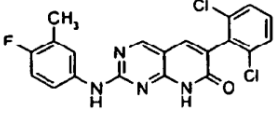
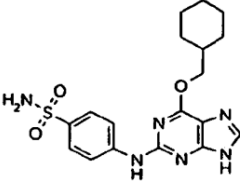
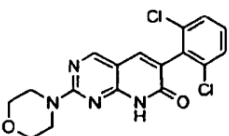
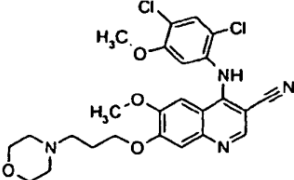
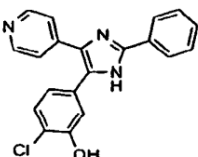
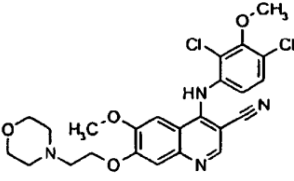
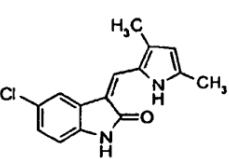
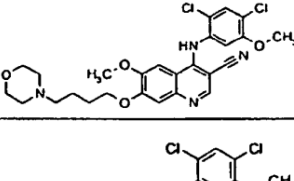
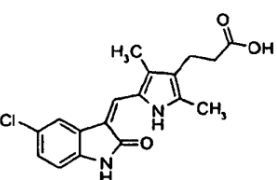
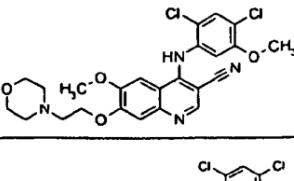
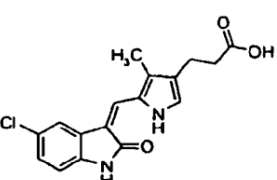
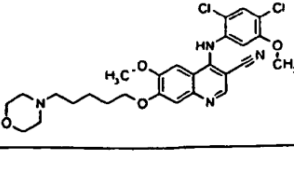
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			

Estructura	Referencia	Estructura	Referencia
			
			
			Bioorg. Med. Chem. Lett. 2001, 11, 1922-1914
			
			

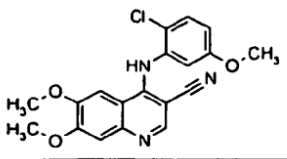
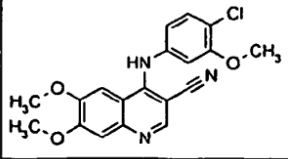
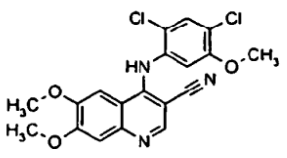
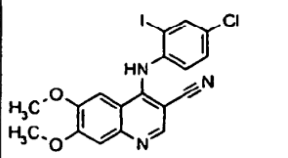
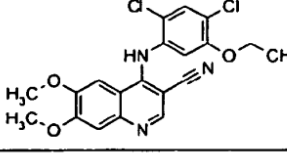
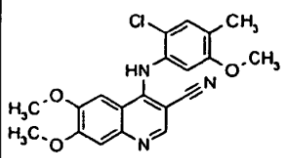
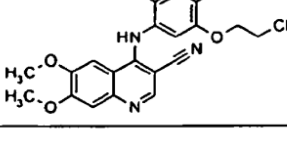
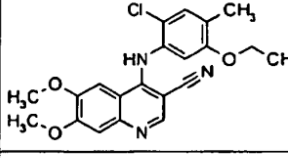
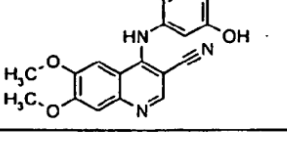
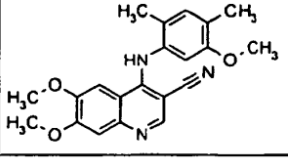
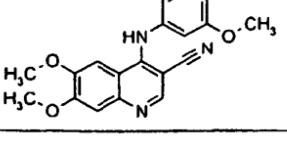
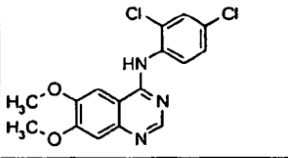
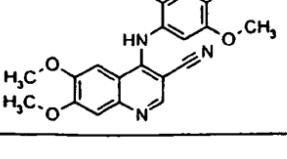
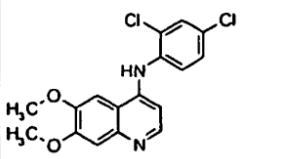
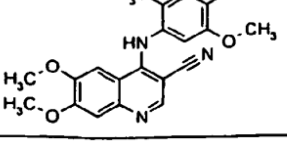
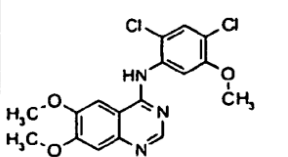
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			

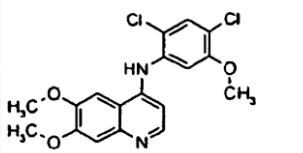
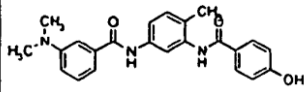
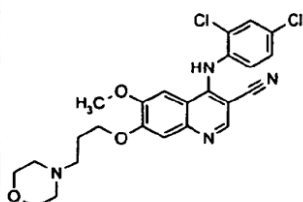
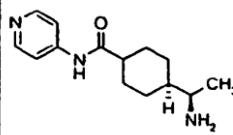
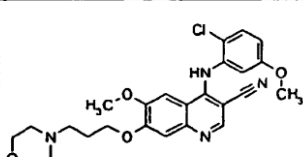
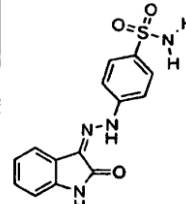
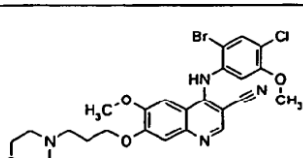
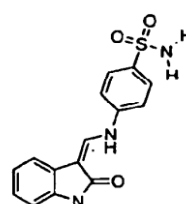
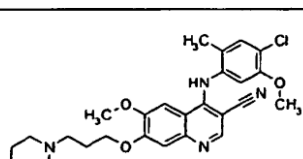
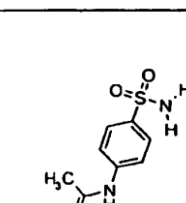
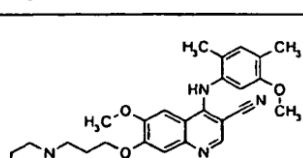
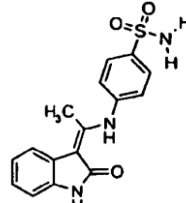
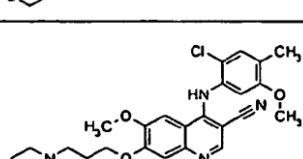
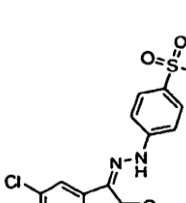
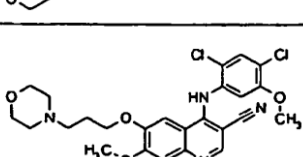
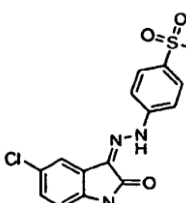
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			

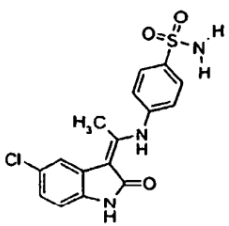
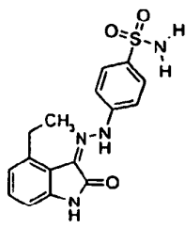
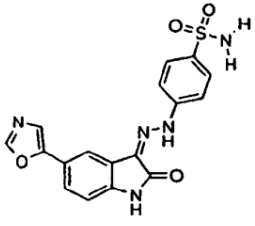
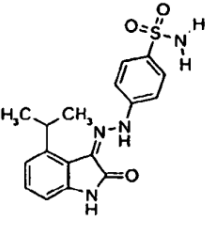
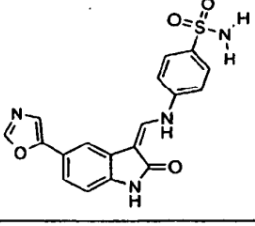
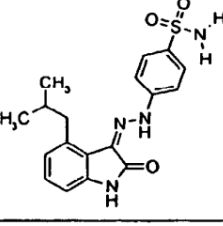
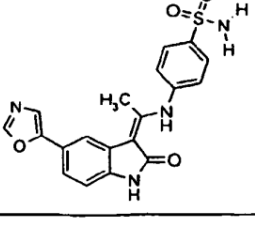
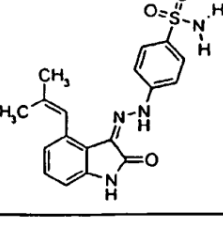
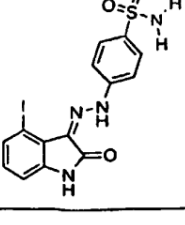
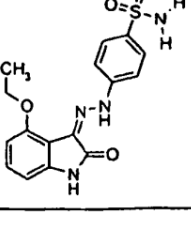
Estructura	Referencia	Estructura	Referencia
			
			
			
			WO 9920624
			
			
			

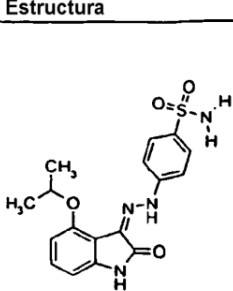
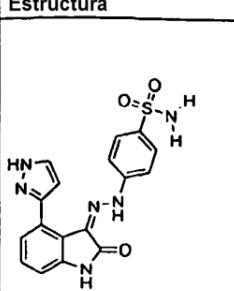
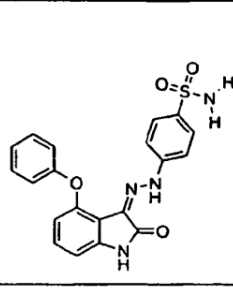
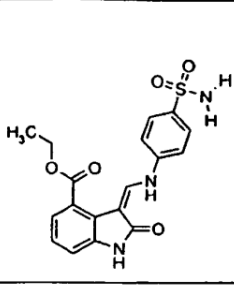
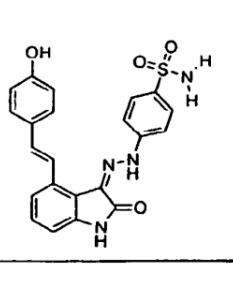
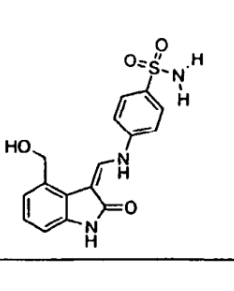
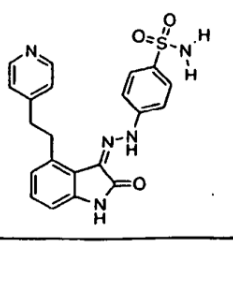
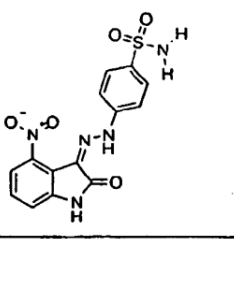
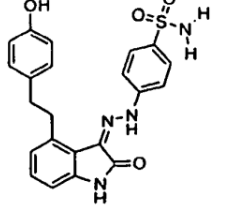
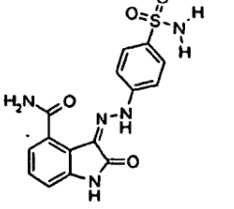
Estructura	Referencia	Estructura	Referencia
			
			
			J. Med. Chem. 2001, 44, 3965-3977
			
			J. Med. Chem. 2001, 44, 3965-3977
			J. Med. Chem. 2001, 44, 3965-3977
			J. Med. Chem. 2001, 44, 3965-3977

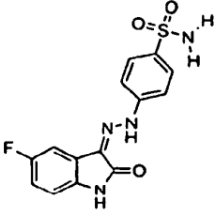
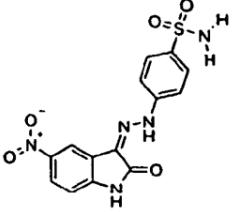
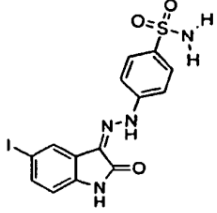
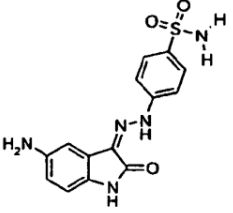
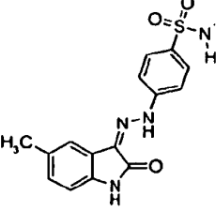
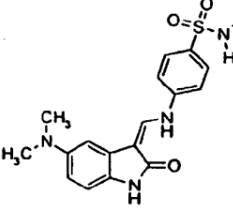
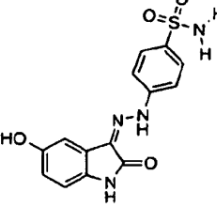
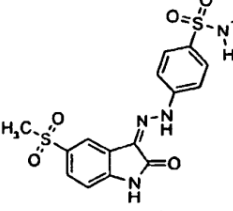
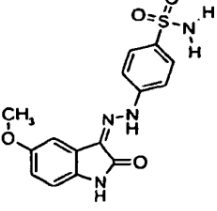
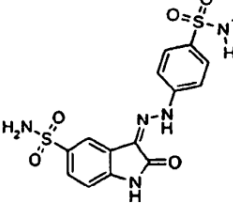
Estructura	Referencia	Estructura	Referencia
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977

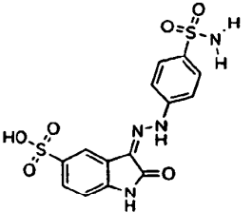
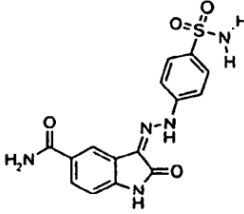
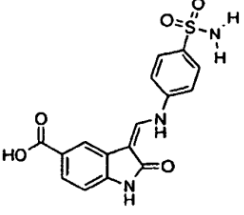
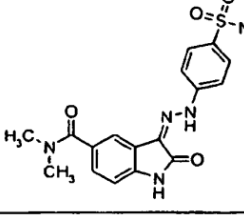
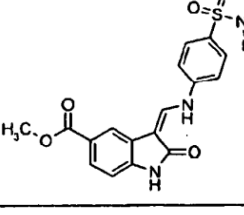
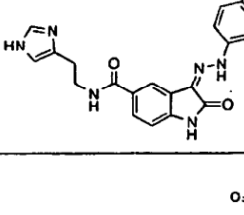
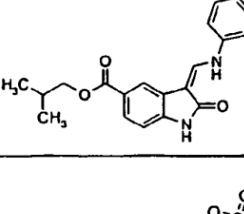
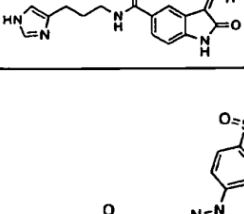
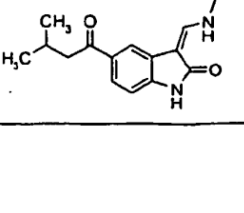
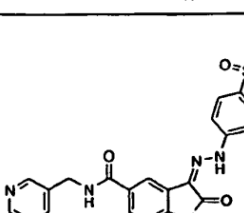
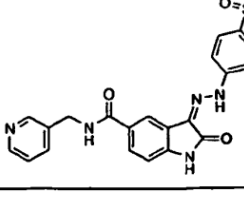
Estructura	Referencia	Estructura	Referencia
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		J. Med. Chem. 2001, 44, 3965-3977
	J. Med. Chem. 2001, 44, 3965-3977		
	J. Med. Chem. 2001, 44, 3965-3977		
			

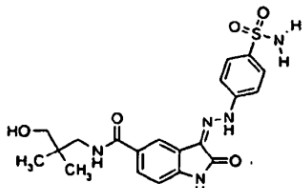
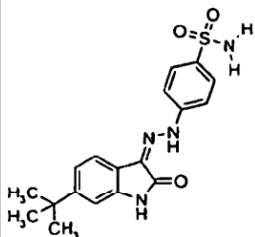
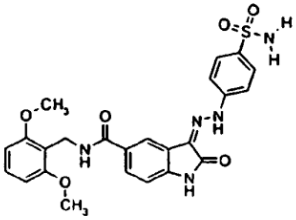
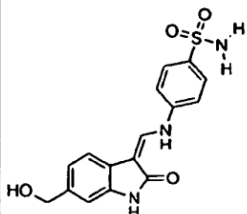
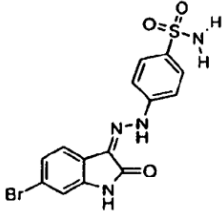
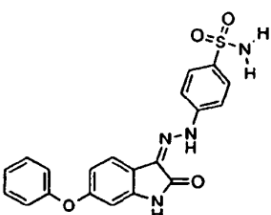
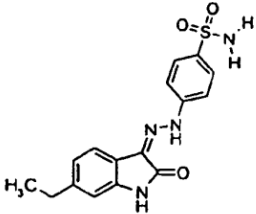
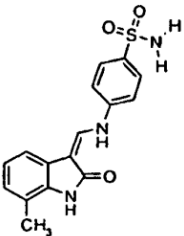
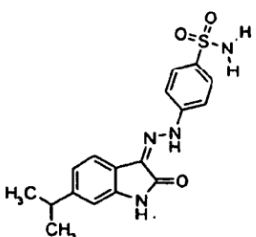
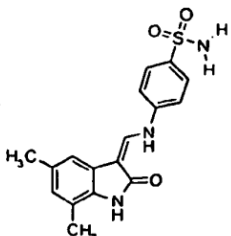
Estructura	Referencia	Estructura	Referencia
			Chem. Biol. 1999, 6, 559 J. Biol. Chem. 1999, 276, 17420, WO 9822103
			Nature, 1998, 389, 990
			
			
			
			
			
			

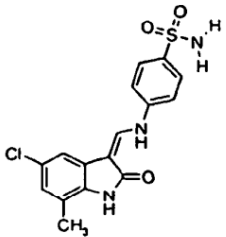
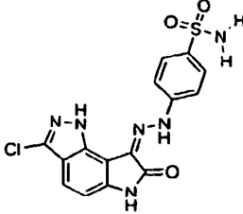
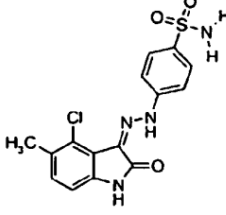
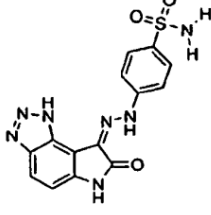
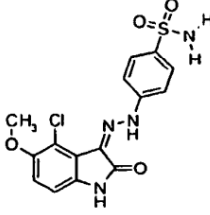
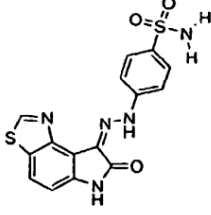
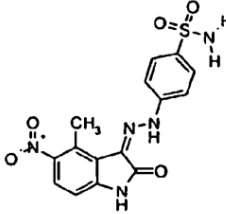
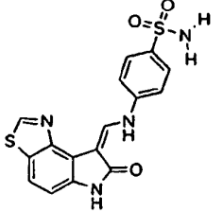
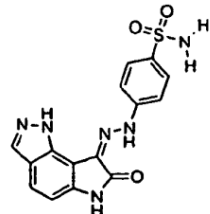
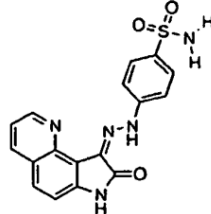
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			

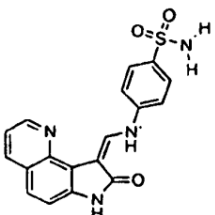
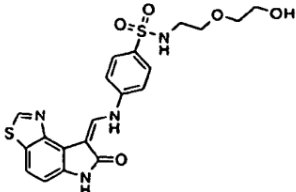
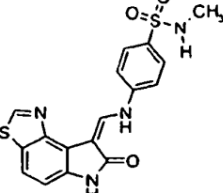
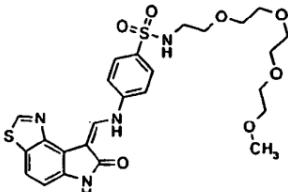
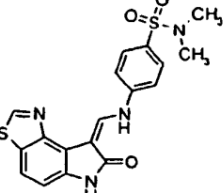
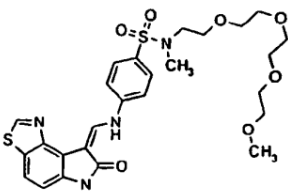
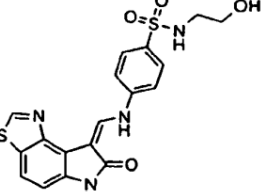
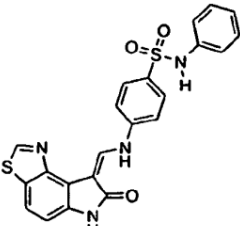
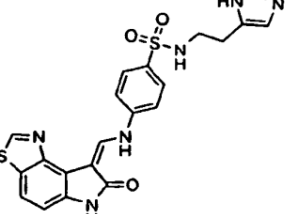
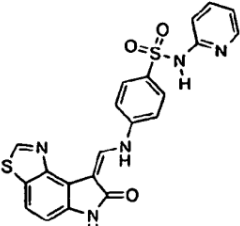
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			

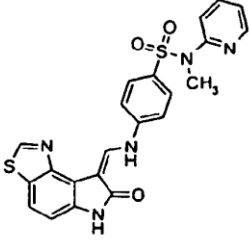
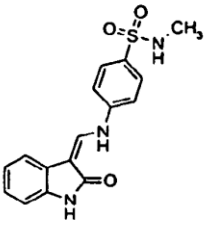
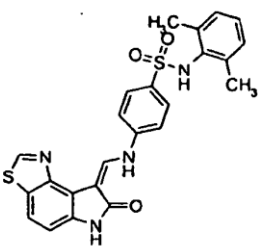
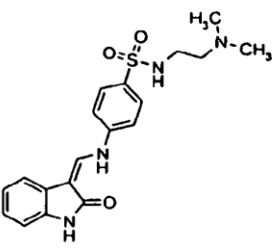
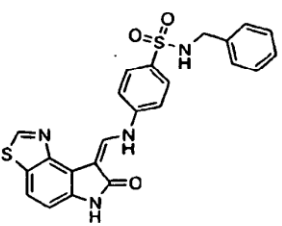
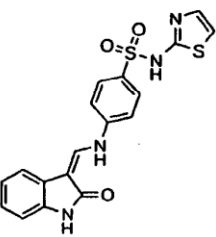
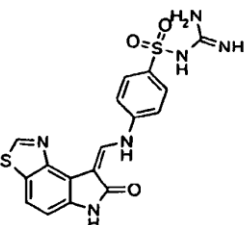
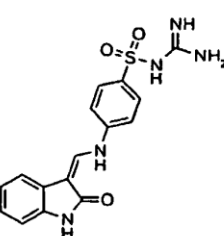
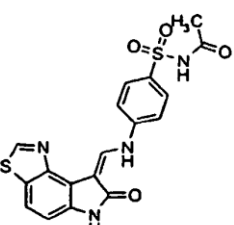
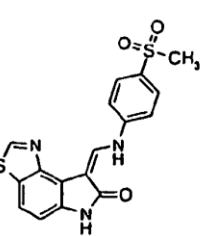
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			

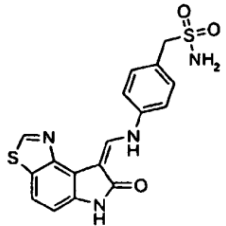
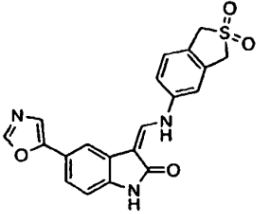
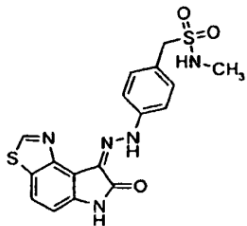
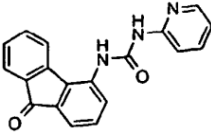
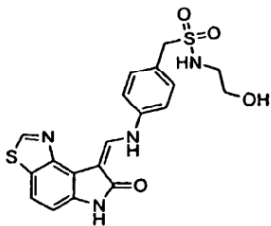
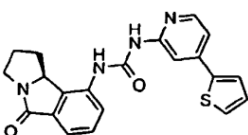
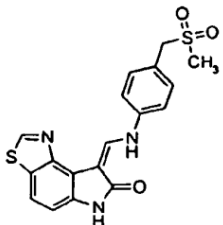
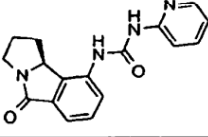
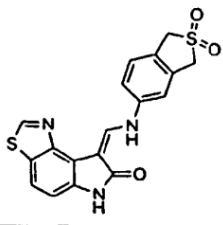
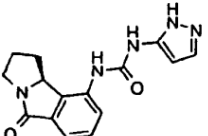
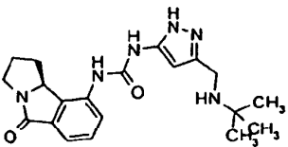
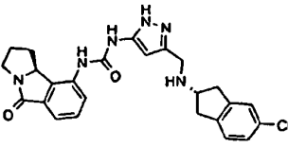
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			
			

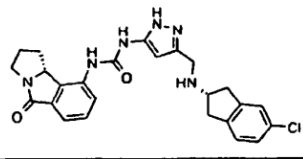
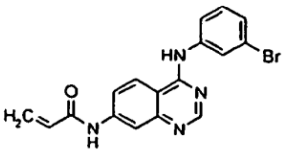
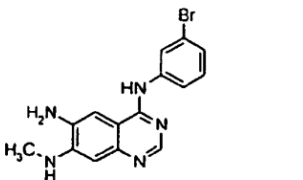
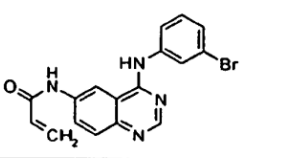
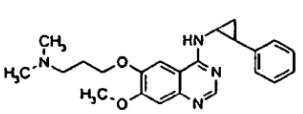
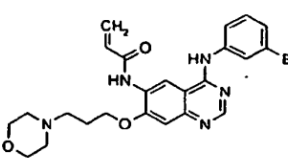
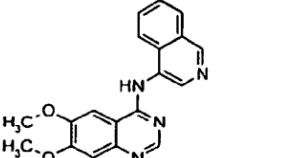
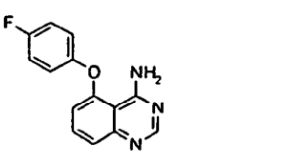
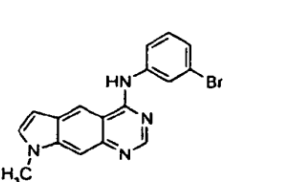
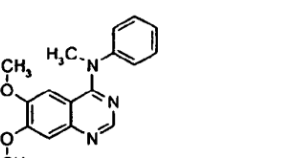
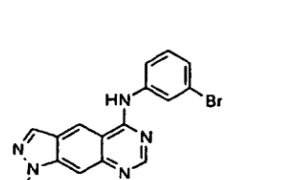
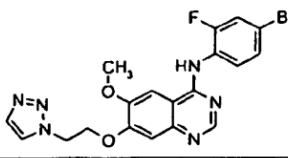
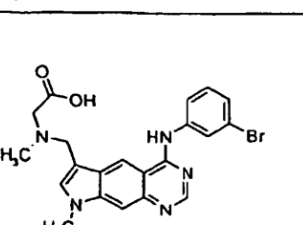
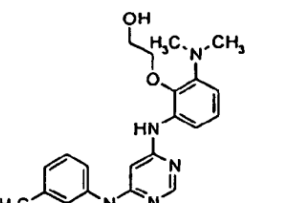
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			

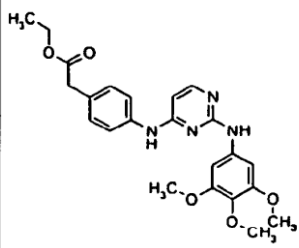
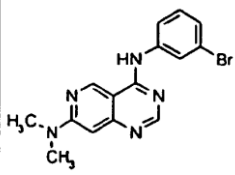
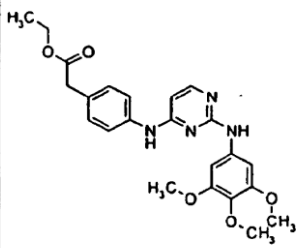
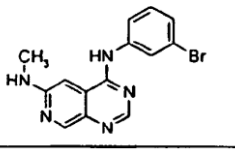
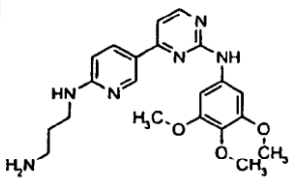
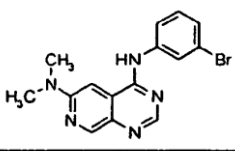
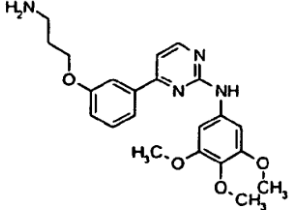
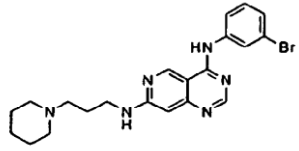
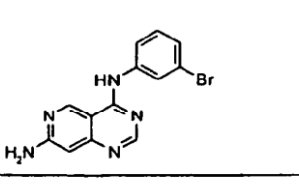
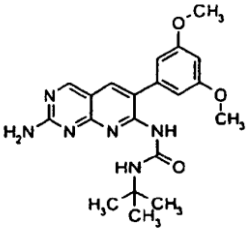
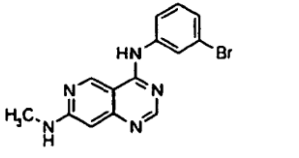
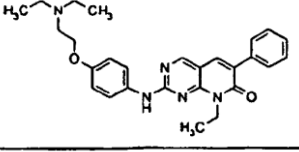
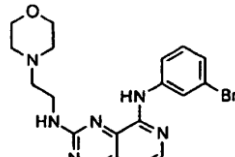
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			

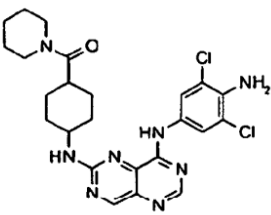
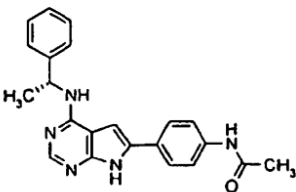
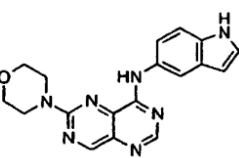
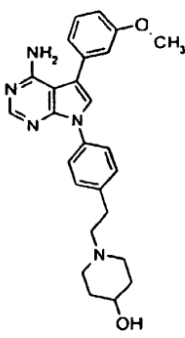
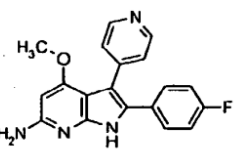
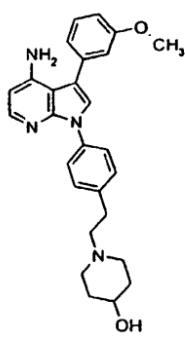
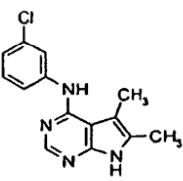
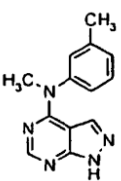
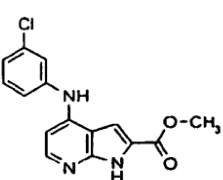
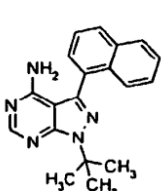
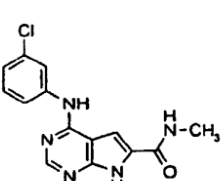
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			

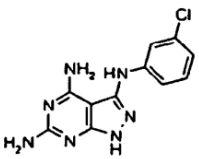
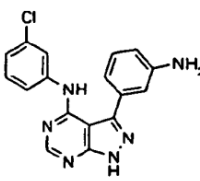
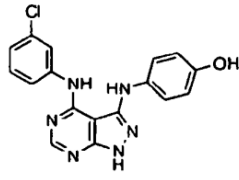
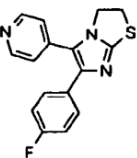
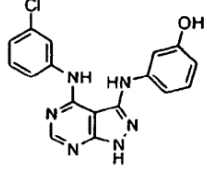
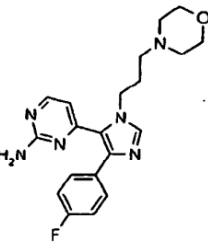
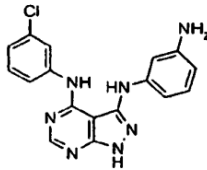
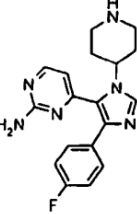
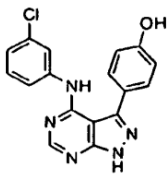
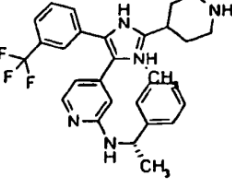
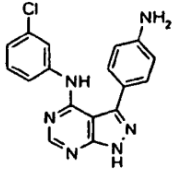
Estructura	Referencia	Estructura	Referencia
			
			
			
			
			

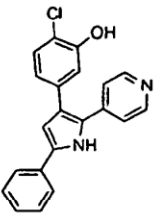
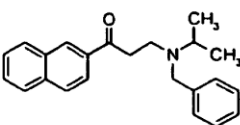
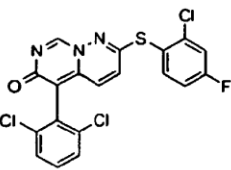
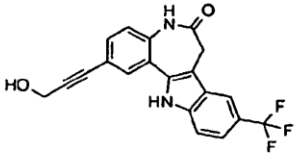
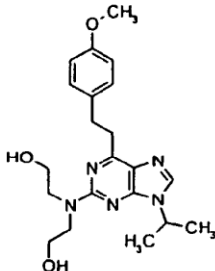
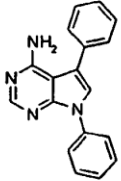
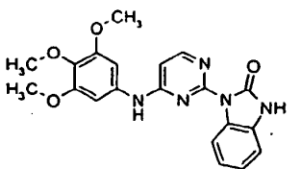
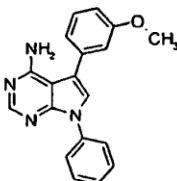
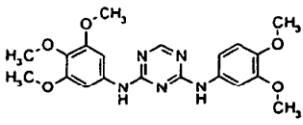
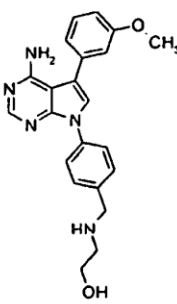
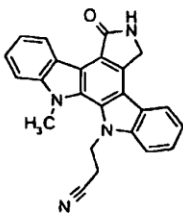
Estructura	Referencia	Estructura	Referencia
			
			J. Med. Chem. 2001,44,4615
			J. Med. Chem. 2001,44,4615
			J. Med. Chem. 2001,44,4615
			J. Med. Chem. 2001,44,4628
			J. Med. Chem. 2001,44,4628
			J. Med. Chem. 2001,44,4628

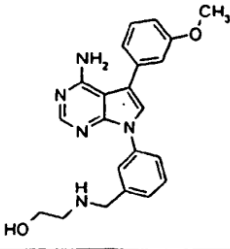
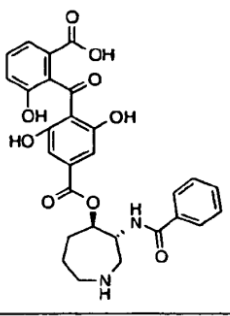
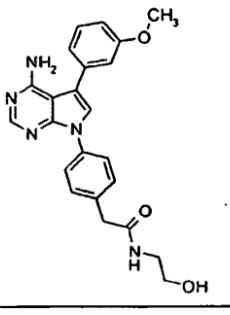
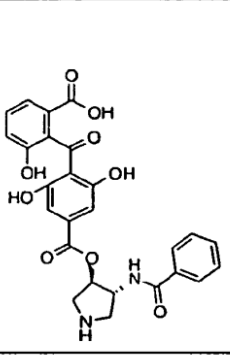
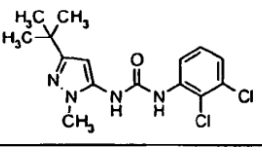
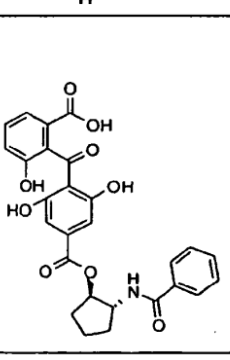
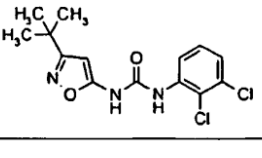
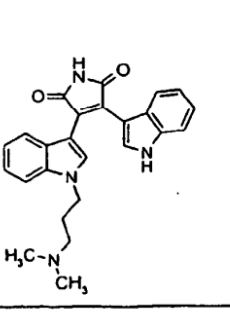
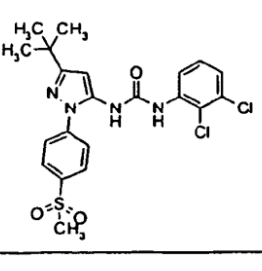
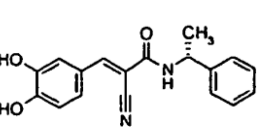
Estructura	Referencia	Estructura	Referencia
	J. Med. Chem. 2001,44,4628		
			
			Proc. Am. Assoc. Can Res. 1999, 40, 117 & 121
			
			
			Proc. Am. Assoc. Can Res. 1999, 90, 69
			

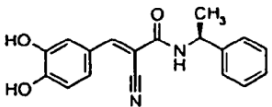
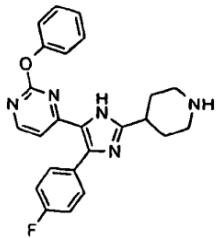
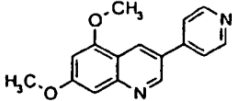
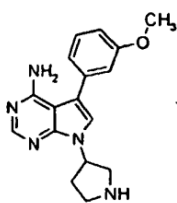
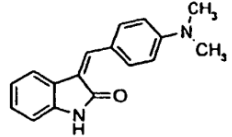
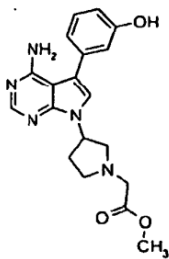
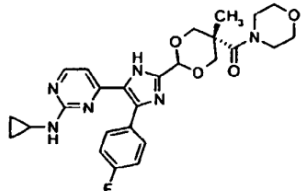
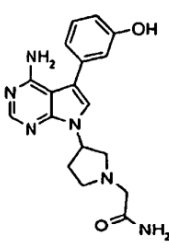
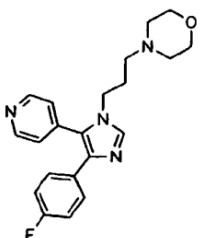
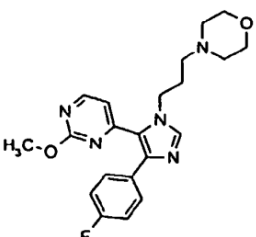
Estructura	Referencia	Estructura	Referencia
			J. Med. Chem. 1995, 38, 3780
			J. Med. Chem. 1995, 38, 3780
			
			J. Med. Chem. 1997, 40, 3915
	J. Med. Chem. 1995, 38, 3780		
	J. Med. Chem. 1995, 38, 3780		J. Med. Chem. 1998, 41, 4365
			J. Med. Chem. 1997, 40, 1820

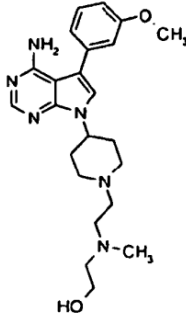
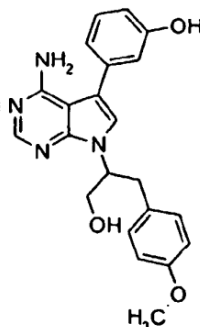
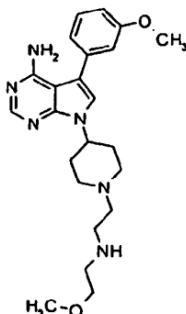
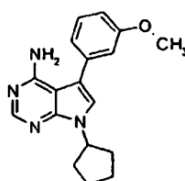
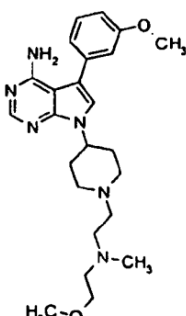
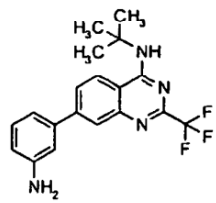
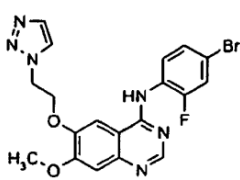
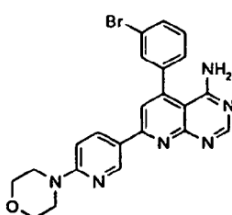
Estructura	Referencia	Estructura	Referencia
			
			
	J. Med. Chem. 1998, 41, 4196		
			
			
			

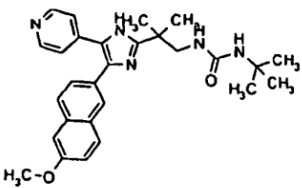
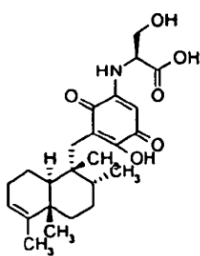
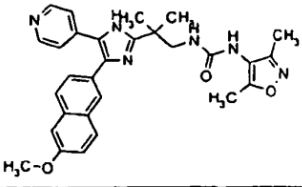
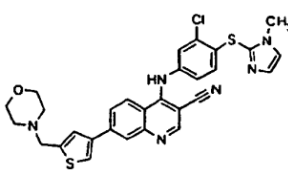
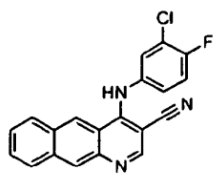
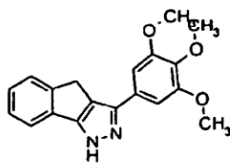
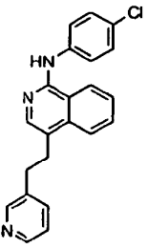
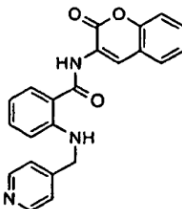
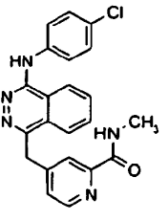
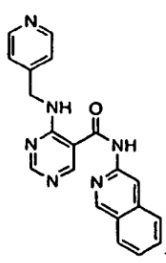
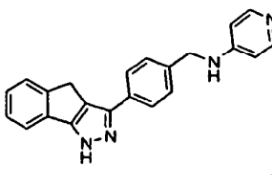
Estructura	Referencia	Estructura	Referencia
			
			Ann. N. Y. Acad. Sci. 1993, 696, 149
			
			
			J. Med. Chem. 1996, 39, 5215
			

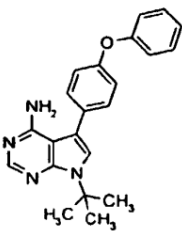
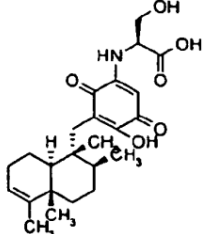
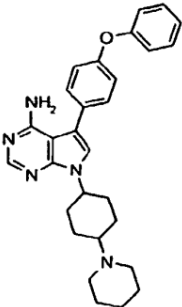
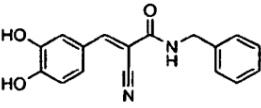
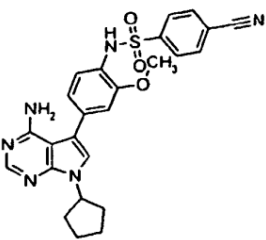
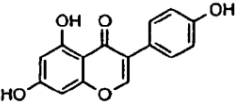
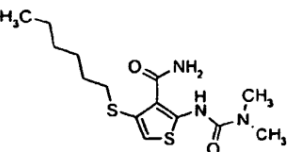
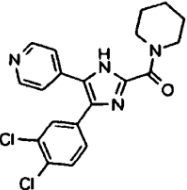
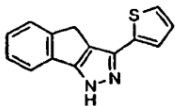
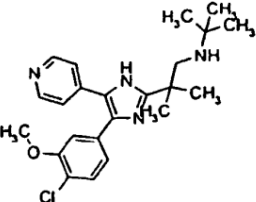
Estructura	Referencia	Estructura	Referencia
	Proc. Am. Assoc. Can. Res. 1998, 39, 558.		Bioorg. Med. Chem. Lett. 2000, 10, 575-579
	J. Med. Chem. 2001, 44, 3417		Bioorg. Med. Chem. Lett. 2000, 10, 567-569
	J. Biol. Chem. 1997, 272, 29207		
	WO 01/060816		Bioorg. Med. Chem. Lett. 2000, 10m 945-949
	WO 01/025220		Bioorg. Med. Chem. Lett. 2000, 10, 945-949
			

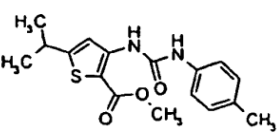
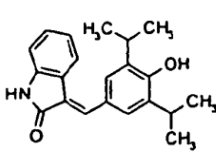
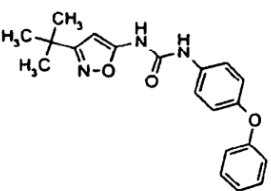
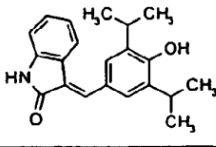
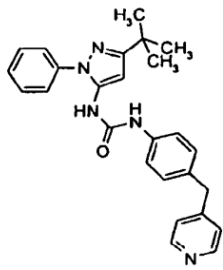
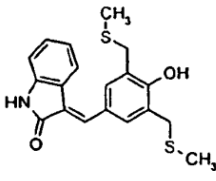
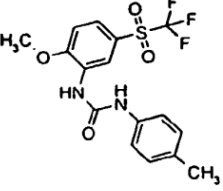
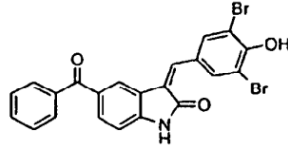
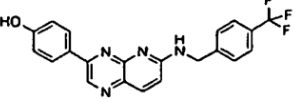
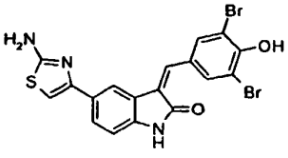
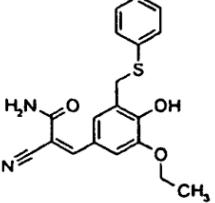
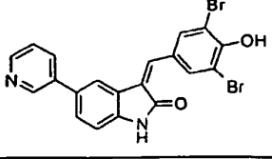
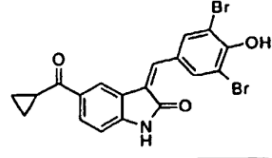
Estructura	Referencia	Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2000, 10, 945-949		
	Bioorg. Med. Chem. Lett. 2000, 10, 945-949		Bioorg. Med. Chem. Lett. 1996, 6, 1759-1764
	Bioorg. Med. Chem. Lett. 2000, 10, 2047-2050		Bioorg. Med. Chem. Lett. 1996, 6, 1759-1764
	Bioorg. Med. Chem. Lett. 2000, 10, 2047-2050		J. Biol. Chem. 1991, 266, 15771
	Bioorg. Med. Chem. Lett. 2000, 10, 2051-2054		J. Med. Chem. 1991, 34, 1896

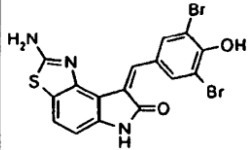
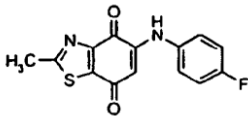
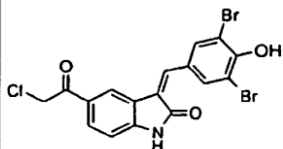
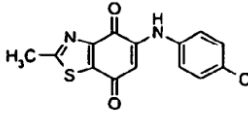
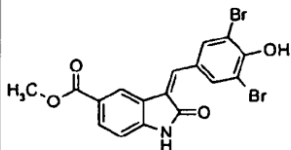
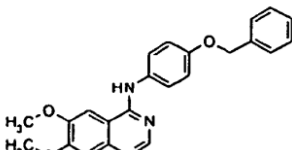
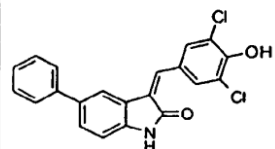
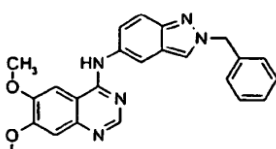
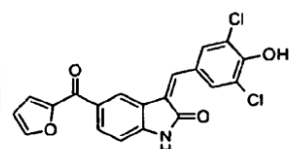
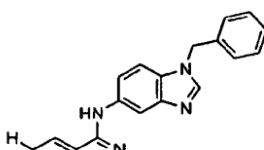
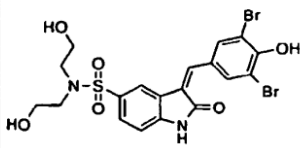
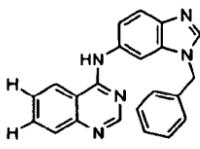
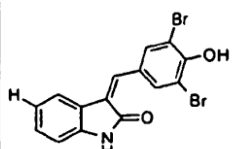
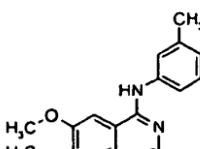
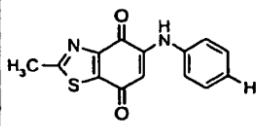
Estructura	Referencia	Estructura	Referencia
			
	J. Med. Chem. 1994, 37, 2627		Bioorg. Med. Chem. Lett. 2001, 11, 1123-1126
	J. Med. Chem. 1999, 41, 2588		Bioorg. Med. Chem. Lett. 2001, 11, 853-856
	Bioorg. Med. Chem. Lett. 2001, 11, 693-696		Bioorg. Med. Chem. Lett. 2001, 11, 853-856
			Bioorg. Med. Chem. Lett. 1998, 8, 3111-3116
	Bioorg. Med. Chem. Lett. 1998, 8, 3111-3116		

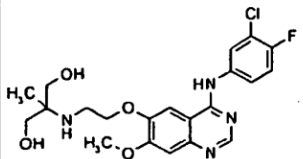
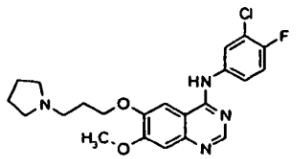
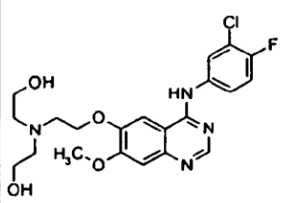
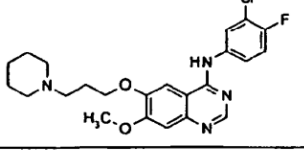
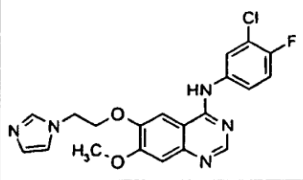
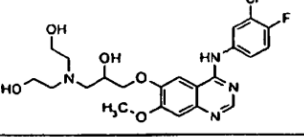
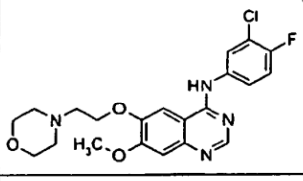
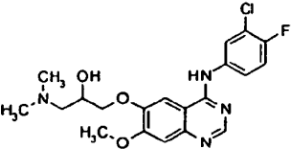
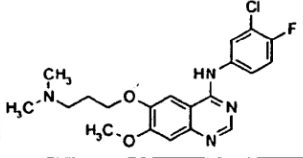
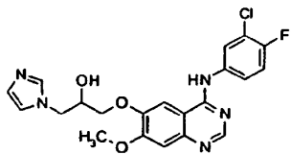
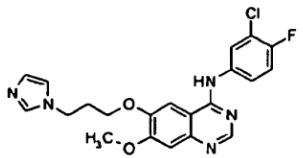
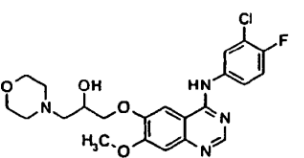
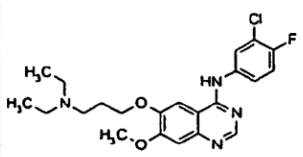
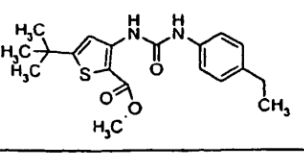
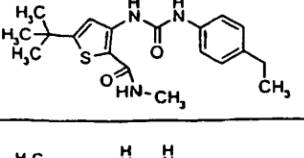
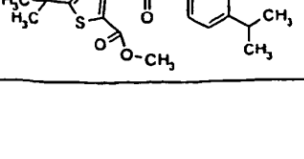
Estructura	Referencia	Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2001, 11, 853-856		Bioorg. Med. Chem. Lett. 2001, 11, 849-852
	Bioorg. Med. Chem. Lett. 2001, 11, 853-856		Bioorg. Med. Chem. Lett. 2001, 11, 840-852
			Bioorg. Med. Chem. Lett. 2001, 11, 1157-1160
			
			J. Med. Chem. 2001, 44, 2133-2138

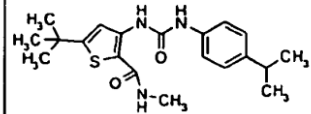
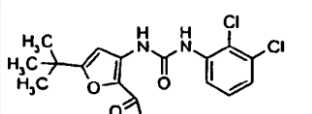
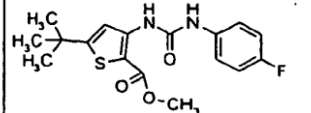
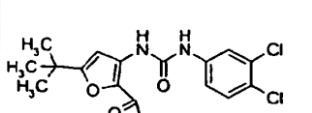
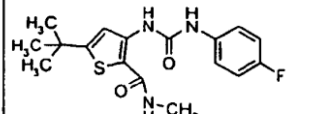
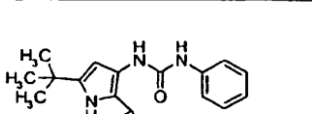
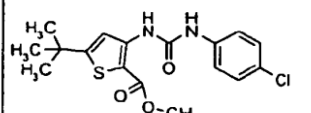
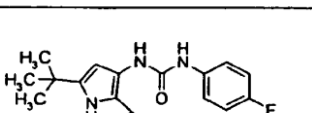
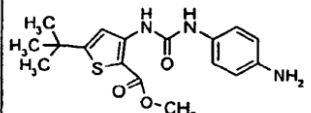
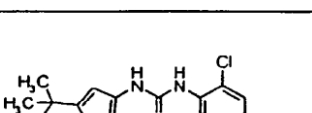
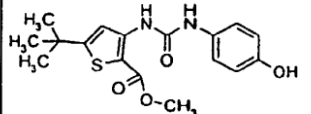
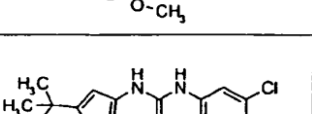
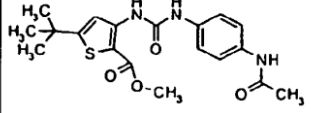
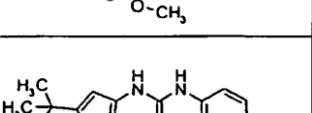
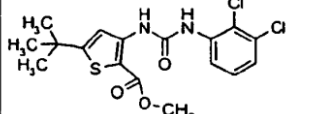
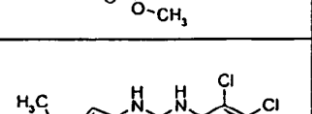
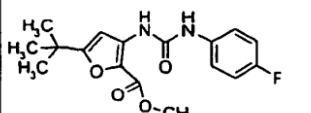
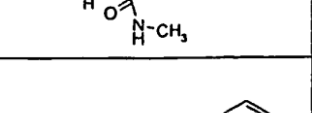
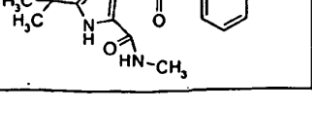
Estructura	Referencia	Estructura	Referencia
	WO 01/37835		J. Am. Chem. Soc. 2001, 123, 11586-11593
			WO 01/072711
	US 01/051620	22	
	WO 01/087846		WO 01/058899
	WO 01/085719		WO 01/010859
	WO 01/085715		WO 00/059901

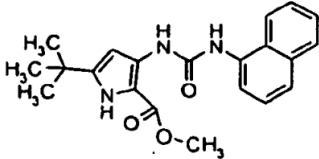
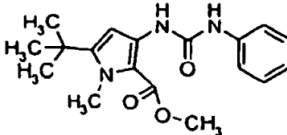
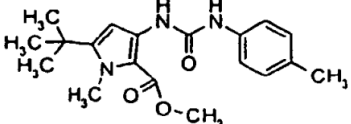
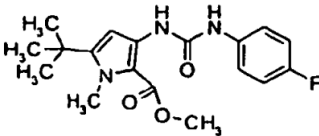
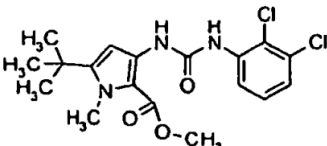
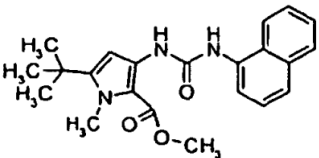
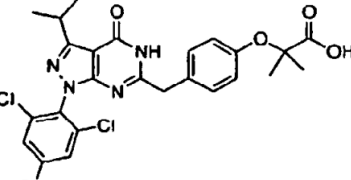
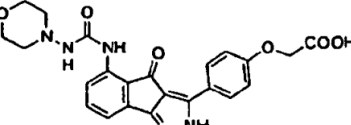
Estructura	Referencia	Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2000, 10, 2167-2170		J. Am. Chem. Soc. 2001, 123, 11586-11593
	WO 00/017203		Atherosclerosis (Shannon, Ireland) (2002), 160(1), 123-132.
	WO 00/017202		
	WO 9962890		WO 01/066540
	WO 9917769		WO 01/066539

Estructura	Referencia	Estructura	Referencia
	US 6187799		Biochem. Biophys. Res. Commun. 1987, 147, 322
	WO 99/32106		Biochem. Biophys. Res. Commun. 1987, 147, 322
	WO 99/32455		Biochem. Biophys. Res. Commun. 1987, 147, 322
	WO 99/32436		Bioorg. Med. Chem. Lett. 2000, 10, 223-226
	WO 99/17759		Bioorg. Med. Chem. Lett. 2000, 10, 223-226
	Biochem. Biophys. Res. Commun. 1987, 147, 322		Bioorg. Med. Chem. Lett. 2000, 10, 223-226
			Bioorg. Med. Chem. Lett. 2000, 10, 223-226

Estructura	Referencia	Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2000, 10, 223-226		Bioorg. Med. Chem. Lett. 2000, 10, 461-464
	Bioorg. Med. Chem. Lett. 2000, 10, 223-226		Bioorg. Med. Chem. Lett. 2000, 10, 461-464
	Bioorg. Med. Chem. Lett. 2000, 10, 223-226		Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405
	Bioorg. Med. Chem. Lett. 2000, 10, 223-226		Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405
	Bioorg. Med. Chem. Lett. 2000, 10, 223-226		Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405
	Bioorg. Med. Chem. Lett. 2000, 10, 223-226		Bioorg. Med. Chem. Lett. 2001, 11, 1401-1405
	Bioorg. Med. Chem. Lett. 2000, 10, 223-226		Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914
	Bioorg. Med. Chem. Lett. 2000, 10, 461-464		

Estructura	Referencia	Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914		Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914
	Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914		Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914
	Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914		Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914
	Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914		Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914
	Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914		Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914
	Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914		Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914
	Bioorg. Med. Chem. Lett. 2001, 11, 1911-1914		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
			Bioorg. Med. Chem. Lett. 2001, 11, 9-12
			Bioorg. Med. Chem. Lett. 2001, 11, 9-12

Estructura	Referencia	Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12		Bioorg. Med. Chem. Lett. 2001, 11, 9-12
			Bioorg. Med. Chem. Lett. 2001, 11, 9-12

Estructura	Referencia
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	Bioorg. Med. Chem. Lett. 2001, 11, 9-12
	
	

REIVINDICACIONES

1. Método de identificación de una secuencia polipeptídica que se une a un ligando especificado por el usuario, que comprende:
 - (i) proporcionar un ligando híbrido que tiene la fórmula general: $R1-Y-R2$, en la que: R1 es un primer ligando, R2 es un ligando especificado por el usuario e Y es un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es un número entero de 2 a 25;
 - (ii) introducir el ligando híbrido en una población de células, conteniendo cada célula un sistema de exploración de ligando híbrido que incluye:
 - (a) un gen indicador unido operativamente a una secuencia reguladora de la transcripción, incluyendo dicha secuencia reguladora una secuencia de ADN que se une a un dominio de unión a ADN;
 - (b) un primer gen quimérico que codifica un primer polipéptido de fusión que comprende: un dominio de unión a ligando P1 y un dominio seleccionado de un dominio de unión a ADN o un dominio de activación de la transcripción, en el que el dominio de unión a ligando se une al primer ligando R1; y
 - (c) un segundo gen quimérico que codifica un segundo polipéptido de fusión que comprende: un dominio de unión a ligando candidato P2 para el ligando especificado por el usuario R2 y un dominio seleccionado de un dominio de unión a ADN o un dominio de activación de la transcripción;

en el que uno de los dos polipéptidos de fusión contiene un dominio de unión a ADN y el otro polipéptido de fusión contiene un dominio de activación de la transcripción;
 - (iii) permitir que el ligando híbrido se una al dominio de unión a ligando del primer polipéptido de fusión a través del primer ligando R1 y contacte con el dominio de unión a ligando candidato del segundo polipéptido de fusión a través del ligando especificado por el usuario R2 de modo que, si R2 se une al dominio de unión a ligando candidato, se produce un aumento en el nivel de transcripción del gen indicador;
 - (iv) identificar una célula de unión a ligando positiva en la que se haya producido un aumento en el nivel de transcripción del gen indicador; y
 - (v) identificar la secuencia de ácido nucleico del segundo gen quimérico que codifica el dominio de unión a ligando candidato que se une al ligando especificado por el usuario R2, identificando de este modo una secuencia polipeptídica que se une a un ligando especificado por el usuario.
2. Método de la reivindicación 1, en el que la secuencia de ácido nucleico que codifica el polipéptido de dominio de unión a ligando candidato del segundo polipéptido de fusión es de una biblioteca seleccionada de: una biblioteca de oligonucleótidos sintéticos, una biblioteca de ADNc, una biblioteca de fragmentos de ADN genómicos bacterianos o una biblioteca de fragmentos de ADN genómicos eucariotas.
3. Método de la reivindicación 1, que comprende además:
 - (a) determinar la afinidad de unión del ligando híbrido a los dominios de unión a ligando P1 y/o P2; o
 - (b) determinar los efectos del ligando híbrido que son independientes de la formación de un complejo trimérico que comprende el ligando híbrido, P1 y P2; o
 - (c) la etapa de realizar al menos un método separado adicional para confirmar que la transcripción del gen indicador depende de la presencia del ligando híbrido y los dominios de unión a ligando P1 y P2.
4. Método de la reivindicación 3, en el que la determinación de la afinidad de unión en la etapa (a) de la reivindicación 3 se realiza por resonancia de plasmón superficial.
5. Método de la reivindicación 3, en el que dicho método separado adicional en la etapa (c) de la reivindicación 3 se selecciona de: un método de ensayo de crecimiento en halo, un ensayo de crecimiento en placa de microtitulación o un ensayo de crecimiento por detección de la fluorescencia.
6. Método de la reivindicación 3, en el que dicho método separado adicional en la etapa (c) de la reivindicación 3 se realiza individualmente sobre más de aproximadamente 10, 100, 1000 ó 10000 tipos celulares de unión a ligando positivos diferentes identificados en la etapa (iv) de la reivindicación 1.
7. Método de identificación de un ligando de un polipéptido especificado por el usuario, que comprende:
 - (i) proporcionar al menos un ligando híbrido candidato que tiene la fórmula general $R1-Y-R2$, en la que R1 es un primer ligando, R2 es un ligando candidato e Y es un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es un número entero de 2 a 25;
 - (ii) introducir el ligando híbrido candidato en al menos una célula que contiene un sistema de exploración de ligando híbrido, que incluye:

- (a) un gen indicador unido operativamente a una secuencia reguladora de la transcripción, incluyendo dicha secuencia reguladora una secuencia de ADN que se une a un dominio de unión a ADN;
- (b) un primer gen quimérico que codifica un primer polipéptido de fusión, que comprende: un dominio de unión a ligando y un dominio seleccionado de: un dominio de unión a ADN o un dominio de activación de la transcripción, en el que el dominio de unión a ligando se une al primer ligando R1; y
- (c) un segundo gen quimérico que codifica un segundo polipéptido de fusión, que comprende: un dominio de unión a ligando especificado por el usuario para el ligando candidato R2 y un dominio seleccionado de: un dominio de unión a ADN o un dominio de activación de la transcripción;
- en el que uno de los dos polipéptidos de fusión contiene un dominio de unión a ADN y el otro polipéptido de fusión contiene un dominio de activación de la transcripción;
- (iii) permitir que el ligando híbrido candidato se una al dominio de unión a ligando del primer polipéptido de fusión a través del primer ligando R1 y contacte con el dominio de unión a ligando especificado por el usuario del segundo polipéptido de fusión a través del ligando candidato R2 de modo que, si el dominio de unión a ligando especificado por el usuario se une al ligando candidato R2, se produce un aumento en el nivel de transcripción del gen indicador;
- (iv) identificar el ligando híbrido candidato que causa un aumento en el nivel de transcripción del gen indicador en la célula, identificando de este modo el ligando candidato en el ligando híbrido candidato como un ligando para el polipéptido especificado por el usuario.
8. Método de la reivindicación 1 ó 7, en el que el gen indicador se selecciona de: HIS3, LEU2, TRP2, TRP1, ADE2, LYS2, URA3, CYH1, CAN1, *lacZ*, *gfp* o CAT.
9. Método *in vitro* para inducir o permitir la detección de un acontecimiento biológicamente detectable, que comprende:
- (i) proporcionar al menos una célula que comprende al menos una secuencia de ácido nucleico que codifica un polipéptido de fusión, que incluye:
- (a) al menos un dominio de unión a ligando; y
- (b) un primer dominio funcional que por sí mismo no es capaz de inducir o permitir la detección del acontecimiento detectable;
- (ii) proporcionar un ligando híbrido de fórmula general R1-Y-R2, en la que R1 es diferente de R2, al menos uno de R1 y R2 no es un péptido, R1 o R2 representa un ligando que se une a dicho dominio de unión a ligando; Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO₂ y n es un número entero de 2 a 25; y en el que la unión de dicho ligando híbrido a dicho dominio de unión a ligando pone al primer dominio funcional en las proximidades de un segundo dominio funcional, induciendo o permitiendo de este modo la detección del acontecimiento detectable; y
- (iii) exponer dicha al menos una célula a una cantidad eficaz de dicho ligando híbrido para poner el primer dominio funcional en las proximidades de un segundo dominio funcional;
- induciendo o permitiendo de este modo la detección del acontecimiento biológicamente detectable.
10. Método *in vitro* para investigar la relación de estructura-actividad de un ligando con un dominio de unión a ligando, que comprende:
- (i) proporcionar un ligando híbrido R1-Y-R2, en el que
- (a) R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, derivado de 2,4-diaminopteridina o ciclosporina, o un derivado de los mismos con modificaciones estructurales minoritarias;
- (b) Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO₂ y n es un número entero de 2 a 25; y
- (c) R2 representa un segundo ligando especificado por el usuario que es diferente de R1 y se selecciona de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o un nucleótido;

- (ii) proporcionar células que comprenden una proteína de fusión que incluye:
- (a) al menos un dominio de unión a ligando; y
- (b) un primer dominio funcional heterólogo respecto al dominio de unión a ligando, que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional;
- 5 en el que se proporciona una pluralidad de ligandos híbridos que comprenden variantes estructurales de dicho segundo ligando R2 en la etapa (i), o se proporciona una pluralidad de proteínas de fusión que comprenden variantes estructurales de dicho dominio de unión a ligando en la etapa (ii);
- 10 (iii) exponer dichas células que comprenden cada proteína de fusión a una cantidad eficaz de cada ligando híbrido, de modo que el primer dominio funcional puede ponerse en las proximidades de un segundo dominio funcional, induciendo o permitiendo de este modo la detección de un acontecimiento detectable;
- (iv) medir la presencia, cantidad o actividad de cualquier acontecimiento detectable así inducido o permitido en la etapa (iii), investigando de este modo la relación de estructura-actividad entre dicho segundo ligando y el
- 15 dominio de unión a ligando.
11. Método de la reivindicación 9 ó 10, en el que dicho primer dominio funcional se selecciona de: un dominio de unión a ADN, un dominio de activación de la transcripción, un subdominio carboxi-terminal de una ubiquitina de tipo salvaje, un subdominio amino-terminal de una ubiquitina o un subdominio amino-terminal de ubiquitina mutante de asociación reducida.
- 20 12. Método de cualquiera de las reivindicaciones 1, 7, 9 y 10, en el que dicha célula o células se seleccionan de células de levadura, mamífero o procariontes.
13. Método de la reivindicación 12, en el que dicha célula o células son levadura.
14. Método de la reivindicación 12, en el que dicha célula o células son de mamífero.
- 25 15. Método de la reivindicación 1, 7, 9 y 10, en el que el dominio de unión a ligando que se une a R1 es la dihidrofolato reductasa (DHFR).
16. Método para identificar un ligando híbrido que tiene la estructura general R1-Y-R2 adecuada para un ensayo *in vivo*, en el que dicho ensayo implica:
- (i) el uso de un ligando híbrido y
- (ii) de al menos un polipéptido de fusión que incluye:
- 30 (a) al menos un dominio de unión a ligando P; y
- (b) un dominio funcional que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable;
- y en el que dicho método implica las etapas de:
- 35 (iii) sintetizar una pluralidad de ligandos híbridos R1-Y-R2 que difieren por una pluralidad de enlazadores Y diferentes, en el que:
- (a) dicho enlazador tiene la estructura general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es un número entero de 2 a 25, y la pluralidad de enlazadores difieren en n,
- (b) R1 y R2 son diferentes, y
- (c) al menos uno de R1 y R2 no es un péptido; y
- 40 (iv) ensayar cada ligando híbrido en dicha pluralidad de ligandos híbridos individualmente para determinar su eficacia para inducir o permitir la detección del acontecimiento detectable; y
- (v) seleccionar un ligando híbrido con un enlazador particular que posea una eficacia adecuada para inducir o permitir la detección del acontecimiento detectable.
17. Composición que comprende:
- 45 (i) Un ligando híbrido representado por la fórmula general: R1-Y-R2, en la que:
- (a) R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, derivado de 2,4-diaminopteridina, novobiocina, maltosa, glutatión, biotina, vitamina D,

dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel o ciclosporina, o un derivado de los mismos con modificaciones estructurales menores;

- (b) Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es un número entero de 2 a 25;

5 (c) R2 representa un segundo ligando especificado por el usuario diferente de R1 seleccionado de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o nucleótido;

- (ii) un polipéptido de fusión que incluye:

(a) al menos un dominio de unión a ligando; y

15 (b) un primer dominio funcional heterólogo al dominio de unión a ligando que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional.

18. Composición de la reivindicación 17, en la que la composición es un complejo.

19. Composición de la reivindicación 17, en la que la composición se proporciona en un entorno seleccionado de: una célula, un recipiente, un kit, una solución o un medio de cultivo.

20. Uso de un compuesto representado por la fórmula general R1-Y-R2, en la que:

- (i) R1 representa un primer ligando seleccionado de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, 2,4-diaminopteridina o ciclosporina, o un derivado de los mismos con modificaciones estructurales menores;

- (ii) Y representa un enlazador de polietileno que tiene la fórmula general $(CH_2-X-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es un número entero de 2 a 25; y

30 (iii) R2 representa un segundo ligando especificado por el usuario diferente de R1 seleccionado de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o un nucleótido;

35 para identificar la unión de un polipéptido a un ligando especificado por el usuario en un sistema de ensayo de triple híbrido;

para inducir o permitir la detección de un acontecimiento biológicamente detectable en un sistema de ensayo de triple híbrido;

para investigar la relación de estructura-actividad de un ligando respecto a un dominio de unión a ligando en un sistema de ensayo de triple híbrido;

40 para identificar un compuesto que inhibe una interacción entre un ligando y un polipéptido en un sistema de ensayo de triple híbrido; o

como ligando híbrido en cualquiera de los métodos de acuerdo con las reivindicaciones 1 a 16 y 36 a 38.

21. Método de cualquiera de las reivindicaciones 1, 7, 9 y 16 en el que:

45 (a) R1 se selecciona de: un esteroide, ácido retinoico, antibiótico beta-lactámico, canabinoide, ácido nucleico, polipéptido, FK506, derivado de FK506, rapamicina, tetraciclina, metotrexato, novobiocina, maltosa, glutatión, biotina, vitamina D, dexametasona, estrógeno, progesterona, cortisona, testosterona, níquel, derivado de 2,4-diaminopteridina o ciclosporina, o un derivado de los mismos con modificaciones estructurales menores; y

50 (b) R2 es diferente de R1 y se selecciona de: un péptido, ácido nucleico, carbohidrato, polisacárido, lípido, prostaglandina, haluro de acilo, alcohol, aldehído, alcano, alqueno, alquino, alquilo, haluro de alquilo, alcaloide, amina, hidrocarburo aromático, éster de sulfonato, carboxilato ácido, haluro de arilo, éster, fenol, éter, nitrilo, anhídrido del ácido carboxílico, amida, sal de amonio cuaternario, imina, enamina, óxido de amina, cianohidrina, organocadmio, aldol, organometálico, hidrocarburo aromático, nucleósido o un

nucleótido;

22. Método de cualquiera de las reivindicaciones 1, 10 y 21, o la composición de la reivindicación 17, o el uso de la reivindicación 20, en el que el primer ligando R1 del ligando híbrido se une al dominio de unión a ligando P1 con gran afinidad.
- 5 23. Método de la reivindicación 22, en el que la afinidad de unión corresponde a una constante de disociación de ligando/proteína de unión a ligando K_D de menos de $1 \mu M$.
24. Método de cualquiera de las reivindicaciones 1, 10 y 21, o la composición de la reivindicación 17, o el uso de la reivindicación 20, en el que el primer ligando es capaz de formar un enlace covalente con el dominio de unión a ligando P1.
- 10 25. Método de cualquiera de las reivindicaciones 1, 10 y 21, o la composición de la reivindicación 17, o el uso de la reivindicación 20, en el que X es O.
26. Método de cualquiera de las reivindicaciones 1, 10 y 21, o la composición de la reivindicación 17, o el uso de la reivindicación 20, en el que Y es $(CH_2-O-CH_2)_n$, en el que $n = 2$ a 5.
27. Método, composición o uso de la reivindicación 26, en el que n es 3, 4 ó 5.
- 15 28. Método, composición o uso de la reivindicación 27, en el que n es 3.
29. Método, composición o uso de la reivindicación 27, en el que n es 5.
30. Método de la reivindicación 10 ó 21, o la composición de la reivindicación 17, o el uso de la reivindicación 20, en el que R1 se selecciona de dexametasona, metotrexato, un derivado de metotrexato, FK506, un derivado de FK506 o un derivado de 2,4-diaminopteridina.
- 20 31. Método, composición o uso de la reivindicación 30, en el que R1 es metotrexato.
32. Método, composición o uso de la reivindicación 30, o el uso de la reivindicación 20, en el que R1 es dexametasona.
33. Método de la reivindicación 10 ó 21, o la composición de la reivindicación 17, o el uso de la reivindicación 20, en el que al menos uno de R1 y R2 no es un péptido.
- 25 34. Método de cualquiera de las reivindicaciones 1, 10 y 21, o la composición de la reivindicación 17, o el uso de la reivindicación 20, en el que R1 es metotrexato, Y es $(CH_2-O-CH_2)_n$ y $n = 2$ a 5.
35. Método de cualquiera de las reivindicaciones 1, 10 y 21, o la composición de la reivindicación 17, o el uso de la reivindicación 20, en el que R2 se une a o inhibe una quinasa.
- 30 36. Método *in vitro* de identificación de un compuesto que inhibe una interacción entre un ligando y un polipéptido, que comprende las etapas de:
- (a) proporcionar un entorno en el que se produce dicha interacción, en el que el entorno es una célula que contiene una composición de cualquiera de las reivindicaciones 17, 18 y 22 a 35.
- (b) poner en contacto el entorno con un compuesto de ensayo; y
- (c) determinar si dicho compuesto de ensayo inhibe dicha interacción, identificando de este modo un compuesto que inhibe la interacción entre un ligando y un polipéptido.
- 35 37. Método para identificar un nuevo modulador de una proteína conocida determinada, incluyendo una quinasa, que usa una célula que contiene una composición de cualquiera de las reivindicaciones 17, 18 y 22 a 35.
38. Kit que comprende:
- (i) al menos un polinucleótido que incluye un fragmento de ADN unido a una secuencia codificante de un primer dominio funcional heterólogo respecto al fragmento de ADN que por sí mismo no es capaz de inducir o permitir la detección de un acontecimiento detectable, pero que es capaz de inducir o permitir la detección de un acontecimiento detectable cuando se pone en las proximidades de un segundo dominio funcional;
- (ii) un ligando híbrido de estructura general R1-Y-R2, en la que R2 es diferente de R1, uno de R1 y R2 es un ligando no peptídico e Y es un enlazador
- 40 45 en el que dicho kit se caracteriza además por el hecho de que:
- (a) uno de R1 y R2 se une a o inhibe una quinasa;
- (b) Y es de estructura general $(CH_2-O-CH_2)_n$, en la que X representa O, S, SO o SO_2 y n es un número entero de 2 a 25; o

- (c) el primer dominio funcional es un subdominio carboxi-terminal de ubiquitina o un subdominio amino-terminal de ubiquitina.

FIG. 1A

Esquema 1

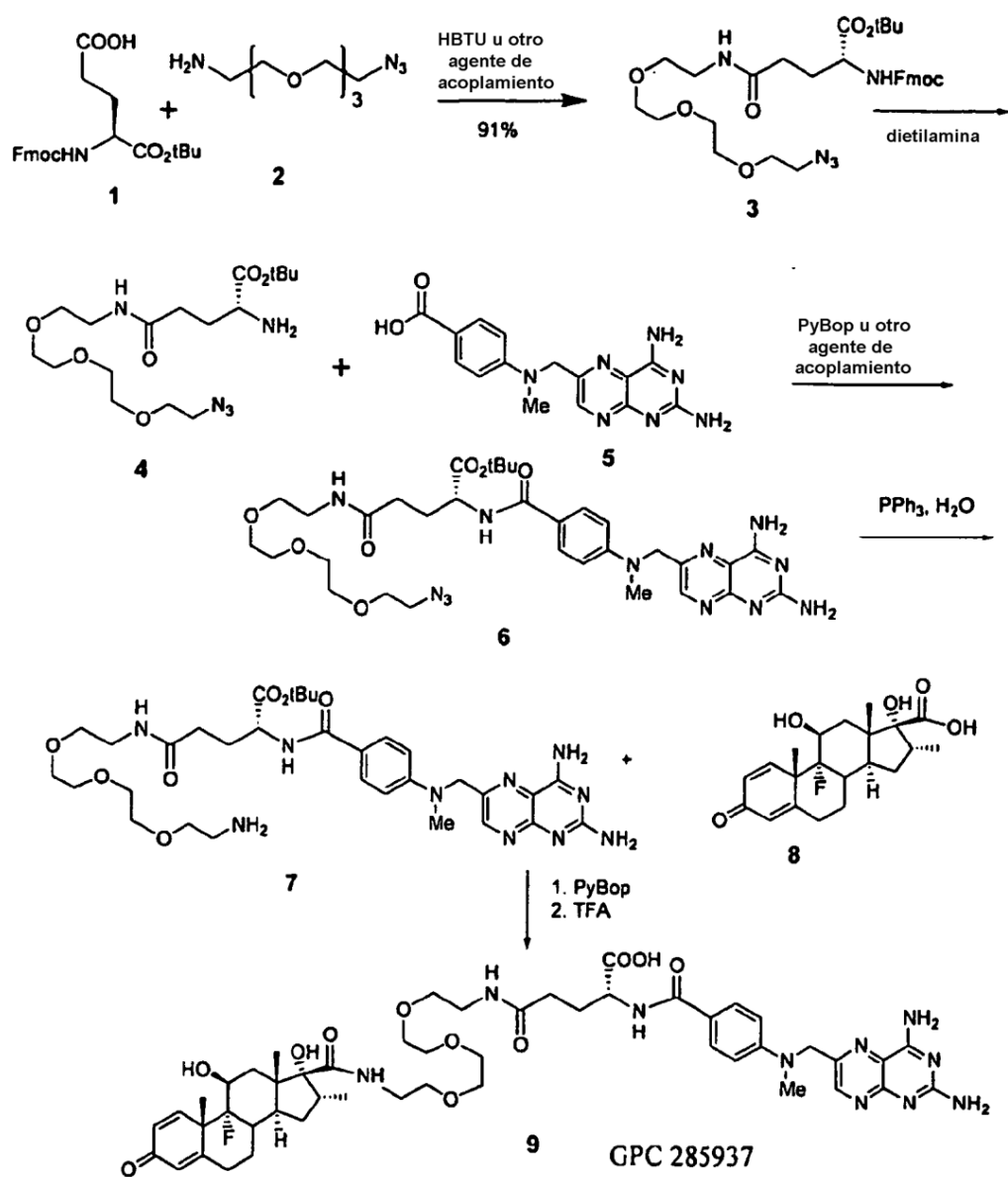


FIG. 1B
Esquema 2

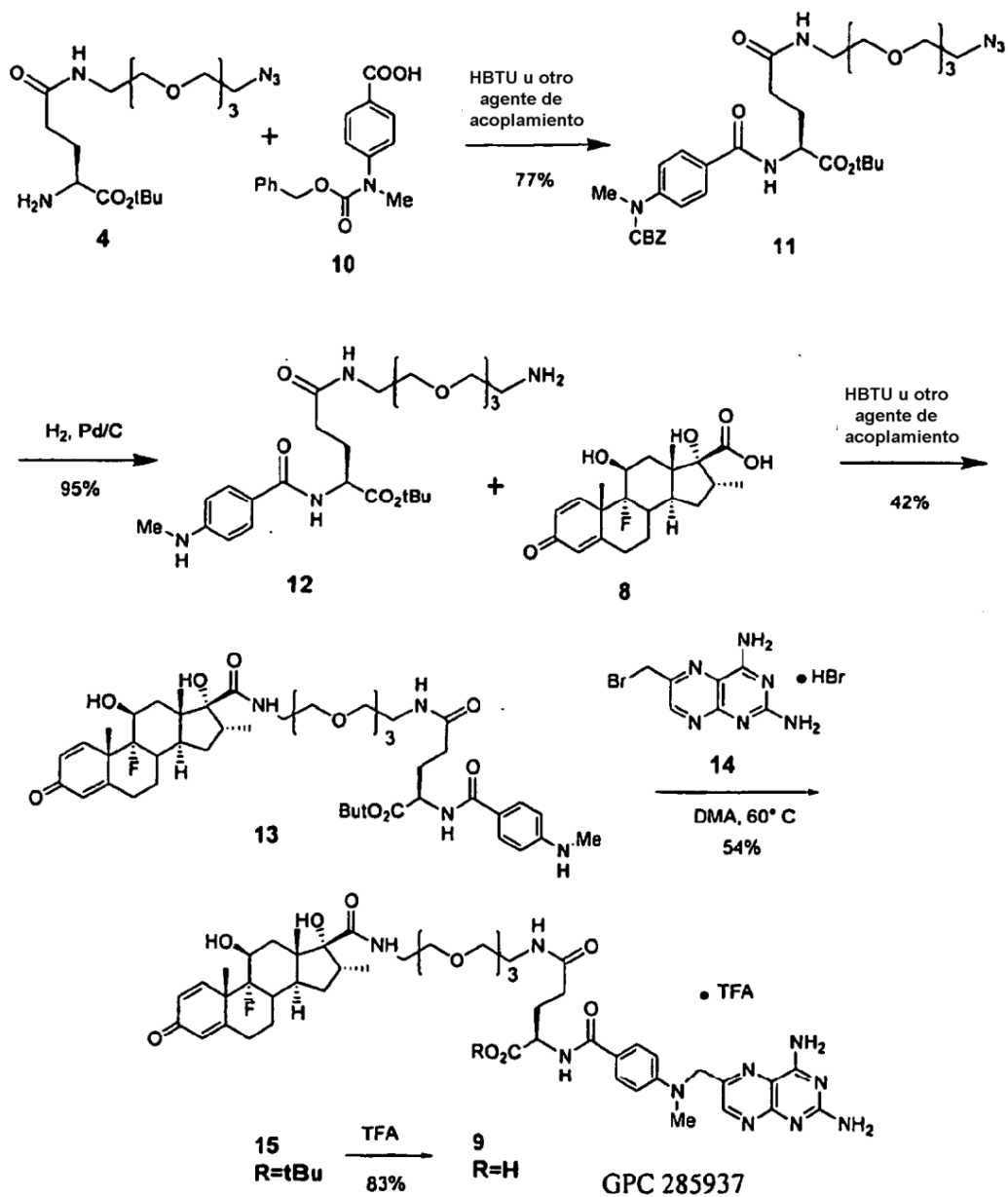


FIG. 1C

Esquema 3

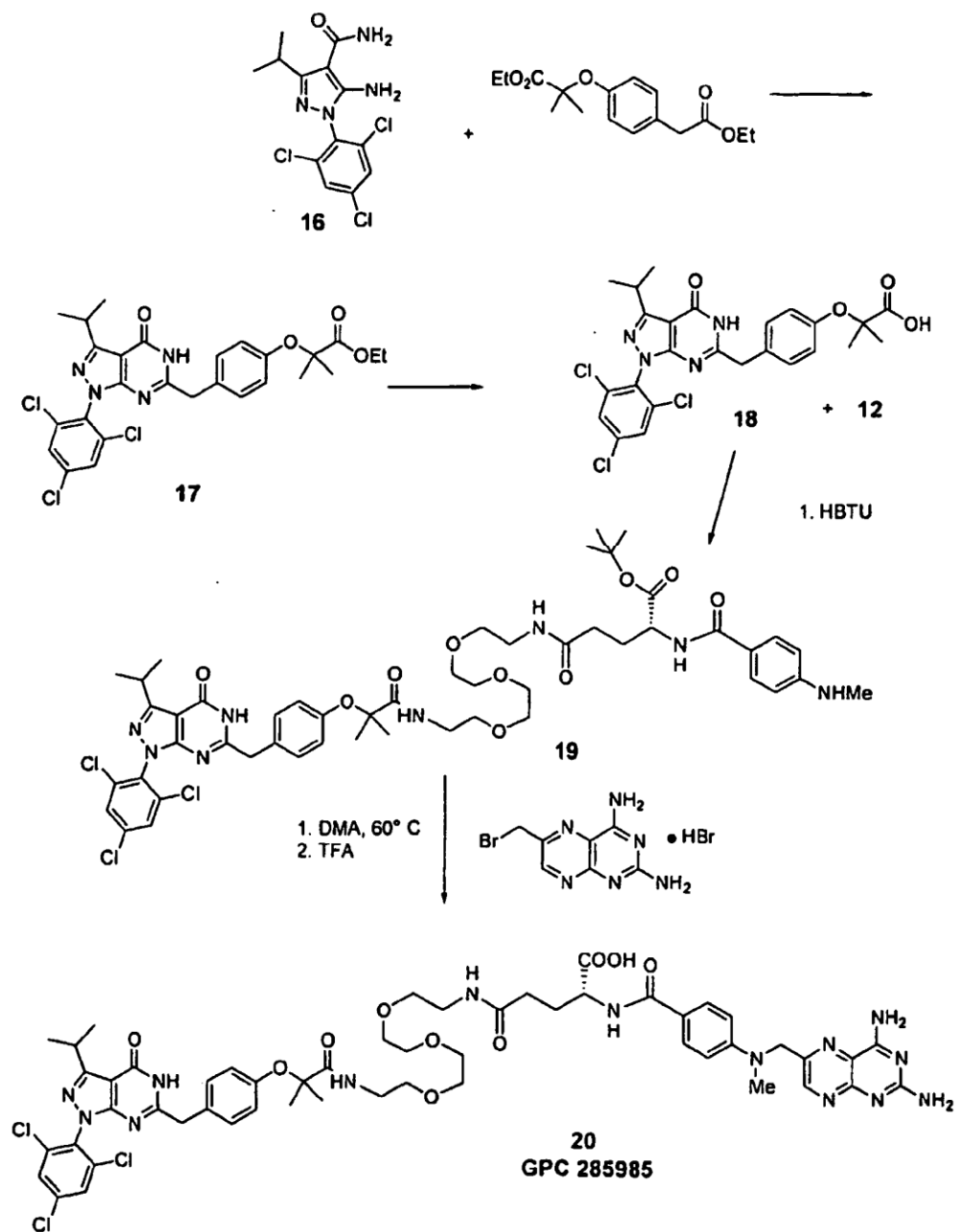


FIG. 1D

Esquema 4

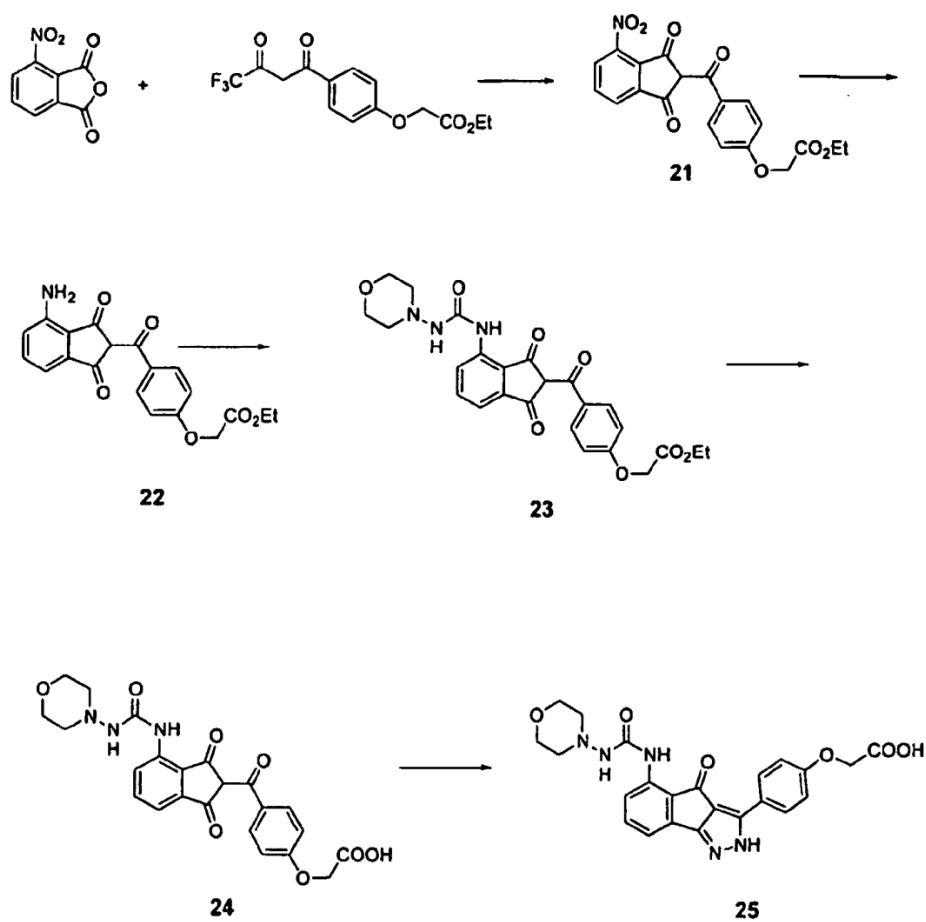


FIG. 1E

Esquema 5

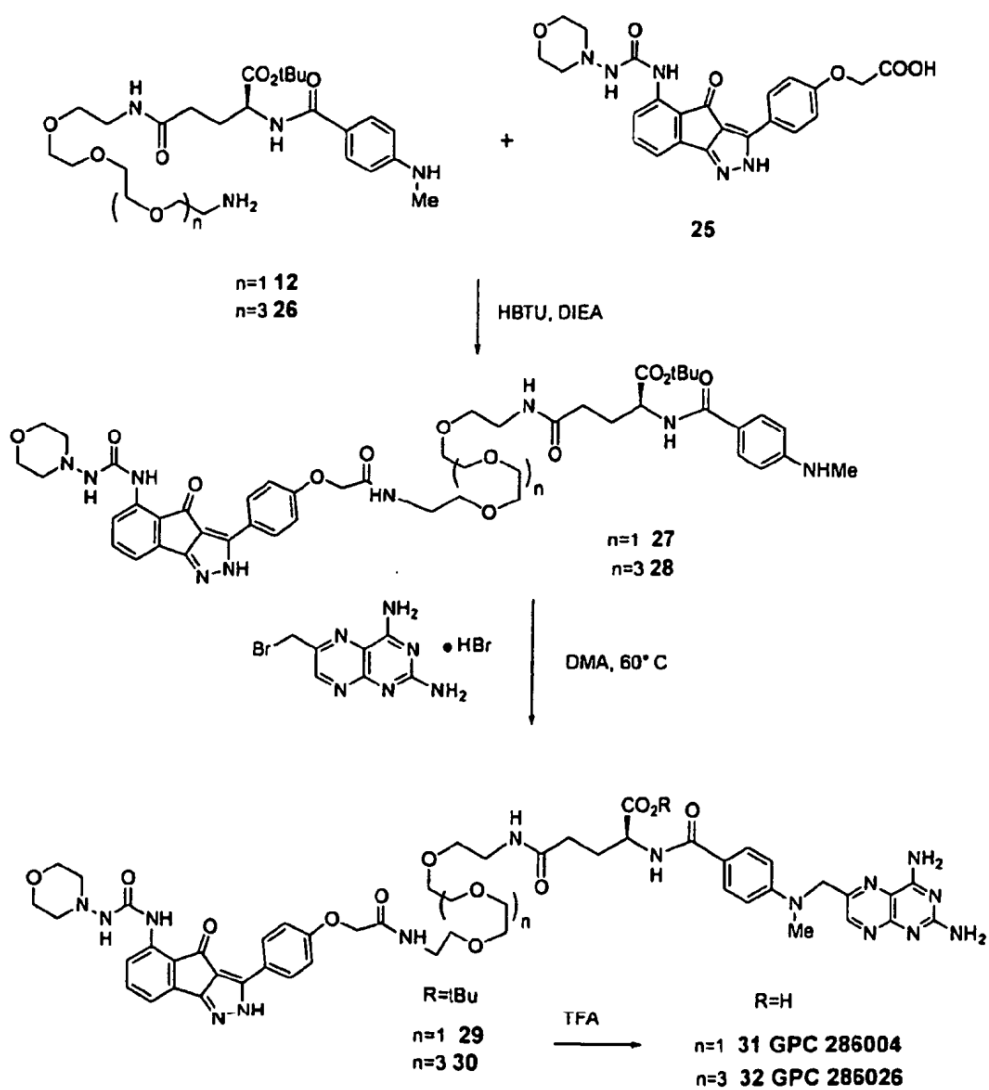


FIG. 1F
Esquema 6

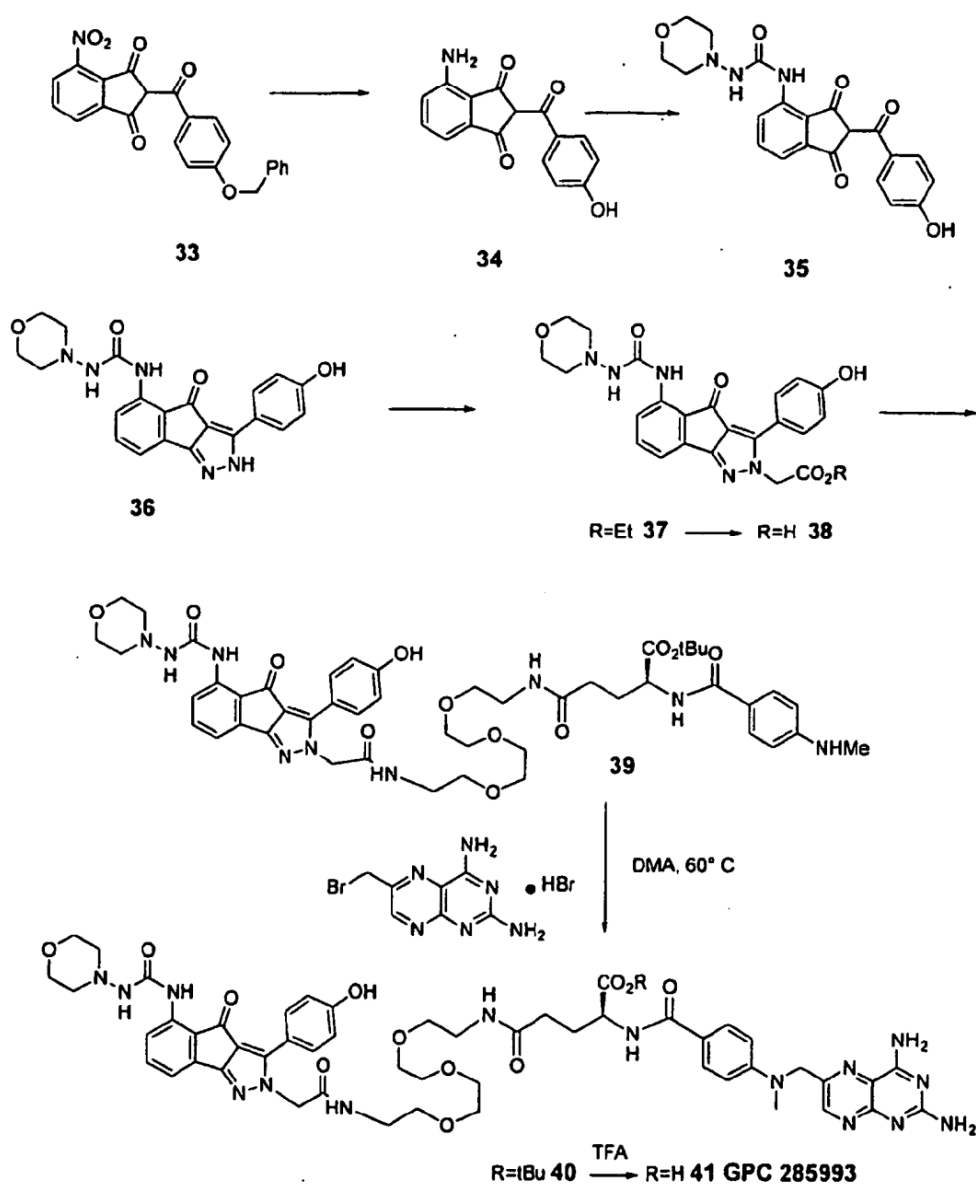


FIG. 2

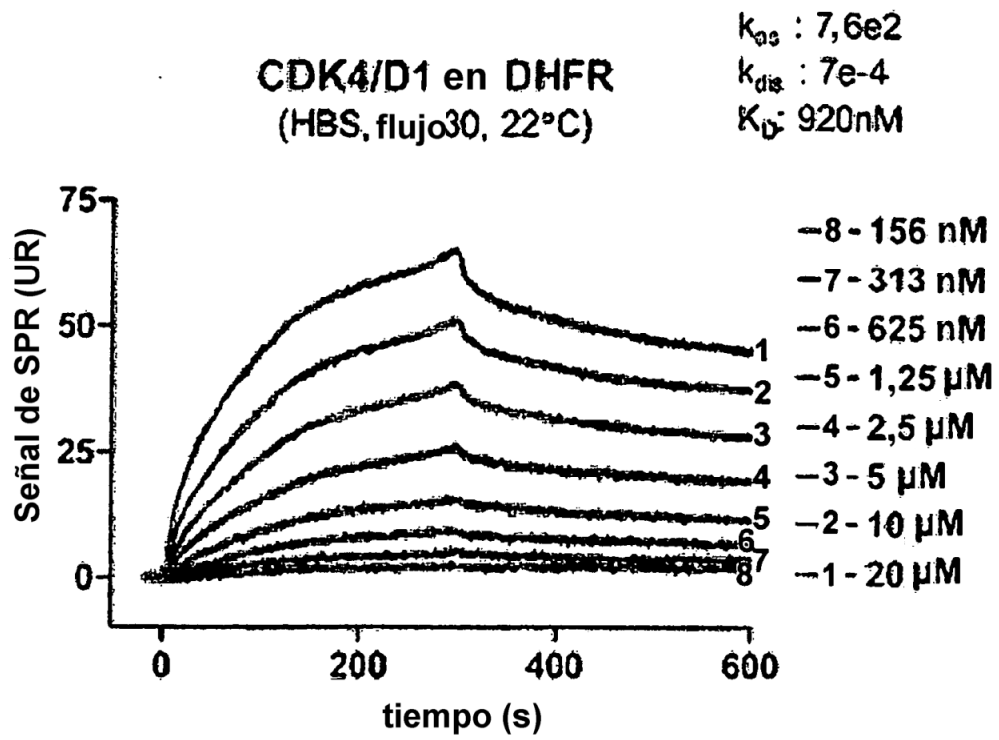
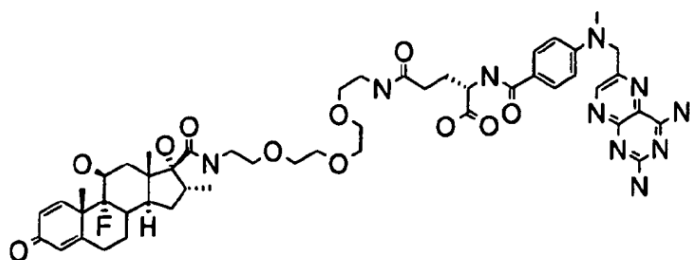


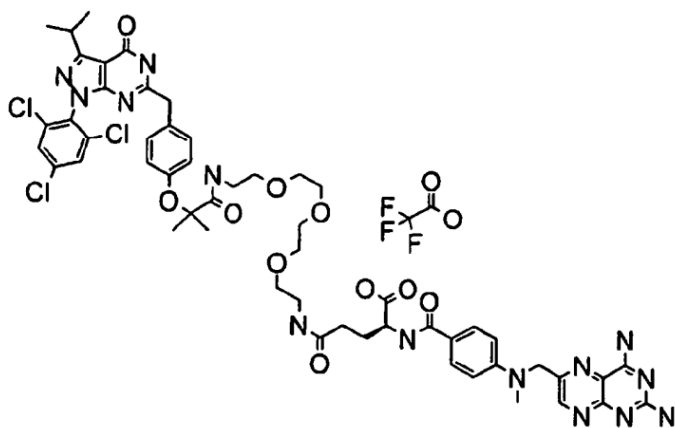
FIG. 3

Representaciones de la estructura de GPC 285937, GPC 285985 y GPC 285993

a) GPC 285937



b) GPC 285985



c) GPC 285993

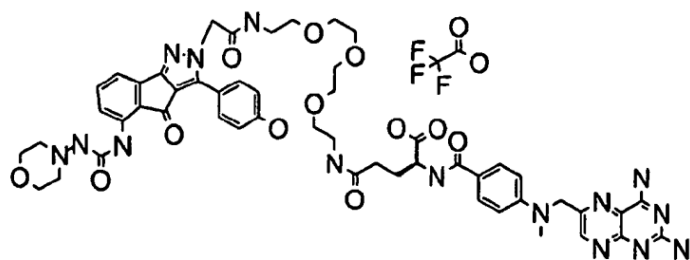


FIG. 4

a. 1 μ l de GPC

**285937 1 mM en
en DMSO**



b. 1 μ l de DMSO

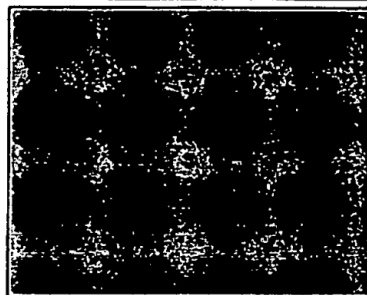


FIG. 5

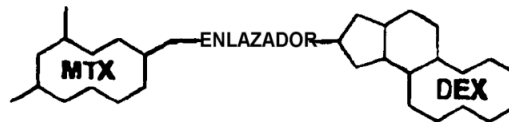
a. MTX-mdbt-Dex



b. GPC 285937



FIG. 6



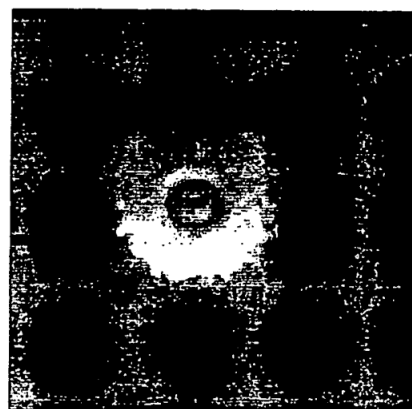
a. GPC 285937



ENLAZADOR=

3 grupos de polietilenglicol

b. Mtx-mdbt-Dex

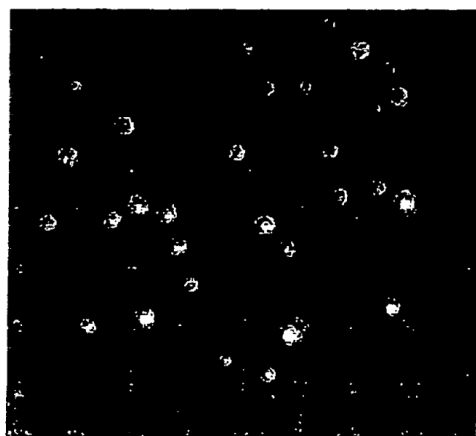


ENLAZADOR=

metadibenzotioéster

FIG. 7

Diferencia en el crecimiento de colonias de levaduras en placas de exploración



a. GPC 285937



b. Mtx-mdbt-Dex

FIG. 8

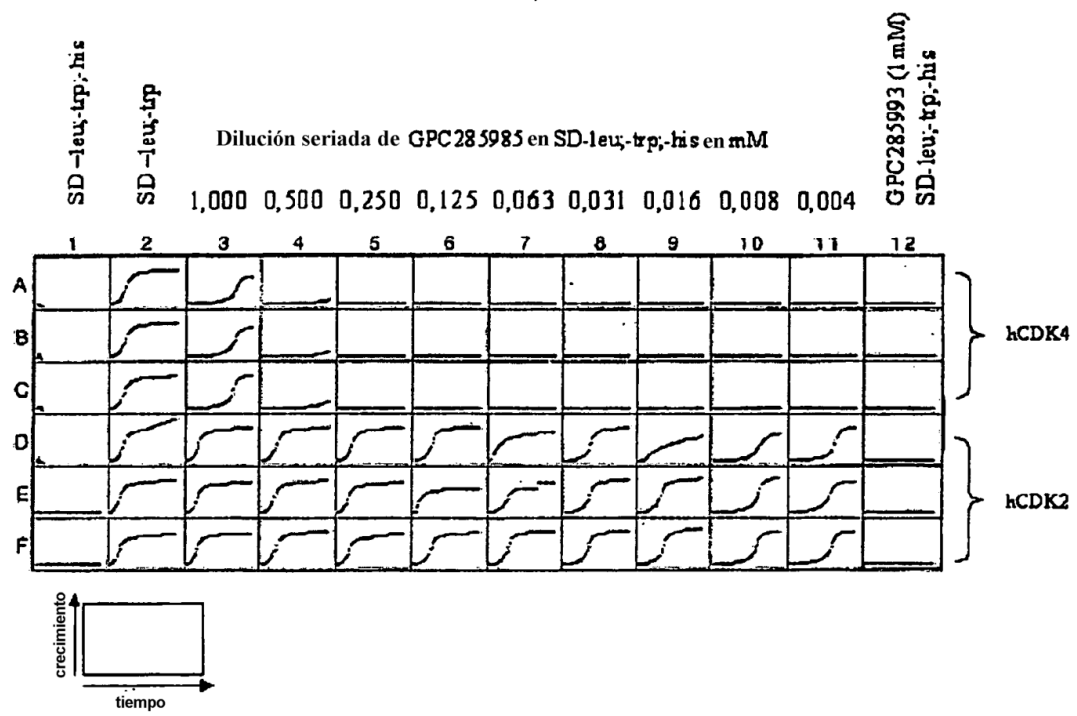


FIG. 9

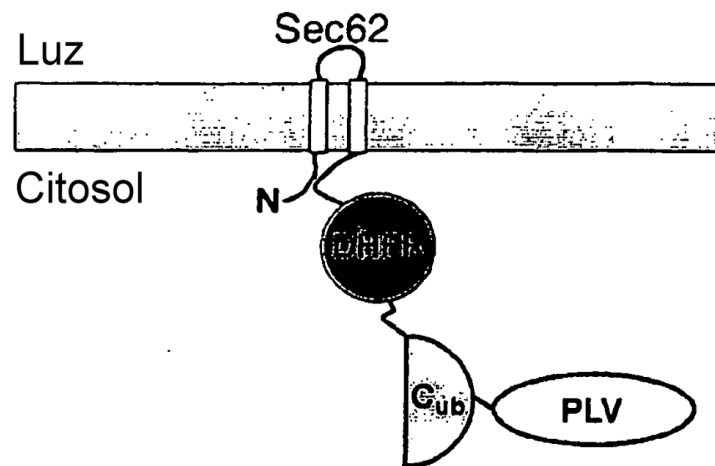
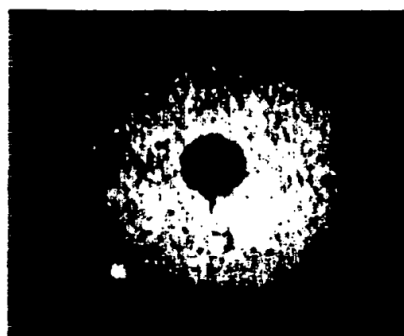


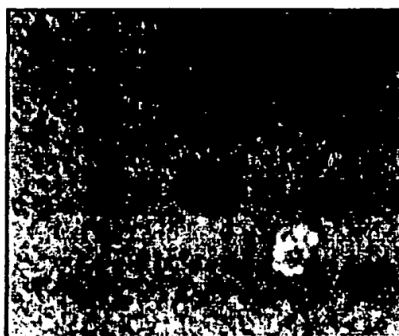
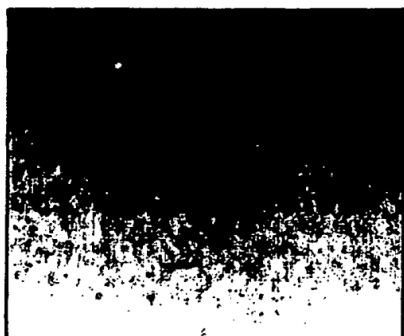
FIG. 10

Ensayo de la especificidad de CDK del compuesto GPC 285985

pBYH-DHFR/pACH-CDK2 pBYH-DHFR/pACH-CDK4



LTH



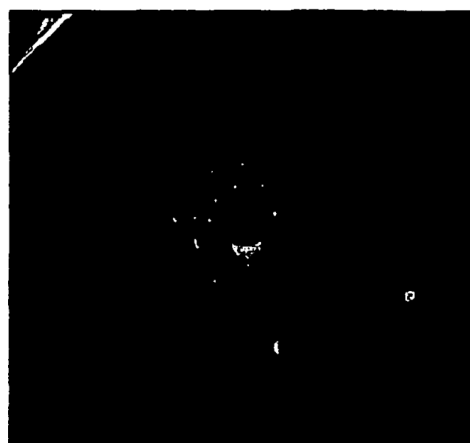
LT

FIG. 11

El compuesto GPC 285985 es específico de CDK2



CDK2



CDK4

FIG. 12

Reensayo de los clones escogidos

GeneD-Forw	Clon ID	LTH 285985	285993
GPC761	MTP1E12	☐	☒
CDK2	MTP1D6	☐	☒
GPC456	MTP1D9	☒	☒
	MTP1E7	☒	☒
GPC672	MTP1C24	☐	☒
GPC556	MTP1C14	☐	☒
GPC173	MTP1F9	☒	☒
GPC553	MTP1F7	☐	☒
GPC644	MTP1C23	☒	☒
	MTP1D16	☐	☐

FIG. 13

GPC761 se une específicamente a GPC2GPC 285985

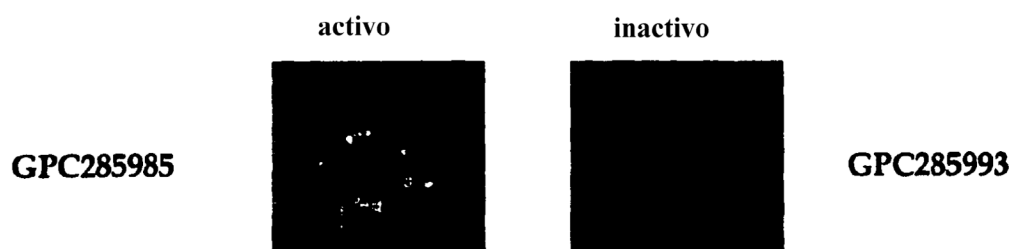


FIG. 14

Sistema de triple híbrido de mamífero

A



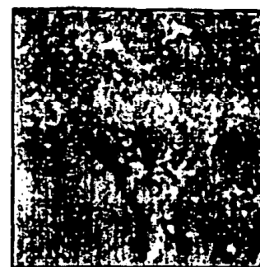
Control positivo

B



**Gal-DHFR
VP16-GR**

C



**Gal-DHFR
VP16-GR**

FIG. 15

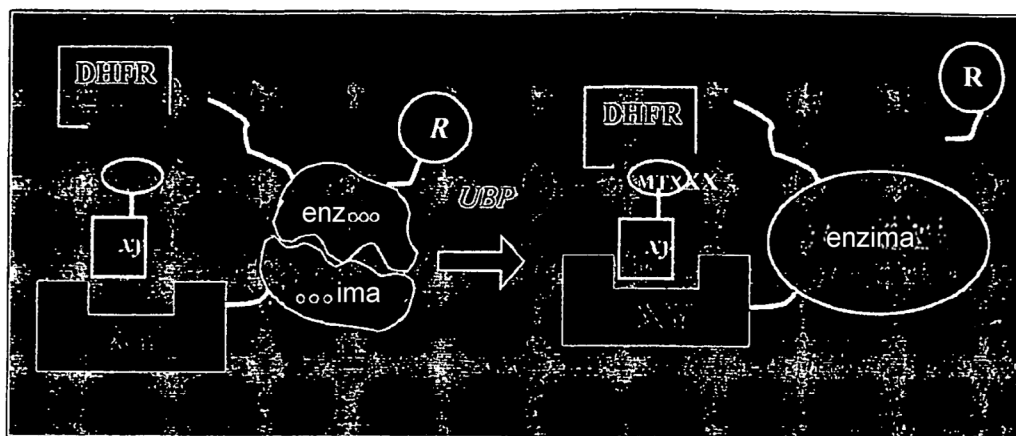
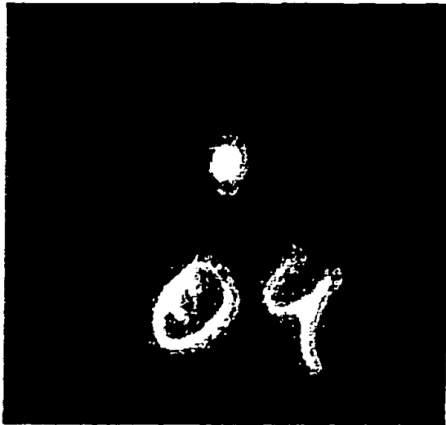


FIG. 16

a. GPC 286004

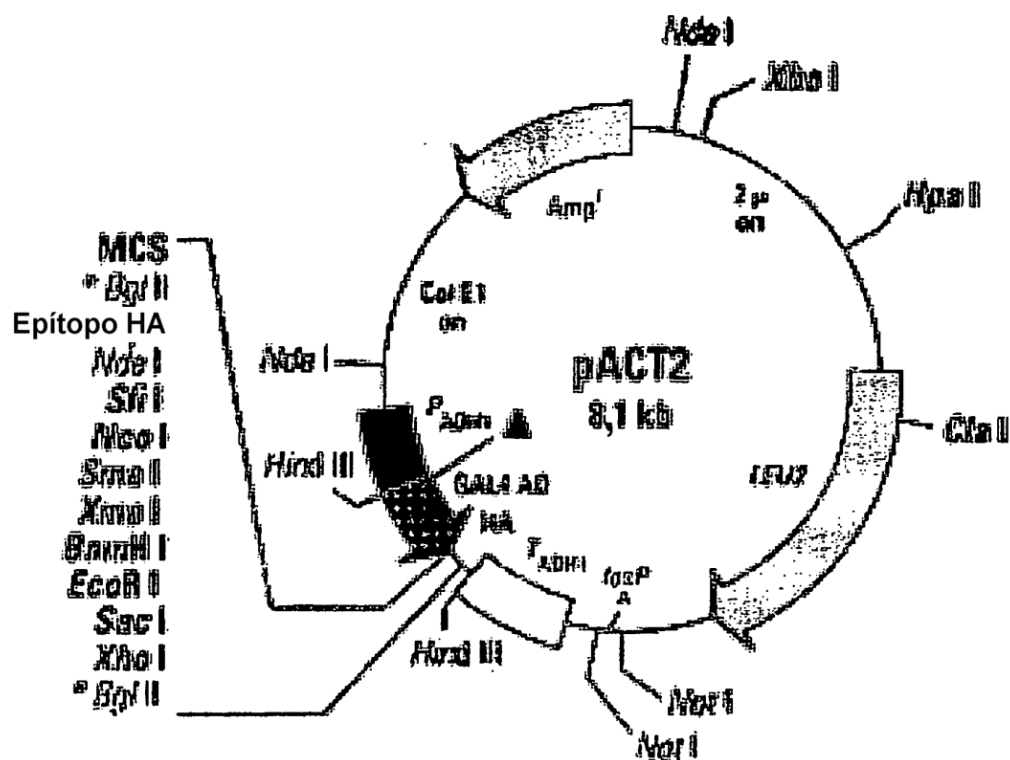


b. GPC 286026

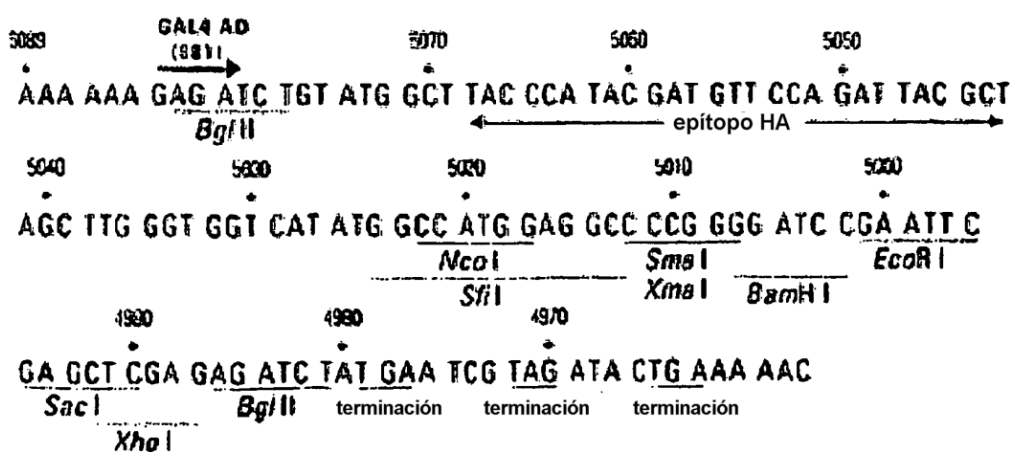


Fig. 17

a.



b



REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante únicamente es para comodidad del lector. Dicha lista no forma parte del documento de patente Europea. Aunque se ha tenido gran cuidado en la recopilación de las referencias, no se pueden excluir errores u omisiones y la EPO rechaza toda responsabilidad a este respecto.

5 Documentos de patentes citados en la descripción

- US 272932 P [0001]
- US 278233 P [0001]
- US 329437 P [0001]
- US 5468614 A [0002]
- 10 • WO 9423025 A [0004]
- WO 9530012 A [0004]
- WO 9418317 A [0006] [0100] [0210] [0229]
- US 5830462 A [0006]
- US 5869337 A [0006]
- 15 • US 6165787 A [0006]
- WO 9606097 A [0006] [0229] [0233]
- WO 9613613 A [0009] [0210] [0229]
- US 5846728 A [0010]
- WO 9741255 A [0012] [0209] [0229]
- 20 • US 5585245 A [0015] [0179] [0180] [0447]
- US 5503977 A [0015] [0179] [0180] [0447]
- US 5837832 A [0108]
- US 5888732 A [0127]
- WO 0212902 A [0180]
- 25 • US 5955280 A [0208]
- US 5965368 A [0208]
- US 5928868 A [0209]
- WO 9602561 A [0210]
- EP 0646644 A [0210]
- 30 • WO 9825947 A [0210] [0216]
- WO 9807845 A [0210]
- WO 9834120 A [0210]
- WO 9839483 A [0210]
- WO 9844350 A [0210]
- 35 • WO 0006722 A [0210]
- WO 0190188 A [0210]
- WO 9941258 A [0210]
- US 6054436 A [0252]
- US 6326155 B [0258]
- 40 • US 5223409 A [0261] [0270] [0313] [0352]

- WO 9002909 A [0263]
- WO 9209690 A [0263] [0313]
- US 5270181 A [0265]
- US 5292646 A [0265]
- 5 • WO 9402502 A [0265]
- US 5198346 A [0270] [0352]
- US 5096815 A [0270] [0352]
- US 4631211 A [0280]
- US 5440016 A [0280]
- 10 • US 5480971 A [0280]
- US 5010175 A [0287] [0288]
- US 5143854 A [0290]
- WO 9816536 A [0310]
- WO 9218619 A [0313]
- 15 • WO 9117271 A [0313]
- WO 9220791 A [0313]
- WO 9215679 A [0313]
- WO 9301288 A [0313]
- WO 9201047 A [0313]
- 20 • WO 9002809 A [0313]
- US 4683195 A [0363]
- WO 9931509 A [0417] [0419]
- WO 9920624 A [0460]
- WO 9822103 A [0460]
- 25 • WO 01060816 A [0460]
- WO 01025220 A [0460]
- WO 0137835 A [0460]
- WO 01072711 A [0460]
- US 01051620 A [0460]
- 30 • WO 01058899 A [0460]
- WO 01087846 A [0460]
- WO 01085719 A [0460]
- WO 01010859 A [0460]
- WO 01085715 A [0460]
- 35 • WO 00059901 A [0460]
- WO 00017203 A [0460]
- WO 00017202 A [0460]
- WO 9962890 A [0460]
- WO 01066540 A [0460]

- WO 9917769 A [0460]
- WO 01066539 A [0460]
- US 6187799 B [0460]
- WO 9932106 A [0460]
- 5 • WO 9932455 A [0460]
- WO 9932436 A [0460]
- WO 9917759 A [0460]

Bibliografía no relativa a patentes citada en la descripción

- Fields ; Song. Nature, 1989, vol. 340, 245-6 [0002];
- 10 • Gyuris et al. Cell, 1993, vol. 75, 791-803 [0002];
- Yang et al. Nucleic Acid Research, 1995, vol. 23,1152-1156 [0002]
- Licitra ; Liu. Proc. Natl. Acad. Sci., 1996, vol. 93,12817-21 [0011]
- Lin et al. J. Am. Chem. Soc., 2000, vol. 122, 4247-8 [0014] [0229] [0366] [0430]
- Eyckerman et al. Nature Cell Biology, 2001, vol. 3, 1114-1119 [0100]
- 15 • Clackson et al. Proc. Natl. Acad. Sci. U.S.A., 1998, vol. 95, 10437-42 [0118]
- Yang et al. J. Med. Chem., 2000, vol. 43, 1135-42 [0118]
- Stagljar et al. PNAS, 1998, vol. 95, 5187-92 [0155]
- Johnsson ; Varshavsky. Proc. Natl. Acad. Sci. U.S.A., 1994, vol. 91, 10340-10344 [0162] [0187]
- Stagljar et al. Proc. Natl. Acad. Sci. U.S.A., vol. 95, 5187-92 [0166]
- 20 • Boeke et al. Methods Enzymol, 1987, vol. 154, 164-75 [0168] [0218]
- McCusker ; Davis. Yeast, 1991, vol. 7, 607-8 [0168]
- Wigler et al. Cell, 1977, vol. 11, 223-32 [0169]
- Borrelli et al. Proc. Natl. Acad. Sci. USA, 1988, vol. 85, 7572-76 [0169]
- Lester et al. Somatic Cell Genet., 1980, vol. 6, 241-59 [0169] [0222]
- 25 • Albertini et al. Nature, 1985, vol. 316, 369-71 [0169] [0222]
- Mullen et al. Proc. Natl. Acad. Sci. USA, 1992, vol. 89, 33-37 [0169] [0222]
- Wei ; Huber. J. Biol. Chem., 1996, vol. 271, 3812-16 [0169] [0222]
- Karreman. Gene, 1998, vol. 218, 57-61 [0169] [0222]
- Strathdee et al. BioTechniques, 2000, vol. 28, 210-14 [0169]
- 30 • Bachmair et al. Science, 1986, vol. 234, 179-86 [0170]
- Varshavsky. Cell, 1992, 725-35 [0170]
- Varshavsky. Cell, 1992, vol. 69, 725-35 [0171]
- Bartel et al. EMBO J., 1990, vol. 9, 3179-89 [0174]
- Dohmen et al. Proc. Natl. Acad. Sci. USA, 1991, vol. 88, 7351-55 [0174] [0176]
- 35 • Kwon et al. Proc. Natl. Acad. Sci. USA, 1998, vol. 95, 7893-903 [0174]
- Baker ; Varshavsky. J Biol Chem, vol. 270, 12065-74 [0177]
- Balzi et al. J. Biol Chem, 1990, vol. 265, 7464-71 [0177]
- Tobias et al. J. Biol. Chem., 1991, vol. 266, 12021 [0185]
- Baker et al. J. Biol. Chem., 1992, vol. 267, 23364 [0185]

- Karreman. BioTechniques, 1988, vol. 24, 736-42 [0198]
- PLV, Stagljär et al. Proc. Natl. Acad. Sci. U.S.A., 1998, vol. 95, 5187-92 [0204]
- Yeast hybrid technologies. 2000 [0207]
- Two-Hybrid Systems: Methods and Protocols. Methods in Molecular Biology. 2001, vol. 177 [0207]
- 5 • The yeast-two-hybrid system. Oxford University Press, 1997 [0210]
- Yeast hybrid technologies. Biotechniques Press, 2000 [0210]
- Van Ostade et al. J. Interf. Cytok. Res., 2000, vol. 20, 79-87 [0210]
- Ghosh et al. J. Am. Chem. Soc., 2000, vol. 122, 5658-9 [0210]
- 10 • Fearon et al. Karyoplasmic interaction selection strategy: A general strategy to detect protein-protein interactions in mammalian cells. Proc. Natl. Acad. Sci. USA, 1992, vol. 89, 7958-7962 [0211]
- Toyn et al. Yeast, 2000, vol. 16, 553-560 [0219]
- Wigler et al. Cell, vol. 11, 223-32 [0222]
- Borrelli et al. Proc. Natl. Acad. Sci. USA, vol. 85, 7572-76 [0222]
- Strathdee et al. BioTechnique, 2000, vol. 28, 210-14 [0222]
- 15 • Rojo-Niersbach et al. mammalian cells as a basis for the selection of protein interactions. Biochem. J., 2000, vol. 348, 585-590 [0222]
- Maier et al. Automation for genome characterisation, 1997 [0225]
- Bergmann et al. J. Steroid Biochem. Molec. Biol., 1994, vol. 49, 139-52 [0229]
- Kathryn et al. Steroid Biochem. Molec. Biol., vol. 49, 139-152 [0233]
- 20 • Haussler et al. Bone, 1995, vol. 17, 33S-38S [0256]
- DeWolf ; Brett. Pharmacol Rev., 2000, vol. 52, 207-36 [0256]
- Houghten. Biotechniques, 1992, vol. 13, 412-421 [0261]
- Lam. Nature, 1991, vol. 354, 82-84 [0261]
- Fodor. Nature, 1993, vol. 364, 555-556 [0261]
- 25 • Cull et al. Proc Natl Acad Sci USA, vol. 89, 1865-1869 [0261]
- Scott ; Smith. Science, 1990, vol. 249, 386-390 [0261]
- Devlin. Science, 1990, vol. 249, 404-406 [0261]
- Cwirla et al. Proc. Natl. Acad. Sci., 1990, vol. 87, 6378-6382 [0261]
- Felici. J. Mol. Biol., 1991, vol. 222, 301-310 [0261]
- 30 • Marks et al. J. Biol. Chem., 1992, vol. 267, 16007-16010 [0263]
- Griffiths et al. EMBO J, 1993, vol. 12, 725-734 [0263]
- Clackson et al. Nature, 1991, vol. 352, 624-628 [0263]
- Barbas et al. PNAS, 1992, vol. 89, 4457-4461 [0263]
- Narang, SA. Tetrahedron, 1983, vol. 39, 3 [0270] [0352]
- 35 • Recombinant DNA, Proc 3rd Cleveland Sympos. Itakura et al. Macromolecules. Elsevier, 1981, 273-289 [0270]
- Itakura et al. Annu. Rev. Biochem., 1984, vol. 53, 323 [0270] [0352]
- Itakura et al. Science, 1984, vol. 198, 1056 [0270] [0352]
- Ike et al. Nucleic Acid Res., 1983, vol. 11, 477 [0270] [0352]
- Scott et al. Science, 1990, vol. 249, 386-390 [0270] [0352]
- 40 • Roberts et al. PNAS, 1992, vol. 89, 2429-2433 [0270] [0352]

- Devlin et al. Science, 1990, vol. 249, 404-406 [0270] [0352]
- Cwirla et al. PNAS, 1990, vol. 87, 6378-6382 [0270] [0352]
- Dower et al. Nucleic Acids Res., 1988, vol. 16, 6127-6145 [0272]
- Merrifield, R.B. J Am Chem Soc, 1963, vol. 85, 2149-2154 [0277]
- 5 • Geysen et al. PNAS, 1984, vol. 81, 3998-4002 [0278]
- Valerio et al. Int J Pept Protein Res, 1993, vol. 42, 1-9 [0278]
- Bray et al. Tetrahedron Lett, 1990, vol. 31, 5811-5814 [0278]
- Valerio et al. Anal Biochem, 1991, vol. 197, 168-177 [0278]
- Bray et al. Tetrahedron Lett, 1991, vol. 32, 6163-6166 [0278]
- 10 • Maeji et al. J Immunol Methods, 1990, vol. 134, 23-33 [0279]
- Gammon et al. J Exp Med, 1991, vol. 173, 609-617 [0279]
- Mutch et al. Pept Res, 1991, vol. 4, 132-137 [0279]
- Houghten. PNAS, 1985, vol. 82, 5131-5135 [0280]
- Houghten et al. PNAS, 1986, vol. 82, 5131-5135 [0281] [0282]
- 15 • Houghten et al. Int J Pept Protein Res, 1986, vol. 27, 673-678 [0281]
- Beck-Sickinger et al. Pept Res, 1991, vol. 4, 88-94 [0281]
- Beck-Sickinger et al. Int J Pept Protein Res, 1990, vol. 36, 522-530 [0282]
- Beck-Sickinger et al. Eur J Biochem, 1990, vol. 194, 449-456 [0282]
- Zimmerman et al. Eur J Biochem, 1991, vol. 200, 519-528 [0282]
- 20 • Gisen. Anal. Chem. Acta, 1972, vol. 58, 248-249 [0283]
- Geysen et al. Mol Immunol, 1986, vol. 23, 709-715 [0287]
- Tjoeng et al. Int J Pept Protein Res, 1990, vol. 35, 141-146 [0287]
- Birkett et al. Anal Biochem, 1991, vol. 196, 137-143 [0287]
- Petithory et al. PNAS, 1991, vol. 88, 11510-11514 [0287]
- 25 • Blankemeyer-Menge et al. Tetrahedron Lett, 1988, vol. 29, 5871-5874 [0289]
- Frank et al. Tetrahedron, 1988, vol. 44, 6031-6040 [0289]
- Eichler et al. Collect Czech Chem Commun, 1989, vol. 54, 1746-1752 [0289]
- Frank, R. Bioorg Med Chem Lett, 1993, vol. 3, 425-430 [0289]
- Eichler et al. Pept Res, 1991, vol. 4, 296-307 [0289]
- 30 • Schmidt et al. Bioorg Med Chem Lett, 1993, vol. 3, 441-446 [0289]
- Frank, R. Tetrahedron, 1992, vol. 48, 9217-9232 [0289]
- Berg et al. J Am Chem Soc, 1989, vol. 111, 8024-8026 [0289]
- Dower et al. Annu Rep Med Chem, 1991, vol. 26, 271-280 [0290]
- Fodor, S.P.A. Science, 1991, vol. 251, 767 [0290]
- 35 • Jacobs et al. Trends Biotechnol, 1994, vol. 12, 19-26 [0290]
- Gallop et al. J Med Chem, 1994, vol. 37, 1233-1251 [0291]
- Brenner et al. PNAS, 1992, vol. 89, 5381-5383 [0294]
- Needles et al. PNAS, 1993, vol. 90, 10700-10704 [0294]
- Nielsen et al. J Am Chem Soc, vol. 115, 9812-9813 [0295]

- Nielsen et al. Methods Compan Methods Enzymol, 1994, vol. 6, 361-371 [0295]
- Kerr JM et al. J Am Chem Soc, 1993, vol. 115, 2529-2531 [0297]
- Nikolaiev et al. Pept Res, 1993, vol. 6, 161-170 [0298]
- Ptek et al. Tetrahedron Lett, 1991, vol. 32, 3891-3894 [0298]
- 5 • Ohlmeyer et al. PNAS, 1993, vol. 90, 10922-10926 [0300]
- Nestler et al. J Org Chem, 1994, vol. 59, 4723-4724 [0300]
- Borchardt et al. J Am Chem Soc, 1994, vol. 116, 373-374 [0301]
- Wennemers et al. J Org Chem, 1995, vol. 60, 1108-1109 [0301]
- Yoon et al. Tetrahedron Lett, 1994, vol. 35, 8557-8560 [0301]
- 10 • Ohlmeyer et al. PNAS, 1995, vol. 92, 6027-6031 [0302]
- Bartel ; Szostak. Science, 1993, vol. 261, 1411-1418 [0306]
- Bock et al. Nature, 1992, vol. 355, 564 [0306]
- Ellington et al. Nature, 1992, vol. 355, 850-852 [0306]
- Oliphant et al. Mol Cell Biol, 1989, vol. 9, 2944-2949 [0306]
- 15 • Tuerk et al. Science, 1990, vol. 249, 505-510 [0307]
- Current Protocols in Molecular Biology. John Wiley & Sons, 1989 [0309] [0320]
- Molecular Cloning:ALaboratory Manual. Cold Spring Harbor press, 1989 [0309]
- Molecular Cloning:ALaboratory Manual. Cold Spring Harbor press, 1989 [0320]
- Goeddel. Gene Expression Technology: Methods in Enzymology. Academic Press, 1990, vol. 185 [0325]
- 20 • Broach et al. Experimental Manipulation ofGeneExpression. Academic Press, 1983, 83 [0327]
- Molecular Cloning A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989 [0329]
- Ben-Bassat et al. J. Bacteriol., 1987, vol. 169, 751-757 [0331]
- Miller et al. PNAS, 1987, vol. 84, 2718-1722 [0331]
- Current Protocols in Molecular Biology. John Wiley & Sons, 1991 [0342]
- 25 • Biochemistry. WFT Freeman and Co, 1981 [0349]
- Recombinant DNA. Itakura et al. Proc 3d Cleveland Sympos. Macromolecules. Elsevier, 1981, 273-289 [0352]
- Freidinger et al. hi Peptides: Chemistry and Biology. ESCOM Publisher, 1988 [0354]
- Huffman et al. Peptides: Chemistry and Biology. ESCOM Publisher, 1988 [0354]
- Garvey et al. Peptides: Chemistry and Biology, 1988 [0354]
- 30 • Ewenson et al. J Med Chem, 1986, vol. 29, 295 [0354]
- Ewenson et al. Peptides: Structure and Function (Proceedings of the American Peptide Symposium). Pierce Chemical Co, 1985 [0354]
- Nagai et al. Tetrahedron Lett, 1985, vol. 26, 647 [0354]
- Sato et al. 3 Chem Soc Perkin Trans, 1986, vol. 1, 1231 [0354]
- 35 • Gordon et al. Biochem Biophys Res Commun, 1985, vol. 26, 419 [0354]
- Dann et al. Biochem Biophys Res Commun, 1986, vol. 134, 71 [0354]
- Molecular Cloning:A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989 [0363]
- DNA Cloning. 1985, vol. I and II [0363]
- Oligonucleotide Synthesis. 1984 [0363]

- Nucleic Acid Hybridization. 1984 [0363]
- Transcription And Translation. 1984 [0363]
- B. Perbal. A Practical Guide To Molecular Cloning, 1984 [0363]
- Methods In Enzymology. Academic Press [0363]
- 5 • Methods In Enzymology. vol. 154 and [0363]
- Immunochemical Methods In Cell And Molecular Biology. Academic Press, 1987 [0363]
- Taghavi ; Pfeleiderer. Tetrahedron Lett., 1997, vol. 38, 6835-36 [0376]
- Taylor ; Portnoy. J. Org. Chem., 1973, vol. 38, 806 [0376]
- Sarcevic et al. J. Biol. Chem., 1997, vol. 272, 33327-37 [0410]
- 10 • Konstantinidis et al. J. Biol. Chem., 1998, vol. 273, 26506-15 [0412]
- Chakraborti et al. J. Biol. Chem., 1991, vol. 266, 22075-22078 [0419]
- Burke. Methods in yeast genetics: A Cold Spring Harbor Laboratory course manual. Cold Spring Harbor Laboratory Press, 2000 [0425] [0428] [0455]
- The Yeast Two-Hybrid System (Advances in Molecular Biology). Oxford University Press, 316-332 [0447]
- 15 • Dohmen et al. JBC, 1995, vol. 270, 18099-109 [0448] [0455]
- Stagljar et al. Proc. Natl. Acad. Sci. U.S.A., 1998, vol. 95, 5187-92 [0450] [0457]
- Reichel et al. PNAS, 1996, vol. 93, 5888-93 [0452]
- Laser et al. PNAS, 2000, vol. 97, 13732-7 [0453] [0454]
- Science, 1997, vol. 278 (5336), 286-90 [0460]
- 20 • J. Med. Chem., 1998, vol. 41, 3276 [0460]
- Bioorg. Med. Chem. Lett., 2002, vol. 12, 221-224 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 1401-1405 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10, 223 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 9-12 [0460]
- 25 • Bioorg. Med. Chem. Lett., 2000, vol. 10, 2051-2054 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10, 567-569 [0460]
- Bioorg. Med. Chem. Lett., 2001, 1401-1405 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 1401-1405 [0460]
- J. Med. Chem., 2000, vol. 43, 4089-4108 [0460]
- 30 • Bioorg. Med. Chem. Lett., 2001, vol. 11, 1922-1914 [0460]
- J. Med. Chem., 2001, vol. 44, 3965-3977 [0460]
- J. Med. Chem., 2001, vol. 44, 3965 [0460]
- J. Med. Chem. [0460]
- Chem. Biol., 1999, vol. 6, 559 [0460]
- 35 • J. Biol. Chem., 1999, vol. 276, 17420 [0460]
- Nature, 1998, vol. 389, 990 [0460]
- J. Med. Chem., 2001, vol. 44, 4615 [0460]
- J. Med. Chem., 2001, vol. 44, 4628 [0460]
- Proc. Am. Assoc. Can Res., 1999 [0460]

- Proc. Am. Assoc. Can. Res., 1999, vol. 90, 69 [0460]
- J. Med. Chem., 1995, vol. 38, 3780 [0460]
- J. Med. Chem., 1998, vol. 41, 4365 [0460]
- J. Med. Chem., 1997, vol. 40, 1820 [0460]
- 5 • J. Med. Chem., 1998, vol. 41, 4196 [0460]
- Ann. N. Y. Acad. Sci., 1993, vol. 696, 149 [0460]
- J. Med. Chem., 1996, vol. 39, 5215 [0460]
- Proc. Am. Assoc. Can. Res., 1998 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10, 575-579 [0460]
- 10 • J. Med. Chem., 2001, vol. 44, 3417 [0460]
- J. Biol. Chem., 1997, vol. 272, 29207 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10m, 945-949 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10, 945-949 [0460]
- Bioorg. Med. Chem. Lett., 1996, vol. 6, 1759-1764 [0460]
- 15 • Bioorg. Med. Chem. Lett., 2000, vol. 10, 2047-2050 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10, 2047-2050 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10, 2051-2054 [0460]
- J. Biol. Chem., 1991, vol. 266, 15771 [0460]
- J. Med. Chem., 1991, vol. 34, 1896 [0460]
- 20 • Bioorg. Med. Chem. Lett., 2001, vol. 11, 1123-1126 [0460]
- J. Med. Chem., 1994, vol. 37, 2627 [0460]
- J. Med. Chem., 1999, vol. 41, 2588 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 853-856 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 693-696 [0460]
- 25 • Bioorg. Med. Chem. Lett., 1998, vol. 8, 3111-3116 [0460]
- Chem. Lett., 2001, vol. 11, 853-856 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 849-852 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 840-852 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 1157-1160 [0460]
- 30 • J. Med. Chem., 2001, vol. 44, 2133-2138 [0460]
- J. Am. Chem. Soc., 2001, vol. 123, 11586-11593 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10, 2167-2170 [0460]
- Atherosclerosis, 2002, vol. 160 (1), 123-132 [0460]
- Biochem. Biophys. Res. Commun., 1987, vol. 147, 322 [0460]
- 35 • Bioorg. Med. Chem. Lett., 2000, vol. 10, 223-226 [0460]
- Bioorg. Med. Chem. Lett., 2000, vol. 10, 461-464 [0460]
- Bioorg. Med. Chem. Lett., 2001 [0460]
- Bioorg. Med. Chem. Lett., 2001, vol. 11, 1911-1914 [0460]