

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 478 622**

51 Int. Cl.:

**C07D 319/20** (2006.01)

**C07C 67/26** (2006.01)

**C07C 243/38** (2006.01)

**A61K 31/357** (2006.01)

**A61P 35/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.02.2004 E 04709841 (3)**

97 Fecha y número de publicación de la concesión europea: **09.04.2014 EP 1603860**

54 Título: **Ligandos de diacilhidrazina para modular la expresión de genes exógenos en sistemas de mamífero a través de un complejo de receptor de ecdisona**

30 Prioridad:

**10.02.2003 US 446233 P**

**09.02.2004 US 775883**

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

**22.07.2014**

73 Titular/es:

**INTREXON CORPORATION (100.0%)**

**1750 Kraft Drive, Suite 1400**

**Blacksburg, VA 24060 , US**

72 Inventor/es:

**HORMANN, ROBERT, EUGENE;**

**TICE, COLIN, M.;**

**CHORTYK, ORESTES;**

**SMITH, HOWARD y**

**METEYER, THOMAS**

74 Agente/Representante:

**LINAGE GONZÁLEZ, Rafael**

**ES 2 478 622 T3**

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

## DESCRIPCIÓN

Ligandos de diacilhidrazina para modular la expresión de genes exógenos en sistemas de mamífero a través de un complejo de receptor de ecdisona

5 Esta solicitud reivindica prioridad sobre la solicitud provisional estadounidense n.º 60/446.233 presentada el 10/2/2003.

**Campo de la invención**

10 Esta invención se refiere al campo de la biotecnología o ingeniería genética. Específicamente, esta invención se refiere al campo de la expresión génica. Más específicamente, esta invención se refiere a ligandos no esteroideos tal como se define en la reivindicación 1 para receptores nucleares naturales y mutados y a su uso en la modulación de la expresión de un gen dentro de una célula huésped.

**Antecedentes de la invención**

15 En el presente documento se citan diversas publicaciones. Sin embargo, la mención de cualquier referencia en el presente documento no debe considerarse un reconocimiento de que tal referencia está disponible como "técnica anterior" a la presente solicitud.

20 En el campo de la ingeniería genética, el control preciso de la expresión génica es una herramienta valiosa para estudiar, manipular y controlar el desarrollo y otros procesos fisiológicos. La expresión génica es un proceso biológico complejo que implica varias interacciones proteína-proteína específicas. Con el fin de desencadenar la expresión génica, de manera que produzca el ARN necesario como primera etapa en la síntesis de proteínas, debe llevarse un activador de la transcripción a las proximidades de un promotor que controla la transcripción génica. Normalmente, el propio activador de la transcripción está asociado con una proteína que tiene al menos un dominio de unión a ADN que se une a sitios de unión del ADN presentes en las regiones promotoras de los genes. Por tanto, para que se produzca la expresión génica, una proteína que comprende un dominio de unión a ADN y un dominio de transactivación ubicado a una distancia apropiada del dominio de unión a ADN debe llevarse a la posición correcta en la región promotora del gen.

25 El enfoque transgénico tradicional utiliza un promotor específico del tipo celular para dirigir la expresión de un transgén diseñado. En primer lugar, se incorpora un constructo de ADN que contiene el transgén en un genoma huésped. Cuando se desencadena por un activador de la transcripción, se produce la expresión del transgén en un tipo celular dado.

30 Otro medio para regular la expresión de genes foráneos en células es a través de promotores inducibles. Los ejemplos del uso de tales promotores inducibles incluyen el promotor PR1-a, sistemas de represor-operador de procariontes, sistemas de inmunosupresor-inmunofilina y sistemas de activación de la transcripción de eucariotas superiores tales como sistemas de receptores de hormonas esteroideas y se describen a continuación.

35 El promotor PR1-a del tabaco se induce durante la respuesta de resistencia sistémica adquirida tras el ataque de patógenos. El uso de PR1-a puede estar limitado porque a menudo responde a materiales endógenos y factores externos tales como patógenos, radiación UV-B y contaminantes. Se han descrito sistemas de regulación génica basados en promotores inducidos por choque térmico, interferón y metales pesados (Wurn *et al.*, 1986, Proc. Natl. Acad. Sci. USA 83:5414-5418; Arnheiter *et al.*, 1990 Cell 62:51-61; Filmus *et al.*, 1992 Nucleic Acids Research 20:27550-27560). Sin embargo, estos sistemas tienen limitaciones debido a su efecto sobre la expresión de genes no diana. Estos sistemas también son rezumantes.

40 Los sistemas de represor-operador de procariontes utilizan proteínas represoras bacterianas y secuencias de ADN de operador único a las que se unen. Se han usado los sistemas de represor-operador tanto de tetraciclina ("Tet") como de lactosa ("Lac") de la bacteria *Escherichia coli* en plantas y animales para controlar la expresión génica. En el sistema Tet, la tetraciclina se une a la proteína represora TetR, dando como resultado un cambio conformacional que libera la proteína represora del operador que como resultado permite que se produzca la transcripción. En el sistema Lac, se activa un operón lac en respuesta a la presencia de lactosa, o análogos sintéticos tales como isopropil-b-D-tiogalactósido. Desgraciadamente, el uso de tales sistemas está limitado por la química inestable de los ligandos, es decir tetraciclina y lactosa, por su toxicidad, por su presencia natural o por los niveles relativamente altos requeridos para la inducción o represión. Por motivos similares, está limitada la utilidad de tales sistemas en animales.

45 Las moléculas inmunosupresoras tales como FK506, rapamicina y ciclosporina A pueden unirse a inmunofilinas FKBP12, ciclofilina, etc. Usando esta información, se ha ideado una estrategia general para reunir dos proteínas cualesquiera colocando simplemente FK506 en cada una de las dos proteínas o colocando FK506 en una y ciclosporina A en otra. Entonces puede usarse un homodímero sintético de FK506 (FK1012) o un compuesto que resulta de la fusión de FK506-ciclosporina (FKCsA) para inducir la dimerización de estas moléculas (Spencer *et al.*,

1993, Science 262:1019-24; Belshaw *et al.*, 1996 Proc Natl Acad Sci U S A 93:4604-7). Se usaron el dominio de unión a ADN de Gal4 fusionado a FKBP12 y el dominio activador de VP16 fusionado a ciclofilina y el compuesto FKCsA para mostrar heterodimerización y activación de un gen indicador bajo el control de un promotor que contiene sitios de unión a Gal4. Desgraciadamente, este sistema incluye inmunosupresores que pueden tener efectos secundarios no deseados y por tanto, se limita su uso a diversas aplicaciones de cambio génico en mamíferos.

También se han empleado sistemas de activación de la transcripción de eucariotas superiores tales como sistemas de receptores de hormonas esteroideas. Los receptores de hormonas esteroideas son miembros de la superfamilia de receptores nucleares y se encuentran en células de vertebrados e invertebrados. Desgraciadamente, el uso de compuestos esteroideos que activan los receptores para la regulación de la expresión génica, particularmente en plantas y mamíferos, está limitado debido a su implicación en muchas otras rutas biológicas naturales en tales organismos. Con el fin de superar tales dificultades, se ha desarrollado un sistema alternativo usando receptores de ecdisona (EcR) de insectos

El crecimiento, la muda y el desarrollo en los insectos están regulados por la hormona esteroidea ecdisona (hormona de la muda) y las hormonas juveniles (Dhadialla, *et al.*, 1998, Annu. Rev. Entomol. 43: 545-569). La diana molecular para ecdisona en insectos consiste en al menos el receptor de ecdisona (EcR) y la proteína ultraespiráculo (USP). El EcR es un miembro de la superfamilia de receptores de esteroides nucleares que se caracteriza por dominios de unión a ADN y a ligandos distintivos y un dominio de activación (Koelle *et al.* 1991, Cell, 67: 59-77). Los receptores EcR responden a varios compuestos esteroideos tales como ponasterona A y muristerona A. Recientemente, se han descrito compuestos no esteroideos con actividad agonista ecdiesteroide, incluyendo los insecticidas disponibles comercialmente tebufenozida y metoxifenozida que se comercializan en todo el mundo por Rohm and Haas Company (véase la solicitud de patente internacional n.º PCT/EP96/00686 y la patente estadounidense 5.530.028). Ambos análogos tienen perfiles de seguridad excepcionales para otros organismos.

El receptor de ecdisona (EcR) de insectos heterodimeriza con ultraespiráculo (USP), el homólogo de insectos de RXR de mamíferos, y se une a elementos de respuesta del receptor de ecdisona y ecdiesteroides y activa la transcripción de genes de respuesta a ecdisona. Los complejos EcR/USP/ligando desempeñan importantes papeles durante el desarrollo y la reproducción de insectos. El EcR es un miembro de la superfamilia de receptores de hormonas esteroideas y tiene cinco dominios modulares, dominios A/B (transactivación), C (unión a ADN, heterodimerización), D (bisagra, heterodimerización), E (unión a ligando, heterodimerización y transactivación) y F (transactivación). Algunos de estos dominios tales como A/B, C y E conservan su función cuando se fusionan a otras proteínas.

Los sistemas de expresión génica inducible regulados estrechamente, o "interruptores génicos", son útiles en diversas aplicaciones, tales como terapia génica, producción a gran escala de proteínas en células, ensayos de selección de alto rendimiento basados en células, genómica funcional y regulación de caracteres en plantas y animales transgénicos.

La primera versión de un interruptor génico basado en EcR utilizó el EcR de *Drosophila melanogaster* (DmEcR) y RXR de *Mus musculus* (MmRXR) y mostró que estos receptores en presencia del esteroide ponasterona A, transactivan genes indicadores en líneas celulares de mamífero y ratones transgénicos (Christopherson K. S., Mark M.R., Baja J. V., Godowski P. J. 1992, Proc. Natl. Acad. Sci. U.S.A. 89: 6314-6318; No D., Yao T.P., Evans R. M., 1996, Proc. Natl. Acad. Sci. U.S.A. 93: 3346-3351). Más tarde, Suhr *et al.* 1998, Proc. Natl. Acad. Sci. 95:7999-8004 mostraron que el agonista no esteroideo de ecdisona, tebufenozida, inducía un alto nivel de transactivación de genes indicadores en células de mamífero a través de EcR de *Bombyx mori* (BmEcR) en ausencia de pareja exógena del heterodímero.

Las solicitudes de patente internacional n.ºs PCT/US97/05330 (documento WO 97/38117) y PCT/US99/08381 (documento WO 99/5815) dan a conocer métodos para modular la expresión de un gen exógeno en que se activa un constructo de ADN que comprende el gen exógeno y un elemento de respuesta a ecdisona por un segundo constructo de ADN que comprende un receptor de ecdisona que, en presencia de un ligando para el mismo, y opcionalmente en presencia de un receptor que puede actuar como pareja silenciosa, se une al elemento de respuesta a ecdisona para inducir la expresión génica. El receptor de ecdisona de elección se aisló de *Drosophila melanogaster*. Normalmente, tales sistemas requieren la presencia de la pareja silenciosa, preferiblemente el receptor X retinoide (RXR), con el fin de proporcionar activación óptima. En células de mamífero, el receptor de ecdisona (EcR) de insectos heterodimeriza con el receptor X retinoide (RXR) y regula la expresión de genes diana de manera dependiente de ligando. La solicitud de patente internacional n.º PCT/US98/14215 (documento WO 99/02683) da a conocer que el receptor de ecdisona aislado del gusano de seda *Bombyx mori* es funcional en sistemas de mamíferos sin necesidad de una pareja exógena del dímero.

La patente estadounidense n.º 6.265.173 B1 da a conocer que diversos miembros de la superfamilia de receptores esteroideos/tiroideos pueden combinarse con el receptor de ultraespiráculo (USP) de *Drosophila melanogaster* o fragmentos del mismo que comprenden al menos el dominio de dimerización de USP para su uso en un sistema de

expresión génica. La patente estadounidense n.º 5.880.333 da a conocer un sistema de heterodímero de EcR y ultraespiráculo (USP) de *Drosophila melanogaster* usado en plantas en que el dominio de transactivación y el dominio de unión a ADN están colocados en dos proteínas híbridas diferentes. Desgraciadamente, estos sistemas basados en USP son constitutivos en células de animales y por lo tanto, no son eficaces para regular la expresión de genes indicadores.

En cada uno de estos casos, el dominio de transactivación y el dominio de unión a ADN (o bien como EcR nativo tal como en la solicitud de patente internacional n.º PCT/US98/14215 o bien como EcR modificado, tal como en la solicitud de patente internacional n.º PCT/US97/05330) se incorporaron en una única molécula y los otros componentes heterodiméricos, o bien USP o bien RXR, se usaron en su estado nativo. Los inconvenientes de los sistemas de regulación génica basados en EcR descritos anteriormente incluyen una actividad de fondo considerable en ausencia de ligandos y la no aplicabilidad de estos sistemas para su uso tanto en plantas como en animales (véase la patente estadounidense n.º 5.880.333). Por lo tanto, existe una necesidad en la técnica de sistemas basados en EcR mejorados para modular con precisión la expresión de genes exógenos tanto en plantas como en animales. Tales sistemas mejorados serían útiles para aplicaciones tales como terapia génica, producción a gran escala de proteína y anticuerpos, ensayos de selección de alto rendimiento basados en células, genómica funcional y regulación de caracteres en animales transgénicos. Para determinadas aplicaciones tales como la terapia génica, puede ser deseable tener un sistema de expresión génica inducible que responda bien a ligandos no esteroideos sintéticos y que simultáneamente sea insensible a los esteroides naturales. Por tanto, sistemas mejorados que sean sencillos, compactos y dependientes de ligandos que sean relativamente económicos, fácilmente disponibles y de baja toxicidad para el huésped demostrarían ser útiles para regular sistemas biológicos.

Recientemente, se ha demostrado que un sistema de expresión génica inducible basado en el receptor de ecdisona en que los dominios de transactivación y de unión a ADN se separan uno de otro al colocarlos en dos proteínas diferentes, da como resultado una actividad de fondo muy reducida en ausencia de un ligando y una actividad significativamente aumentada con respecto al fondo en presencia de un ligando (solicitud de patente en trámite PCT/US01/09050). Este sistema de dos híbridos es un sistema de modulación de la expresión génica inducible significativamente mejorado en comparación con los dos sistemas dados a conocer en las solicitudes PCT/US97/05330 y PCT/US98/14215. El sistema de dos híbridos se aprovecha de la capacidad de un par de proteínas que interaccionan para llevar el dominio de activación de la transcripción a una posición más favorable con respecto al dominio de unión a ADN, de manera que cuando el dominio de unión a ADN se une al sitio de unión del ADN en el gen, el dominio de transactivación activa más eficazmente el promotor (véase, por ejemplo, la patente estadounidense n.º 5.283.173). Brevemente, el sistema de expresión génica de dos híbridos comprende dos casetes de expresión génica: codificando el primero para un dominio de unión a ADN fusionado a un polipéptido de receptor nuclear, y codificando el segundo para un dominio de transactivación fusionado a un polipéptido de receptor nuclear diferente. En presencia de ligando, la interacción del primer polipéptido con el segundo polipéptido ancla eficazmente el dominio de unión a ADN al dominio de transactivación. Puesto que los dominios de unión a ADN y de transactivación residen en dos moléculas diferentes, la actividad de fondo en ausencia de ligando está muy reducida.

Un sistema de dos híbridos también proporciona sensibilidad mejorada a ligandos no esteroideos, por ejemplo, diacilhidrazinas, cuando se compara con ligandos esteroideos, por ejemplo, ponasterona A ("Pon A") o muristerona A ("MurA"). Es decir, cuando se compara con ligandos esteroideos, los ligandos no esteroideos proporcionan una actividad superior a una concentración inferior. Además, puesto que la transactivación basada en interruptores génicos de EcR con frecuencia es dependiente de la línea celular, es más fácil adaptar los sistemas de interruptor para obtener la máxima capacidad de transactivación para cada aplicación. Además, el sistema de dos híbridos evita algunos efectos secundarios debidos a la sobreexpresión de RXR que se producen a menudo cuando se usa RXR no modificado como pareja de interruptor. En un sistema de dos híbridos preferido, los dominios nativos de unión a ADN y de transactivación de EcR o RXR se eliminan y como resultado, estas moléculas híbridas tienen menos posibilidad de interaccionar con otros receptores de hormonas esteroideas presentes en la célula dando como resultado efectos secundarios reducidos.

El documento WO 01/70816 da a conocer un sistema de expresión génica inducible y métodos de modulación de la expresión génica en una célula huésped.

El documento EP0965644 da a conocer un método para modular la expresión génica exógena en el que se pone en contacto un complejo de receptor de ecdisona que comprende: un dominio de unión a ADN; un dominio de unión a ligando; un dominio de transactivación; y un ligando, con un constructo de ADN que comprende: el gen exógeno y un elemento de respuesta; en el que el gen exógeno está bajo el control del elemento de respuesta y la unión del dominio de unión a ADN al elemento de respuesta en presencia del ligando da como resultado la activación o supresión del gen.

Con la mejora en los sistemas de regulación génica basados en el receptor de ecdisona hay un aumento en su uso en diversas aplicaciones, dando como resultado una demanda aumentada de ligandos con actividad superior a los que existen actualmente. La patente estadounidense 6.258.603 B1 (y patentes citadas en la misma) dan a conocer ligandos de dibenzoilhidrazina, sin embargo, existe la necesidad de ligandos adicionales con diferentes estructuras y

propiedades fisicoquímicas. Se han descubierto ligandos de diacilhidrazina novedosos que no se ha descrito ni mostrado anteriormente que tengan capacidad para modular la expresión de transgenes.

### Sumario de la invención

5

La presente invención proporciona un compuesto seleccionado del grupo que consiste en:

- a) N-terc-butil-N'-(3-hidroximetil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- 10 b) N-terc-butil-N'-[3-(terc-butil-dimetil-silaniloximetil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil]-hidrazida del ácido 3,5-dimetil-benzoico,
- 15 c) ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico,
- d) éster metílico del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico,
- 20 e) N-terc-butil-N'-(3-semicarbazidometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- f) éster 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetílico del ácido fenil-carbámico,
- 25 g) N'-[3-(2-amino-etil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil]-N-terc-butil-hidrazida del ácido 3,5-dimetil-benzoico,
- h) éster pentafluorofenílico del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico,
- 30 i) metilamida del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico,
- 35 j) N-terc-butil-N'-(3-formil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- k) éster 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetílico del ácido tolueno-4-sulfónico,
- 40 l) N-terc-butil-N'-[3-(hidroxiimino-metil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil]-hidrazida del ácido 3,5-dimetil-benzoico,
- m) N-terc-butil-N'-(3-cianometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- 45 n) N-terc-butil-N'-(5-metil-3-metilsulfanilmetil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- 50 o) N-terc-butil-N'-(3-metanosulfonilmetil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, y
- p) N-terc-butil-N'-(3-fluorometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico.

55 En otro aspecto, la presente invención proporciona una composición farmacéutica que comprende un compuesto de la presente invención y un portador farmacéuticamente aceptable.

60 En un aspecto adicional, la presente invención proporciona un compuesto de la presente invención para su uso en la modulación de la expresión de un gen diana en una célula huésped aislada, en el que se pone en contacto dicha célula huésped con dicho compuesto, y en el que dicha célula huésped incluye un primer casete de expresión génica que comprende un primer polinucleótido que codifica para un primer polipéptido que comprende:

(i) un dominio de transactivación;

(ii) un dominio de unión a ADN; y

65

(iii) un dominio de unión a ligando del receptor nuclear de grupo H; y

un segundo casete de expresión génica que comprende:

- (i) un elemento de respuesta que puede unirse a dicho dominio de unión a ADN;
- (ii) un promotor que se activa mediante dicho dominio de transactivación; y
- (iii) dicho gen diana.

La presente invención proporciona en otro aspecto un compuesto de la presente invención para su uso en el tratamiento de una enfermedad en un sujeto en el que dicho compuesto modula la expresión de uno o más genes exógenos que codifican para proteínas biológicamente activas.

La presente invención proporciona, en un aspecto adicional, un compuesto de la presente invención para su uso en la regulación de la expresión génica endógena o heteróloga en un sujeto transgénico no animal o no humano que comprende poner en contacto dicho compuesto con un complejo de receptor de ecdisona dentro de las células de dicho sujeto, en el que dichas células contienen además una secuencia de unión a ADN para dicho complejo de receptor de ecdisona cuando está en combinación con dicho compuesto y en el que la formación de un complejo de receptor de ecdisona-compuesto-secuencia de unión a ADN induce la expresión de dicho gen.

En un aspecto adicional, la presente invención proporciona un compuesto de la presente invención para su uso en la modulación de la expresión de un gen en una célula huésped aislada que comprende las etapas de:

a) introducir en dicha célula huésped un sistema de modulación de la expresión génica que comprende:

i) un primer casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que codifica para un primer polipéptido híbrido que comprende:

(a) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y

(b) un dominio de unión a ligando del receptor de ecdisona;

ii) un segundo casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que codifica para un segundo polipéptido híbrido que comprende:

(a) un dominio de transactivación; y

(b) un dominio de unión a ligando del receptor X retinoide quimérico; y

iii) un tercer casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que comprende:

(a) un elemento de respuesta reconocido por dicho dominio de unión a ADN de dicho primer polipéptido híbrido;

(b) un promotor que se activa mediante dicho dominio de transactivación de dicho segundo polipéptido híbrido; y

(c) un gen cuya expresión va a modularse; y

b) introducir en dicha célula huésped dicho compuesto.

En otro aspecto, la presente invención proporciona un compuesto de la presente invención para su uso en la producción de un polipéptido en una célula huésped aislada que comprende las etapas de:

a) seleccionar una célula que es sustancialmente insensible a la exposición a dicho compuesto;

b) introducir en dicha célula:

1) un constructo de ADN que comprende:

i) un gen exógeno que codifica para el polipéptido; y

ii) un elemento de respuesta;

en el que dicho gen está bajo el control de dicho elemento de respuesta; y

2) un constructo de ADN que codifica para un complejo de receptor de ecdisona que comprende:

i) un dominio de unión a ADN;

5 ii) un dominio de unión a ligando del receptor de ecdisona; y

iii) un dominio de transactivación; y

c) exponer dicha célula a dicho compuesto.

10 En otro aspecto, la presente invención proporciona el uso de un compuesto de la presente invención para la fabricación de un medicamento para:

15 1) modular la expresión de un gen diana en una célula huésped, en el que la célula huésped incluye un primer casete de expresión génica que comprende un primer polinucleótido que codifica para un primer polipéptido que comprende:

(i) un dominio de transactivación;

20 (ii) un dominio de unión a ADN; y

(iii) un dominio de unión a ligando del receptor nuclear de grupo H; y

un segundo casete de expresión génica que comprende:

25 (i) un elemento de respuesta que puede unirse a dicho dominio de unión a ADN;

(ii) un promotor que se activa mediante dicho dominio de transactivación; y

30 (iii) dicho gen diana;

2) modular la expresión de uno o más genes exógenos en un sujeto;

35 3) regular la expresión génica endógena o heteróloga en un sujeto transgénico no humano que comprende poner en contacto dicho compuesto con un complejo de receptor de ecdisona dentro de las células de dicho sujeto, en el que dichas células contienen además una secuencia de unión a ADN para dicho complejo de receptor de ecdisona cuando está en combinación con dicho compuesto y en el que la formación de un complejo de receptor de ecdisona-compuesto-secuencia de unión a ADN induce la expresión de dicho gen;

40 4) modular la expresión de un gen en una célula huésped que comprende las etapas de:

a) introducir en dicha célula huésped un sistema de modulación de la expresión génica que comprende:

45 i) un primer casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que codifica para un primer polipéptido híbrido que comprende:

(a) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y

50 (b) un dominio de unión a ligando del receptor de ecdisona;

ii) un segundo casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que codifica para un segundo polipéptido híbrido que comprende:

55 (a) un dominio de transactivación; y

(b) un dominio de unión a ligando del receptor X retinoide quimérico; y

60 iii) un tercer casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que comprende:

(a) un elemento de respuesta reconocido por dicho dominio de unión a ADN de dicho primer polipéptido híbrido;

65 (b) un promotor que se activa mediante dicho dominio de transactivación de dicho segundo polipéptido híbrido; y

(c) un gen cuya expresión va a modularse; y

b) introducir en dicha célula huésped dicho compuesto; o

5) producir un polipéptido que comprende las etapas de:

5

a) seleccionar una célula que es sustancialmente insensible a la exposición a dicho compuesto;

b) introducir en dicha célula:

10

1) un constructo de ADN que comprende:

i) un gen exógeno que codifica para el polipéptido; y

15

ii) un elemento de respuesta;

en el que dicho gen está bajo el control de dicho elemento de respuesta; y

2) un constructo de ADN que codifica para un complejo de receptor de ecdisona que comprende:

20

i) un dominio de unión a ADN;

ii) un dominio de unión a ligando del receptor de ecdisona; y

25

iii) un dominio de transactivación; y

c) exponer dicha célula a dicho compuesto.

La presente invención se refiere a ligandos no esteroideos para su uso en un sistema de expresión génica inducible basado en receptores nucleares y en métodos de modulación de la expresión de un gen dentro de una célula huésped usando estos ligandos con sistemas de expresión génica inducible basados en receptores nucleares.

30

La descripción de los solicitantes también se refiere a métodos de modulación de la expresión génica en una célula huésped usando un sistema de modulación de la expresión génica con un ligando de la presente descripción. Específicamente, la descripción de los solicitantes proporciona un método de modulación de la expresión de un gen en una célula huésped que comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica según la descripción; b) introducir en la célula huésped un casete de expresión génica que comprende i) un elemento de respuesta que comprende un dominio al que se une el dominio de unión a ADN del primer polipéptido híbrido del sistema de modulación de la expresión génica; ii) un promotor que se activa mediante el dominio de transactivación del segundo polipéptido híbrido del sistema de modulación de la expresión génica; y iii) un gen cuya expresión va a modularse; y c) introducir en la célula huésped un ligando; con lo que con la introducción del ligando en la célula huésped, se modula la expresión del gen. La descripción de los solicitantes también proporciona un método de modulación de la expresión de un gen en una célula huésped que comprende un casete de expresión génica que comprende un elemento de respuesta que comprende un dominio al que se une el dominio de unión a ADN del primer polipéptido híbrido del sistema de modulación de la expresión génica; un promotor que se activa mediante el dominio de transactivación del segundo polipéptido híbrido del sistema de modulación de la expresión génica; y un gen cuya expresión va a modularse; en el que el método comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica según la descripción; y b) introducir en la célula huésped un ligando; con lo que con la introducción del ligando en el huésped, se modula expresión del gen.

50

### Descripción detallada de la invención

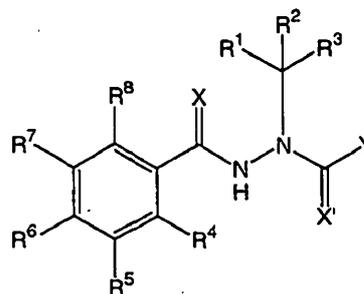
Los solicitantes han descubierto ligandos novedosos para receptores nucleares naturales y mutados. Por tanto, la descripción de los solicitantes proporciona un ligando para su uso con el sistema de expresión génica inducible basado en el receptor de ecdisona útil para modular la expresión de un gen de interés en una célula huésped. En una realización particularmente deseable, la descripción de los solicitantes proporciona un sistema de expresión génica inducible que tiene un nivel reducido de expresión génica de fondo y responde a concentraciones submicromolares de ligando no esteroideo. Por tanto, los ligandos novedosos de los solicitantes y el sistema de expresión génica inducible y su uso en métodos de modulación de la expresión génica en una célula huésped superan las limitaciones de los sistemas de expresión inducible disponibles actualmente y proporcionan al experto en la técnica un medio eficaz para controlar la expresión génica.

60

La presente invención es útil para aplicaciones tales como terapia génica, producción a gran escala de proteínas y anticuerpos, ensayos de selección de alto rendimiento basados en células, genómica, proteómica, metabolómica funcionales y regulación de caracteres en organismos transgénicos, donde son deseables niveles de control de la expresión génica. Una ventaja de la invención de los solicitantes es que proporciona un medio para regular la

65

expresión génica y adaptar los niveles de expresión para ajustarse a las necesidades del usuario.



La presente invención se refiere a compuestos de fórmula general:

5 en la que X y X' son independientemente O o S;

Y es:

10 (a) fenilo sustituido o no sustituido en el que los sustituyentes son independientemente 1-5 H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), alquenilo (C<sub>2</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), hidroxilo, amino, ciano o nitro; o

(b) 2-piridilo, 3-piridilo o 4-piridilo sustituidos o no sustituidos, en los que los sustituyentes son independientemente 1-4 H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), alquenilo (C<sub>2</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), hidroxilo, amino, ciano o nitro;

15 R<sup>1</sup> y R<sup>2</sup> son independientemente: H; ciano; alquilo de cadena lineal o ramificada (C<sub>1</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; alquenilo de cadena lineal o ramificada (C<sub>2</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; alquenilalquilo de cadena lineal o ramificada (C<sub>3</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; o juntas las valencias de R<sup>1</sup> y R<sup>2</sup> forman un grupo alquilideno (C<sub>1</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido (R<sup>a</sup>R<sup>b</sup>C=) en los que la suma de carbonos sin sustituyentes en R<sup>a</sup> y R<sup>b</sup> es de 0-6;

R<sup>3</sup> es H, metilo, etilo, n-propilo, isopropilo o ciano;

25 R<sup>4</sup>, R<sup>7</sup> y R<sup>8</sup> son independientemente: H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), alquenilo (C<sub>2</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), hidroxilo, amino, ciano o nitro; y

30 R<sup>5</sup> y R<sup>6</sup> son independientemente: H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alquenilo (C<sub>2</sub>-C<sub>4</sub>), alquenilalquilo (C<sub>3</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), hidroxilo, amino, ciano, nitro, o juntos como un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-) forman un anillo con los carbonos de fenilo a los que están unidos; en los que R<sup>9</sup> y R<sup>10</sup> son independientemente: H, halo, alquilo (C<sub>1</sub>-C<sub>3</sub>), alquenilo (C<sub>2</sub>-C<sub>3</sub>), alcoxi (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), benzoiloxi-alquilo (C<sub>1</sub>-C<sub>3</sub>), hidroxiloalquilo (C<sub>1</sub>-C<sub>3</sub>), haloalquilo (C<sub>1</sub>-C<sub>3</sub>), formilo, formilalquilo (C<sub>1</sub>-C<sub>3</sub>), ciano, cianoalquilo (C<sub>1</sub>-C<sub>3</sub>), carboxilo, carboxialquilo (C<sub>1</sub>-C<sub>3</sub>), alcoxicarbonil (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), alquilcarbonil (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), alcanoiloxi (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), aminoalquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-aminoalquilo (C<sub>1</sub>-C<sub>3</sub>)  $-(CH_2)_nR^cR^d$ , oximo (-CH=NOH), oximoalquilo (C<sub>1</sub>-C<sub>3</sub>), alcoximo (C<sub>1</sub>-C<sub>3</sub>) (-C=NOR<sup>d</sup>), alcoximoalquilo (C<sub>1</sub>-C<sub>3</sub>), carboxamido (C<sub>1</sub>-C<sub>3</sub>) (-C(O)NR<sup>e</sup>R<sup>f</sup>), carboxamido (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), semicarbazido (C<sub>1</sub>-C<sub>3</sub>) (-C=NNHC(O)NR<sup>e</sup>R<sup>f</sup>), semicarbazidoalquilo (C<sub>1</sub>-C<sub>3</sub>), aminocarboniloxilo (-OC(O)NHR<sup>9</sup>), aminocarboniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), pentafluorofeniloxicarbonilo, pentafluorofeniloxicarbonilalquilo (C<sub>1</sub>-C<sub>3</sub>), p-toluenosulfoniloxialquilo (C<sub>1</sub>-C<sub>3</sub>); arilsulfoniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), tio (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-sulfóxidoalquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-sulfonilalquilo (C<sub>1</sub>-C<sub>3</sub>) o siloxi trisustituido (C<sub>1</sub>-C<sub>5</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>)  $-(CH_2)_nSiOR^gR^hR^i$ ; en los que n=1-3, R<sup>c</sup> y R<sup>d</sup> representan cadenas hidrocarbonadas lineales o ramificadas de la longitud indicada, R<sup>e</sup>, R<sup>f</sup> representan H o cadenas hidrocarbonadas lineales o ramificadas de la longitud indicada, R<sup>g</sup> representa alquilo (C<sub>1</sub>-C<sub>3</sub>) o arilo opcionalmente sustituido con halo o alquilo (C<sub>1</sub>-C<sub>3</sub>) y R<sup>c</sup>, R<sup>d</sup>, R<sup>e</sup>, R<sup>f</sup> y R<sup>g</sup> son independientes unos de otros;

siempre que

45 i cuando R<sup>9</sup> y R<sup>10</sup> son ambos H, o

ii cuando cualquiera de R<sup>9</sup> o R<sup>10</sup> es halo, alquilo (C<sub>1</sub>-C<sub>3</sub>), alcoxi (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>) o benzoiloxialquilo (C<sub>1</sub>-C<sub>3</sub>), o

50 iii cuando R<sup>5</sup> y R<sup>6</sup> no forman juntos un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-),

entonces el número de átomos de carbono, excluyendo los de la sustitución de ciano, para cualquiera o ambos de los grupos R<sup>1</sup> o R<sup>2</sup> sea mayor que 4, y el número de átomos de carbono, excluyendo los de la sustitución de ciano, para la suma de los grupos R<sup>1</sup>, R<sup>2</sup> y R<sup>3</sup> sea de 10, 11 ó 12.

55

Se prefieren compuestos de la fórmula general cuando:

X y X' son O;

Y es:

5 (a) fenilo sustituido o no sustituido en el que los sustituyentes son independientemente 1-5 H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), ciano o nitro; o

10 (b) 2-piridilo, 3-piridilo o 4-piridilo sustituidos o no sustituidos, en los que los sustituyentes son independientemente 1-4 H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), ciano o nitro;

15 R<sup>1</sup> y R<sup>2</sup> son independientemente: H; ciano; alquilo de cadena lineal o ramificada (C<sub>1</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; alqueno de cadena lineal o ramificada (C<sub>2</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; alquenalquilo de cadena lineal o ramificada (C<sub>3</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; o juntas las valencias de R<sup>1</sup> y R<sup>2</sup> forman un grupo alquilideno (C<sub>1</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido (R<sup>a</sup>R<sup>b</sup>C=) en los que la suma de carbonos sin sustituyentes en R<sup>a</sup> y R<sup>b</sup> es de 0-6;

R<sup>3</sup> es H, metilo, etilo o ciano;

20 R<sup>4</sup>, R<sup>7</sup> y R<sup>8</sup> son independientemente: H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), ciano o nitro; y

25 R<sup>5</sup> y R<sup>6</sup> son independientemente: H, alquilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo C<sub>1</sub>-C<sub>4</sub>, alcoxilo (C<sub>1</sub>-C<sub>4</sub>), hidroxilo, amino, ciano, nitro, o juntos como un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-) forman un anillo con los carbonos de fenilo a los que están unidos; en los que R<sup>9</sup> o R<sup>10</sup> es H, y el otro de R<sup>9</sup> o R<sup>10</sup> es: H, haloalquilo (C<sub>1</sub>-C<sub>3</sub>), formilo, formilalquilo (C<sub>1</sub>-C<sub>3</sub>); ciano, cianoalquilo (C<sub>1</sub>-C<sub>3</sub>), carboxilo, carboxialquilo (C<sub>1</sub>-C<sub>3</sub>), aminoalquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-aminoalquilo (C<sub>1</sub>-C<sub>3</sub>) (-CH<sub>2</sub>)<sub>n</sub>R<sup>c</sup>R<sup>e</sup>), oximo (-CH=NOH), oximoalquilo (C<sub>1</sub>-C<sub>3</sub>), alcoximo (C<sub>1</sub>-C<sub>3</sub>) (-C=NOR<sup>d</sup>), alcoximoalquilo (C<sub>1</sub>-C<sub>3</sub>), carboxamido (C<sub>1</sub>-C<sub>3</sub>) (-C(O)NR<sup>e</sup>R<sup>f</sup>), carboxamido (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), semicarbazido (C<sub>1</sub>-C<sub>3</sub>) (-C=NNHC(O)NR<sup>e</sup>R<sup>f</sup>), semicarbazidoalquilo (C<sub>1</sub>-C<sub>3</sub>), aminocarboniloxilo (-OC(O)NHR<sup>9</sup>), aminocarboniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), pentafluorofeniloxycarbonilo, pentafluorofeniloxycarbonilalquilo (C<sub>1</sub>-C<sub>3</sub>), p-toluenosulfoniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), arilsulfoniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), tio (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-sulfoxidoalquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-sulfonilalquilo (C<sub>1</sub>-C<sub>3</sub>) o siloxi trisustituido (C<sub>1</sub>-C<sub>5</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>) (-CH<sub>2</sub>)<sub>n</sub>SiOR<sup>d</sup>R<sup>e</sup>R<sup>9</sup>); en los que n=1-3, R<sup>c</sup> y R<sup>d</sup> representan cadenas hidrocarbonadas lineales o ramificadas de la longitud indicada, R<sup>e</sup>, R<sup>f</sup> representan H o cadenas hidrocarbonadas lineales o ramificadas de la longitud indicada, R<sup>9</sup> representa alquilo (C<sub>1</sub>-C<sub>3</sub>) o arilo opcionalmente sustituido con halo o alquilo (C<sub>1</sub>-C<sub>3</sub>) y R<sup>c</sup>, R<sup>d</sup>, R<sup>e</sup>, R<sup>f</sup> y R<sup>9</sup> son independientes unos de otros;

siempre que i cuando R<sup>9</sup> y R<sup>10</sup> son ambos H, o

ii cuando R<sup>5</sup> y R<sup>6</sup> no forman juntos un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-),

40 entonces el número de átomos de carbono, excluyendo los de la sustitución de ciano, para cualquiera o ambos de los grupos R<sup>1</sup> o R<sup>2</sup> sea mayor que 4, y el número de átomos de carbono, excluyendo los de la sustitución de ciano, para la suma de los grupos R<sup>1</sup>, R<sup>2</sup> y R<sup>3</sup> sea de 10, 11 ó 12.

45 Se prefieren más compuestos de la fórmula general cuando:

X y X' son O;

Y es:

50 (a) fenilo sustituido o no sustituido en el que los sustituyentes son independientemente 1-5 H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), ciano o nitro; o

55 (b) 2-piridilo, 3-piridilo o 4-piridilo sustituidos o no sustituidos, en los que los sustituyentes son independientemente 1-4 H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), ciano o nitro;

60 R<sup>1</sup> y R<sup>2</sup> son independientemente: H; ciano; alquilo de cadena lineal o ramificada (C<sub>1</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; alqueno de cadena lineal o ramificada (C<sub>2</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; alquenalquilo de cadena lineal o ramificada (C<sub>3</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; o juntas las valencias de R<sup>1</sup> y R<sup>2</sup> forman un grupo alquilideno (C<sub>1</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido (R<sup>a</sup>R<sup>b</sup>C=) en los que la suma de carbonos sin sustituyentes en R<sup>a</sup> y R<sup>b</sup> es de 0-6;

R<sup>3</sup> es H, metilo, etilo o ciano;

65 R<sup>4</sup>, R<sup>7</sup> y R<sup>8</sup> son independientemente: H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>), ciano o nitro; y

$R^5$  y  $R^6$  son independientemente: H, alquilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo C<sub>1</sub>-C<sub>4</sub>, alcoxilo (C<sub>1</sub>-C<sub>4</sub>), hidroxilo, amino, ciano, nitro, o juntos como un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-) forman un anillo con los carbonos de fenilo a los que están unidos; en los que  $R^9$  o  $R^{10}$  es H, y el otro de  $R^9$  o  $R^{10}$  es: H, haloalquilo (C<sub>1</sub>-C<sub>3</sub>), formilo, formilalquilo (C<sub>1</sub>-C<sub>3</sub>), ciano, cianoalquilo (C<sub>1</sub>-C<sub>3</sub>), carboxilo, carboxialquilo (C<sub>1</sub>-C<sub>3</sub>), aminoalquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-aminoalquilo (C<sub>1</sub>-C<sub>3</sub>)  $-(CH_2)_nR^cR^e$ , oximo (-CH=NOH), oximoalquilo (C<sub>1</sub>-C<sub>3</sub>), alcoximo (C<sub>1</sub>-C<sub>3</sub>) (-C=NOR<sup>d</sup>), alcoximoalquilo (C<sub>1</sub>-C<sub>3</sub>), carboxamido (C<sub>1</sub>-C<sub>3</sub>) (-C(O)NR<sup>e</sup>R<sup>f</sup>), carboxamido (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), semicarbazido (C<sub>1</sub>-C<sub>3</sub>) (-C=NNHC(O)NR<sup>e</sup>R<sup>f</sup>), semicarbazidoalquilo (C<sub>1</sub>-C<sub>3</sub>), aminocarboniloxilo (-OC(O)NHR<sup>g</sup>), aminocarboniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), pentafluorofeniloxycarbonilo, pentafluorofeniloxycarbonilalquilo (C<sub>1</sub>-C<sub>3</sub>), p-toluenosulfoniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), arilsulfoniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), tio (C<sub>1</sub>-C<sub>3</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-sulfoxidoalquilo (C<sub>1</sub>-C<sub>3</sub>), alquil (C<sub>1</sub>-C<sub>3</sub>)-sulfonilalquilo (C<sub>1</sub>-C<sub>3</sub>) o siloxi trisustituido (C<sub>1</sub>-C<sub>5</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>)  $-(CH_2)_nSiOR^dR^eR^g$ ; en los que n=1-3, R<sup>c</sup> y R<sup>d</sup> representan cadenas hidrocarbonadas lineales o ramificadas de la longitud indicada, R<sup>e</sup>, R<sup>f</sup> representan H o cadenas hidrocarbonadas lineales o ramificadas de la longitud indicada, R<sup>g</sup> representa alquilo (C<sub>1</sub>-C<sub>3</sub>) o arilo opcionalmente sustituido con halo o alquilo (C<sub>1</sub>-C<sub>3</sub>) y R<sup>c</sup>, R<sup>d</sup>, R<sup>e</sup>, R<sup>f</sup> y R<sup>g</sup> son independientes unos de otros;

15 siempre que

cuando i  $R^9$  y  $R^{10}$  son ambos H, o

20 ii cuando  $R^5$  y  $R^6$  no forman juntos un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-),

entonces el número de átomos de carbono, excluyendo los de la sustitución de ciano, para cualquiera o ambos de los grupos  $R^1$  o  $R^2$  sea mayor que 4, y el número de átomos de carbono, excluyendo los de la sustitución de ciano, para la suma de los grupos  $R^1$ ,  $R^2$  y  $R^3$  sea de 10, 11 ó 12; y cuando  $R^5$  y  $R^6$  juntos como un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-) forman un anillo con los carbonos de fenilo a los que están unidos y  $R^9$  y  $R^{10}$  no son ambos H, entonces  $R^1$  y  $R^2$  sean alquilo de cadena lineal o ramificada (C<sub>1</sub>-C<sub>4</sub>) y  $R^3$  sea H o metilo.

Se prefieren incluso más compuestos de la fórmula general cuando:

30 X y X' son O;

Y es:

35 (a) fenilo sustituido o no sustituido en el que los sustituyentes son independientemente 1-5 H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>); o

(b) 3-piridilo sustituido o no sustituido, en el que los sustituyentes son independientemente 1-4 H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>);

40  $R^1$  y  $R^2$  son independientemente: H; ciano; alquilo de cadena lineal o ramificada (C<sub>1</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; alqueno de cadena lineal o ramificada (C<sub>2</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; alquenalquilo de cadena lineal o ramificada (C<sub>3</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido; o juntas las valencias de  $R^1$  y  $R^2$  forman un grupo alquilideno (C<sub>1</sub>-C<sub>7</sub>) sustituido con ciano o no sustituido ( $R^aR^bC=$ ) en los que la suma de carbonos sin sustituyentes en  $R^a$  y  $R^b$  es de 0-3;

45  $R^3$  es metilo;

$R^4$ ,  $R^7$  y  $R^8$  se seleccionan independientemente de: H, alquilo (C<sub>1</sub>-C<sub>4</sub>), alcoxilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo (C<sub>1</sub>-C<sub>4</sub>); y

50  $R^5$  y  $R^6$  son independientemente: H, alquilo (C<sub>1</sub>-C<sub>4</sub>), halo (F, Cl, Br, I), haloalquilo C<sub>1</sub>-C<sub>4</sub>, alcoxilo (C<sub>1</sub>-C<sub>4</sub>), o juntos como un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-) forman un anillo con los carbonos de fenilo a los que están unidos; en los que  $R^9$  o  $R^{10}$  es H, y el otro de  $R^9$  o  $R^{10}$  es: H, haloalquilo (C<sub>1</sub>-C<sub>2</sub>), formilo, cianoalquilo (C<sub>1</sub>-C<sub>2</sub>), carboxilo, aminoalquilo (C<sub>1</sub>-C<sub>2</sub>), oximo (-CH=NOH), carboxamido (C<sub>1</sub>-C<sub>3</sub>) (-C(O)NR<sup>e</sup>R<sup>f</sup>), semicarbazido (C<sub>1</sub>-C<sub>2</sub>) (-C=NNHC(O)NR<sup>e</sup>R<sup>f</sup>), aminocarboniloxilo (-OC(O)NHR<sup>g</sup>), pentafluorofeniloxycarbonilo, p-toluenosulfoniloxialquilo (C<sub>1</sub>-C<sub>3</sub>), metilalquilo (C<sub>1</sub>-C<sub>2</sub>), metilsulfoxidoalquilo (C<sub>1</sub>-C<sub>2</sub>), metilsulfonilalquilo (C<sub>1</sub>-C<sub>2</sub>) o siloxi trisustituido (C<sub>1</sub>-C<sub>5</sub>)-alquilo (C<sub>1</sub>-C<sub>3</sub>)  $-(CH_2)_nSiOR^dR^eR^g$ ; en los que n=1-3, R<sup>d</sup> representa una cadena hidrocarbonada lineal o ramificada de la longitud indicada, R<sup>e</sup>, R<sup>f</sup> representan H o cadenas hidrocarbonadas lineales o ramificadas de la longitud indicada, R<sup>g</sup> representa alquilo (C<sub>1</sub>-C<sub>3</sub>) o arilo opcionalmente sustituido con halo o alquilo (C<sub>1</sub>-C<sub>3</sub>) y R<sup>c</sup>, R<sup>d</sup>, R<sup>e</sup>, R<sup>f</sup> y R<sup>g</sup> son independientes unos de otros;

60 siempre que i) cuando  $R^9$  y  $R^{10}$  son ambos H, o

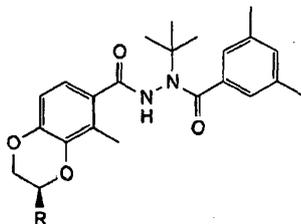
ii) cuando  $R^5$  y  $R^6$  no forman juntos un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-),

65 entonces el número de átomos de carbono, excluyendo los de la sustitución de ciano, para cualquiera o ambos de

los grupos R<sup>1</sup> o R<sup>2</sup> sea mayor que 4, y el número de átomos de carbono, excluyendo los de la sustitución de ciano, para la suma de los grupos R<sup>1</sup>, R<sup>2</sup> y R<sup>3</sup> sea de 10, 11 ó 12; y cuando R<sup>5</sup> y R<sup>6</sup> juntos como un enlace del tipo (-OCHR<sup>9</sup>CHR<sup>10</sup>O-) forman un anillo con los carbonos de fenilo a los que están unidos y R<sup>9</sup> y R<sup>10</sup> no son ambos H, entonces R<sup>1</sup> y R<sup>2</sup> sean metilo.

5

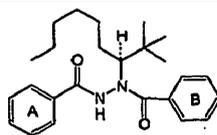
Los compuestos de la presente invención son los siguientes:



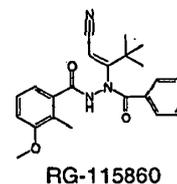
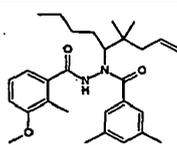
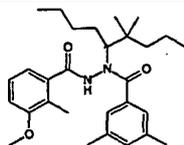
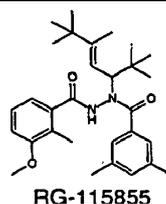
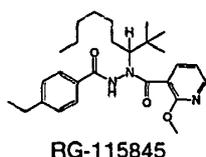
N.º de referencia de compuesto	R
RG-115789	-CH <sub>2</sub> OH
RG-115790	-CH <sub>2</sub> OSi(tBu)(CH <sub>3</sub> ) <sub>2</sub>
RG-115805	-CO <sub>2</sub> H
RG-115806	-CO <sub>2</sub> Me
RG-115807	-C=NNHCONH <sub>2</sub>
RG-115808	-CH <sub>2</sub> OC(O)NHPh
RG-115809	-CH <sub>2</sub> CH <sub>2</sub> NH <sub>2</sub>
RG-115810	-C(O)OC <sub>6</sub> F <sub>5</sub>
RG-115811	-CONHMe
RG-115812	-CHO
RG-115813	-CH <sub>2</sub> OS(O) <sub>2</sub> Ph-4-CH <sub>3</sub>
RG-115814	-C=NOH
RG-115815	-CH <sub>2</sub> F
RG-115816	-CH <sub>2</sub> CN
RG-115817	-CH <sub>2</sub> SCH <sub>3</sub>
RG-115818	-CH <sub>2</sub> S(O) <sub>2</sub> CH <sub>3</sub>

10

También se dan a conocer los siguientes compuestos:



N.º de referencia de compuesto	Sustitución en el anillo A	Sustitución en el anillo B
RG-115843	4-Et	3,5-di-CH <sub>3</sub>
RG-115844	4-Et	3,5-di-OCH <sub>3</sub> , 4-CH <sub>3</sub>
RG-115853	2-CH <sub>3</sub> , 3-OCH <sub>3</sub>	3,5-di-CH <sub>3</sub>
RG-115854	2-CH <sub>3</sub> , 3-OCH <sub>3</sub>	3,5-di-OCH <sub>3</sub> , 4-CH <sub>3</sub>



N.º RG	Sustitución en el anillo A	Sustitución en el anillo B
RG-115845	4-Et	2-OCH <sub>3</sub> -3-piridil
RG-115855	2-CH <sub>3</sub> , 3-OCH <sub>3</sub>	3,5-di-CH <sub>3</sub>
RG-115860	2-CH <sub>3</sub> , 3-OCH <sub>3</sub>	3,5-di-CH <sub>3</sub>
RG-115877	2-CH <sub>3</sub> , 3-OCH <sub>3</sub>	3,5-di-CH <sub>3</sub>
RG-115878	2-CH <sub>3</sub> , 3-OCH <sub>3</sub>	3,5-di-CH <sub>3</sub>

5 Puesto que los compuestos de la fórmula general de la presente descripción pueden contener varios átomos de carbono estereogénicos, los compuestos pueden existir como enantiómeros, diastereómeros, estereoisómeros o sus mezclas, aunque un centro estereogénico esté explícitamente especificado.

#### DEFINICIONES

10 Cuando se especifica un grupo R<sup>x</sup>, en el que x representa una letra a-g, y también se especifica el mismo grupo R<sup>x</sup> con una longitud de cadena de grupo alquilo tal como "(C<sub>1</sub>-C<sub>3</sub>)", se entiende que la longitud de cadena especificada sólo se refiere a casos en los que R<sup>x</sup> puede ser alquilo, y no se refiere a casos en los que R<sup>x</sup> puede ser un grupo distinto de alquilo, tal como H o arilo.

15 El término "alquilo" incluye grupos alquilo tanto de cadena lineal como ramificada. Los grupos alquilo típicos incluyen, por ejemplo, metilo, etilo, n-propilo, isopropilo, n-butilo, sec-butilo, isobutilo, terc-butilo, n-pentilo, isopentilo, n-hexilo, n-heptilo, isoocilo, nonilo y decilo.

El término "halo" se refiere a fluoro, cloro, bromo o yodo.

20 El término "haloalquilo" se refiere a un grupo alquilo sustituido con uno o más grupos halo tales como, por ejemplo, clorometilo, 2-bromoetilo, 3-yodopropilo, trifluorometilo y perfluoropropilo.

25 El término "cicloalquilo" se refiere a una estructura de anillo alifático cíclico, opcionalmente sustituido con alquilo, hidroxilo o halo, tal como ciclopropilo, metilciclopropilo, ciclobutilo, 2-hidroxiciclopentilo, ciclohexilo y 4-clorociclohexilo.

El término "hidroxialquilo" se refiere a un grupo alquilo sustituido con uno o más grupos hidroxilo tales como, por ejemplo, hidroximetilo y 2,3-dihidroxibutilo.

30 El término "alquilsulfonilo" se refiere a un resto sulfonilo sustituido con un grupo alquilo tal como, por ejemplo, mesilo y n-propilsulfonilo.

35 El término "alqueno" se refiere a un grupo hidrocarbonado etilénicamente insaturado, de cadena lineal o ramificada, que tiene 1 ó 2 uniones etilénicas tales como, por ejemplo, vinilo, alilo, 1-butenilo, 2-butenilo, isopropenilo y 2-pentenilo.

El término "haloalqueno" se refiere a un grupo alqueno sustituido con uno o más grupos halo.

40 El término "alquino" se refiere a un grupo hidrocarbonado insaturado, lineal o ramificado, que tiene 1 ó 2 uniones acetilénicas tales como, por ejemplo, etinilo y propargilo.

El término “alquilcarbonilo” se refiere a una funcionalidad alquilceto, por ejemplo acetilo, *n*-butirilo y similares.

El término “heterociclilo” o “heterociclo” se refiere a un anillo sustituido o no sustituido; saturado, parcialmente insaturado o insaturado, de 5 ó 6 miembros, que contiene uno, dos o tres heteroátomos, preferiblemente uno o dos heteroátomos seleccionados independientemente del grupo que consiste en oxígeno, nitrógeno y azufre. Los ejemplos de heterocicilos incluyen, por ejemplo, piridilo, tienilo, furilo, pirimidinilo, pirazinilo, quinolinilo, isoquinolinilo, pirrolilo, indolilo, tetrahidrofurilo, pirrolidinilo, piperidinilo, tetrahidropirranilo, morfolinilo, piperazinilo, dioxolanilo y dioxanilo.

El término “alcoxilo” incluye grupos alquilo tanto de cadena lineal como ramificada unidos a un átomo de oxígeno terminal. Los grupos alcoxilo típicos incluyen, por ejemplo, metoxilo, etoxilo, *n*-propoxilo, isopropoxilo y terc-butoxilo.

El término “haloalcoxilo” se refiere a un grupo alcoxilo sustituido con uno o más grupos halo tales como, por ejemplo clorometoxilo, trifluorometoxilo, difluorometoxilo y perfluoroisobutoxilo.

El término “alquiltio” incluye grupos alquilo tanto de cadena lineal como ramificada, unidos a un átomo de azufre terminal tal como, por ejemplo metiltio.

El término “haloalquiltio” se refiere a un grupo alquiltio sustituido con uno o más grupos halo tales como, por ejemplo trifluorometiltio.

El término “alcoxialquilo” se refiere a un grupo alquilo sustituido con un grupo alcoxilo tal como, por ejemplo, isopropoximetilo.

“Cromatografía en gel de sílice” se refiere a un método de purificación en el que se aplica una sustancia química de interés como una muestra concentrada a la parte superior de una columna vertical de gel de sílice o gel de sílice químicamente modificado contenido en un cilindro de vidrio, plástico o metal, y la elución de tal columna con un disolvente o mezcla de disolventes.

“Cromatografía ultrarrápida” se refiere a cromatografía en gel de sílice realizada bajo presión de aire, argón o nitrógeno normalmente en el intervalo de 10 a 50 psi.

“Cromatografía en gradiente” se refiere a cromatografía en gel de sílice en la que la sustancia química se eluye de una columna con una composición que cambia progresivamente de una mezcla de disolventes.

“Rf” es un término de cromatografía en capa fina que se refiere a la distancia fraccional de movimiento de una sustancia química de interés sobre una placa de cromatografía en capa fina, con respecto a la distancia de movimiento del sistema de disolventes en elución.

“Hidrogenador Parr” y “agitador Parr” se refieren a aparatos disponibles de Parr Instrument Company, Moline IL, que están diseñado para facilitar un mezclado vigoroso de una disolución que contiene una sustancia química de interés con un catalizador suspendido sólido opcional y una atmósfera contenida, presurizada, de un gas reactivo. Normalmente, el gas es hidrógeno y el catalizador es paladio, platino u óxidos de los mismos depositados sobre partículas de carbón pequeñas. La presión de hidrógeno está normalmente en el intervalo de 30 a 70 psi.

“Reactivo de Dess-Martin” se refiere a (1,1,1-triacetoxi)-1,1-dihidro-1,2-benciodoxol-3(1H)-ona como disolución en diclorometano disponible de Acros Organics/Fisher Scientific Company, L.L.C.

“PS-NMM” se refiere a una resina de poliestireno funcionalizada con -SO<sub>2</sub>NH(CH<sub>2</sub>)<sub>3</sub>-morfolina disponible de Argonaut Technologies, San Carlos, CA.

“AP-NCO” se refiere a una resina funcionalizada con isocianato disponible de Argonaut Technologies, San Carlos, CA.

“AP-trisamina” se refiere a una resina de poliestireno-CH<sub>2</sub>NHCH<sub>2</sub>CH<sub>3</sub>NH(CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>)<sub>2</sub> disponible de Argonaut Technologies, San Carlos, CA.

El término “aislado” para los fines de la presente invención designa un material biológico (ácido nucleico o proteína) que se ha retirado de su entorno original (el entorno en que está presente de manera natural). Por ejemplo, un polinucleótido presente en estado natural en una planta o un animal no está aislado, sin embargo el mismo polinucleótido separado de los ácidos nucleicos adyacentes en que está presente de manera natural, se considera “aislado”. El término “purificado” no requiere que el material esté presente en una forma que muestre pureza absoluta, excluyendo la presencia de otros compuestos. Es más bien una definición relativa.

Un polinucleótido está en el estado “purificado” tras la purificación del material de partida o del material natural en al menos un orden de magnitud, preferiblemente 2 ó 3 y preferiblemente 4 ó 5 órdenes de magnitud.

Un “ácido nucleico” es un compuesto polimérico compuesto por subunidades unidas covalentemente denominadas nucleótidos. Ácido nucleico incluye poli(ácido ribonucleico) (ARN) y poli(ácido desoxirribonucleico) (ADN), pudiendo ser ambos monocatenarios o bicatenarios. El ADN incluye, pero no se limita a, ADNc, ADN genómico, ADN de plásmidos, ADN sintético y ADN semi-sintético. El ADN puede ser lineal, circular o superenrollado.

Una “molécula de ácido nucleico” se refiere a la forma polimérica de éster fosfato de ribonucleósidos (adenosina, guanosina, uridina o citidina; “moléculas de ARN”) o desoxirribonucleósidos (desoxiadenosina, desoxiguanosina, desoxitimidina o desoxicitidina; “moléculas de ADN”), o cualquier análogo de fosfoéster de los mismos, tales como fosforotioatos y tioésteres, o bien en forma monocatenaria o bien como una hélice bicatenaria. Son posibles hélices de ADN-ADN, ADN-ARN y ARN-ARN bicatenarias. El término molécula de ácido nucleico, y en particular molécula de ADN o ARN, se refiere sólo a la estructura primaria o secundaria de la molécula y no la limita a ninguna forma terciaria particular. Por tanto, este término incluye ADN bicatenario encontrado, entre otros en moléculas de ADN lineales o circulares (por ejemplo, fragmentos de restricción), plásmidos y cromosomas. Al comentar la estructura de moléculas de ADN bicatenarias particulares, las secuencias pueden describirse en el presente documento según el convenio normal de facilitar sólo la secuencia en el sentido de 5’ a 3’ a lo largo de la cadena no transcrita de ADN (es decir, la cadena que tiene una secuencia homóloga al ARNm). Una “molécula de ADN recombinante” es una molécula que se ha sometido a manipulación de biología molecular.

El término “fragmento” se entenderá que significa una secuencia de nucleótido de longitud reducida con respecto al ácido nucleico de referencia y que comprende, en la parte común, una secuencia de nucleótido idéntica al ácido nucleico de referencia. Un fragmento de ácido nucleico de este tipo según la descripción puede estar incluido, cuando sea apropiado, en un polinucleótido más largo del que es un constituyente. Tales fragmentos comprenden, o alternativamente consisten en, oligonucleótidos que oscilan en longitud entre al menos 6, 8, 9, 10, 12, 15, 18, 20, 21, 22, 23, 24, 25, 30, 39, 40, 42, 45, 48, 50, 51, 54, 57, 60, 63, 66, 70, 75, 78, 80, 90, 100, 105, 120, 135, 150, 200, 300, 500, 720, 900, 1000 o 1500 nucleótidos consecutivos de un ácido nucleico según la descripción.

Tal como se usa en el presente documento, un “fragmento de ácido nucleico aislado” es un polímero de ARN o ADN que es mono o bicatenario, que contiene opcionalmente bases nucleotídicas sintéticas, no naturales o alteradas. Un fragmento de ácido nucleico aislado en forma de un polímero de ADN puede estar compuesto por uno o más segmentos de ADNc, ADN genómico o ADN sintético.

Un “gen” se refiere a un conjunto de nucleótidos que codifican para un polipéptido, e incluye ácidos nucleicos de ADNc y ADN genómico. “Gen” también se refiere a un fragmento de ácido nucleico que expresa una proteína o polipéptido específico, incluyendo secuencias reguladoras que preceden (secuencias no codificantes en 5’) y que siguen (secuencias no codificantes en 3’) a la secuencia codificante. “Gen nativo” se refiere a un gen tal como se encuentra en la naturaleza con sus propias secuencias reguladoras. “Gen quimérico” se refiere a cualquier gen que no es un gen nativo, que comprende secuencias reguladoras y/o codificantes que no se encuentran juntas en la naturaleza. Por consiguiente, un gen quimérico puede comprender secuencias reguladoras y secuencias codificantes que se derivan de diferentes fuentes, o secuencias reguladoras y secuencias codificantes derivadas de la misma fuente, pero dispuestas de manera diferente a la encontrada en la naturaleza. Un gen quimérico puede comprender secuencias codificantes derivadas de diferentes fuentes y/o secuencias reguladoras derivadas de diferentes fuentes. “Gen endógeno” se refiere a un gen nativo en su ubicación natural en el genoma de un organismo. Un gen “foráneo” o gen “heterólogo” se refiere a un gen normalmente no encontrado en el organismo huésped, sino que se introduce en el organismo huésped mediante transferencia génica. Los genes foráneos pueden comprender genes nativos insertados en un organismo no nativo, o genes quiméricos. Un “transgén” es un gen que se ha introducido en el genoma mediante un procedimiento de transformación.

ADN “heterólogo” se refiere a ADN no ubicado de manera natural en la célula, o en un sitio cromosómico de la célula. Preferiblemente, el ADN heterólogo incluye un gen foráneo a la célula.

El término “genoma” incluye ADN o ARN cromosómico, así como mitocondrial, de cloroplasto y viral.

Una molécula de ácido nucleico “puede hibridarse” con otra molécula de ácido nucleico, tal como un ADNc, ADN genómico o ARN, cuando una forma monocatenaria de la molécula de ácido nucleico puede aparearse con otra molécula de ácido nucleico en condiciones apropiadas de temperatura y fuerza iónica de la disolución (véase Sambrook *et al.*, 1989 citado posteriormente). Se conocen bien las condiciones de hibridación y lavado y se facilitan a modo de ejemplo en Sambrook, J., Fritsch, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*, segunda edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989), particularmente capítulo 11 y tabla 11.1 en ese documento. Las condiciones de temperatura y fuerza iónica determina la “rigurosidad” de la hibridación. Las condiciones de rigurosidad pueden ajustarse para seleccionar fragmentos moderadamente similares, tales como secuencias homólogas de organismos relacionados de manera distante, hasta fragmentos altamente similares, tales como genes que duplican las enzimas funcionales de organismos estrechamente relacionados. Para la selección preliminar de ácidos nucleicos homólogos, pueden usarse condiciones de hibridación de baja rigurosidad, que se corresponden con una  $T_m$  de 55°, por ejemplo, 5x SSC, SDS al 0,1%, leche al 0,25% y sin formamida; o formamida al 30%, 5x SSC, SDS al 0,5%). Las condiciones de hibridación de rigurosidad moderada se corresponden con una

$T_m$  superior, por ejemplo, formamida al 40%, con 5x o 6x SCC. Las condiciones de hibridación de alta rigurosidad se corresponden con la  $T_m$  más alta, por ejemplo, formamida al 50%, 5x o 6x SCC.

5 La hibridación requiere que los dos ácidos nucleicos contengan secuencias complementarias, aunque dependiendo de la rigurosidad de la hibridación, son posibles apareamientos erróneos entre bases. El término "complementario" se usa para describir la relación entre bases de nucleótido que pueden hibridarse entre sí. Por ejemplo, con respecto al ADN, adenosina es complementaria a timina y citosina es complementaria a guanina. Por consiguiente, la presente invención también incluye fragmentos de ácido nucleico aislados que son complementarios a las secuencias completas tal como se da a conocer o se usa en el presente documento, así como aquellas secuencias de ácido nucleico sustancialmente similares.

15 En una realización específica de la descripción, los polinucleótidos se detectan empleando condiciones de hibridación que comprenden una etapa de hibridación a  $T_m$  de 55°C, y utilizando condiciones tal como se expusieron anteriormente. En una realización preferida, la  $T_m$  es de 60°C; en una realización más preferida, la  $T_m$  es de 63°C; en una realización incluso más preferida, la  $T_m$  es de 65°C.

20 Los lavados posteriores a la hibridación también determinan las condiciones de rigurosidad. Un conjunto de condiciones preferidas usa una serie de lavados partiendo de 6X SSC, SDS al 0,5% a temperatura ambiente durante 15 minutos (min.), repetidos después con 2X SSC, SDS al 0,5% a 45°C durante 30 minutos, y repetidos después dos veces con 0,2X SSC, SDS al 0,5% a 50°C durante 30 minutos. Un conjunto más preferido de condiciones rigurosas usa temperaturas superiores en las que los lavados son idénticos a los indicados anteriormente excepto por la temperatura de los dos lavados finales de 30 min. en 0,2 X SSC, SDS al 0,5%, que se aumentó hasta 60°C. Otro conjunto preferido de condiciones altamente rigurosas usa dos lavados finales en 0,1X SSC, SDS al 0,1% a 65°C. La hibridación requiere que los dos ácidos nucleicos comprendan secuencias complementarias, aunque dependiendo de la rigurosidad de la hibridación, son posibles apareamientos erróneos entre bases.

30 La rigurosidad apropiada para hibridar ácidos nucleicos depende de la longitud de los ácidos nucleicos y del grado de complementación, variables bien conocidas en la técnica. Cuanto mayor es el grado de similitud u homología entre dos secuencias de nucleótido, mayor es el valor de  $T_m$  para híbridos de ácidos nucleicos que tienen esas secuencias. La estabilidad relativa (que se corresponde con una  $T_m$  superior) de las hibridaciones de ácidos nucleicos disminuye en el orden siguiente: ARN:ARN, ADN:ARN y ADN:ADN. Para híbridos de longitud mayor de 100 nucleótidos, se han obtenido ecuaciones para calcular la  $T_m$  (véase Sambrook *et al.*, citado anteriormente, 9.50-0.51). Para la hibridación con ácidos nucleicos más cortos, es decir oligonucleótidos, la posición de los apareamientos erróneos se vuelve más importante y la longitud del oligonucleótido determina su especificidad (véase Sambrook *et al.*, citado anteriormente, 11.7-11.8).

40 En una realización específica de la descripción, los polinucleótidos se detectan empleando condiciones de hibridación que comprenden una etapa de hibridación en sal a menos de 500 mM y a al menos 37 grados centígrados, y una etapa de lavado en 2X SSPE a al menos 63 grados centígrados. En una realización preferida, las condiciones de hibridación comprenden sal a menos de 200 mM y al menos 37 grados centígrados para la etapa de hibridación. En una realización más preferida, las condiciones de hibridación comprenden 2X SSPE y 63 grados centígrados para las etapas tanto de hibridación como de lavado.

45 En una realización, la longitud para un ácido nucleico que puede hibridarse es de al menos aproximadamente 10 nucleótidos. Preferiblemente, una longitud mínima para un ácido nucleico que puede hibridarse es de al menos aproximadamente 15 nucleótidos; más preferiblemente de al menos aproximadamente 20 nucleótidos y lo más preferiblemente la longitud es de al menos 30 nucleótidos. Además, el experto en la técnica reconocerá que la temperatura y la concentración de sal de la solución de lavado pueden ajustarse según sea necesario según factores tales como la longitud de la sonda

50 El término "sonda" se refiere a una molécula de ácido nucleico monocatenario que puede formar pares de bases con un ácido nucleico diana monocatenario complementario para formar una molécula bicatenaria.

55 Tal como se usa en el presente documento, el término "oligonucleótido" se refiere a un ácido nucleico, en general de al menos 18 nucleótidos, que puede hibridarse con una molécula de ADN genómico, una molécula de ADNc, un ADN de plásmido o una molécula de ARNm. Los oligonucleótidos pueden marcarse, por ejemplo, con <sup>32</sup>P-nucleótidos o nucleótidos con los que se ha conjugado covalentemente un marcador, tal como biotina. Un oligonucleótido marcado puede usarse como sonda para detectar la presencia de un ácido nucleico. Los oligonucleótidos (pudiendo estar uno o ambos marcados) pueden usarse como cebadores de PCR, para clonar o bien la longitud completa o bien un fragmento de un ácido nucleico, o para detectar la presencia de un ácido nucleico. Un oligonucleótido también puede usarse para formar una triple hélice con una molécula de ADN. En general, los oligonucleótidos se preparan de manera sintética, preferiblemente en un sintetizador de ácidos nucleicos. Por consiguiente, los oligonucleótidos pueden prepararse con uniones análogas a fosfoéster que no se producen de manera natural, tales como uniones de tioéster, etc.

65 Un "cebador" es un oligonucleótido que se hibrida con una secuencia de ácido nucleico diana para crear una región

de ácido nucleico bicatenario que puede servir como punto de inicio para la síntesis de ADN en condiciones adecuadas. Tales cebadores pueden usarse en una reacción en cadena de la polimerasa.

5 “Reacción en cadena de la polimerasa” se abrevia como PCR y significa un método *in vitro* para amplificar enzimáticamente secuencias específicas de ácido nucleico. La PCR implica una serie repetitiva de ciclos de temperatura comprendiendo cada ciclo tres fases: desnaturalización del ácido nucleico de molde para separar las cadenas de la molécula diana, apareamiento de un cebador oligonucleotídico de PCR monocatenario con el ácido nucleico de molde y extensión del(de los) cebador(es) apareado(s) mediante ADN polimerasa. La PCR proporciona un medio para detectar la presencia de la molécula diana y, en condiciones cuantitativas o semicuantitativas, 10 determinar la cantidad relativa de esa molécula diana dentro del grupo de partida de ácidos nucleicos.

15 “Reacción en cadena de la polimerasa-transcripción inversa” se abrevia como RT-PCR y significa un método *in vitro* para producir enzimáticamente una molécula o moléculas de ADNc diana a partir de una molécula o moléculas de ARN, seguido por la amplificación enzimática de una secuencia o secuencias específicas de ácido nucleico dentro de la molécula o moléculas de ADNc diana tal como se describió anteriormente. La RT-PCR también proporciona un medio para detectar la presencia de la molécula diana y, en condiciones cuantitativas o semicuantitativas, determinar la cantidad relativa de esa molécula diana dentro del grupo de partida de ácidos nucleicos.

20 Una “secuencia codificante” de ADN es una secuencia de ADN bicatenario que se transcribe y traduce para dar un polipéptido en una célula *in vitro* o *in vivo* cuando se coloca bajo el control de secuencias reguladoras apropiadas. “Secuencias reguladoras adecuadas” se refiere a secuencias de nucleótido ubicadas aguas arriba (secuencias no codificantes en 5’), dentro de o aguas abajo (secuencias no codificantes en 3’) de una secuencia codificante, y que influyen en la transcripción, procesamiento o estabilidad de ARN, o traducción de la secuencia codificante asociada. Las secuencias reguladoras pueden incluir promotores, secuencias líder de traducción, intrones, secuencias de reconocimiento de poliadenilación, sitio de procesamiento de ARN, sitio de unión de efectores y estructura de tallo-bucle. Los límites de la secuencia codificante están determinados por un codón de inicio en el extremo 5’ terminal (amino) y por un codón de terminación de la traducción en el extremo 3’ terminal (carboxilo). Una secuencia codificante puede incluir, pero no se limita a, secuencias de procariotas, ADNc a partir del ARNm, secuencias de ADN genómico e incluso secuencias de ADN sintético. Si la secuencia codificante está destinada a la expresión en una célula eucariota, habitualmente se ubicará una señal de poliadenilación y una secuencia de terminación de la transcripción en 3’ con respecto a la secuencia codificante

35 “Marco de lectura abierto” se abrevia ORF y significa una longitud de secuencia de ácido nucleico, o bien de ADN, ADNc o bien de ARN, que comprende una señal de inicio o un codón de inicio de la traducción, tal como ATG o AUG, y un codón de terminación y que puede traducirse potencialmente en una secuencia de polipéptido

40 El término “cabeza con cabeza” se usa en el presente documento para describir la orientación de dos secuencias de polinucleótido una en relación con la otra. Dos polinucleótidos están colocados en una orientación cabeza con cabeza cuando el extremo 5’ de la cadena codificante en un polinucleótido es adyacente al extremo 5’ de la cadena codificante en el otro polinucleótido, por lo que el sentido de transcripción de cada polinucleótido avanza alejándose del extremo 5’ del otro polinucleótido. El término “cabeza con cabeza” puede abreviarse (5’) con (5’) y también puede indicarse mediante los símbolos ( $\leftarrow \rightarrow$ ) o ( $3' \leftarrow 5' 5' \rightarrow 3'$ ).

45 El término “cola con cola” se usa en el presente documento para describir la orientación de dos secuencias de polinucleótido una en relación con la otra. Dos polinucleótidos están colocados en una orientación cola con cola cuando el extremo 3’ de la cadena codificante de un polinucleótido es adyacente al extremo 3’ de la cadena codificante del otro polinucleótido, por lo que el sentido de transcripción de cada polinucleótido avanza hacia el otro polinucleótido. El término “cola con cola” puede abreviarse (3’) con (3’) y también puede indicarse mediante los símbolos ( $\rightarrow \leftarrow$ ) o ( $5' \rightarrow 3' 3' \leftarrow 5'$ ).

50 El término “cabeza con cola” se usa en el presente documento para describir la orientación de dos secuencias de polinucleótido una en relación con la otra. Dos polinucleótidos están colocados en una orientación cabeza con cola cuando el extremo 5’ de la cadena codificante de un polinucleótido es adyacente al extremo 3’ de la cadena codificante del otro polinucleótido, por lo que el sentido de transcripción de cada polinucleótido avanza en el mismo sentido que el del otro polinucleótido. El término “cabeza con cola” puede abreviarse (5’) con (3’) y también puede indicarse mediante los símbolos ( $\rightarrow \rightarrow$ ) o ( $5' \rightarrow 3' 5' \rightarrow 3'$ ).

60 El término “aguas abajo” se refiere a una secuencia de nucleótido que está ubicada hacia 3’ con respecto a la secuencia de nucleótido de referencia. En particular, secuencias de nucleótido aguas abajo se refiere en general a secuencias que siguen al punto de inicio de la transcripción. Por ejemplo, el codón de inicio de la transcripción de un gen está ubicado aguas abajo del sitio de inicio de la transcripción.

65 El término “aguas arriba” se refiere a una secuencia de nucleótido que está ubicada hacia 5’ con respecto a la secuencia de nucleótido de referencia. En particular, secuencias de nucleótido aguas arriba se refiere en general a secuencias que están ubicadas en el lado de 5’ de una secuencia codificante o punto de inicio de la transcripción. Por ejemplo, la mayoría de los promotores están ubicados aguas arriba del sitio de inicio de la transcripción.

Los términos “endonucleasa de restricción” y “enzima de restricción” se refieren a una enzima que se une y corta dentro de una secuencia de nucleótido específica dentro del ADN bicatenario.

5 “Recombinación homóloga” se refiere a la inserción de una secuencia de ADN foránea en otra molécula de ADN, por ejemplo, la inserción de un vector en un cromosoma. Preferiblemente, el vector selecciona como diana un sitio cromosómico específico para recombinación homóloga. Para la recombinación homóloga específica, el vector contendrá regiones de homología suficientemente largas con secuencias del cromosoma para permitir la unión complementaria y la incorporación del vector en el cromosoma. Regiones de homología más largas y mayores  
10 grados de similitud de secuencia, pueden aumentar la eficacia de la recombinación homóloga.

Pueden usarse varios métodos conocidos en la técnica para propagar un polinucleótido según la descripción. Una vez que se ha establecido un sistema huésped y condiciones de crecimiento adecuados, pueden propagarse vectores de expresión recombinantes y prepararse en cantidad. Tal como se describe en el presente documento, los  
15 vectores de expresión que pueden usarse incluyen, pero no se limitan a, los siguientes vectores o sus derivados: virus de seres humanos o animales tales como el virus vaccinia o adenovirus; virus de insectos tales como baculovirus; vectores de levaduras; vectores de bacteriófagos (por ejemplo, lambda) y vectores de ADN de plásmido y cósmido, por nombrar sólo algunos.

20 Un “vector” es cualquier medio para la clonación de y/o la transferencia de un ácido nucleico en una célula huésped. Un vector puede ser un replicón al que puede unirse otro segmento de ADN para provocar la replicación del segmento unido. Un “replicón” es cualquier elemento genético (por ejemplo, plásmido, fago, cósmido, cromosoma, virus) que funciona como una unidad autónoma de replicación de ADN *in vivo*, es decir, que puede replicarse bajo su propio control. El término “vector” incluye tanto medios virales como no virales para introducir el ácido nucleico en  
25 una célula *in vitro*, *ex vivo* o *in vivo*. Puede usarse un gran número de vectores conocidos en la técnica para manipular ácidos nucleicos, incorporar elementos de respuesta y promotores en genes, etc. Los vectores posibles incluyen, por ejemplo, plásmidos o virus modificados incluyendo, por ejemplo bacteriófagos tales como derivados de lambda, o plásmidos tales como los derivados de plásmido pBR322 o pUC, o el vector Bluescript. Por ejemplo, la inserción de los fragmentos de ADN que se corresponden con elementos de respuesta y promotores en un vector  
30 adecuado puede llevarse a cabo mediante el ligamiento de fragmentos de ADN apropiados en un vector escogido que tiene extremos terminales cohesivos complementarios. Alternativamente, los extremos de las moléculas de ADN pueden modificarse enzimáticamente o puede producirse cualquier sitio mediante el ligamiento de secuencias de nucleótido (ligadores) en los extremos terminales del ADN. Tales vectores pueden modificarse por ingeniería genética para contener genes marcadores seleccionables que proporcionan la selección de células que han  
35 incorporado el marcador en el genoma celular. Tales marcadores permiten la identificación y/o selección de células huésped que incorporan y expresan las proteínas codificadas por el marcador.

Se han usado vectores virales, y particularmente vectores retrovirales, en una amplia variedad de aplicaciones de suministro génico en células, así como en sujetos animales vivos. Los vectores virales que pueden usarse incluyen,  
40 pero no se limitan a, vectores de retrovirus, virus adenoasociados, poxvirus, baculovirus, virus vaccinia, virus del herpes simple, virus de Epstein-Barr, adenovirus, geminivirus y caulimovirus. Los vectores no virales incluyen plásmidos, liposomas, lípidos cargados eléctricamente (citofectinas), complejos de ADN-proteínas y biopolímeros. Además de un ácido nucleico, un vector puede comprender también una o más regiones reguladoras y/o marcadores seleccionables útiles en la selección, medición y monitorización de resultados de transferencia de ácido  
45 nucleico (transferencia a tejidos, duración de expresión, etc.).

El término “plásmido” se refiere a un elemento extracromosómico que porta a menudo un gen que no forma parte del metabolismo central de la célula y habitualmente en forma de moléculas de ADN bicatenario circular. Tales  
50 elementos pueden ser secuencias que se replican de manera autónoma, secuencias de integración de genoma, secuencias de fagos o nucleótidos, lineales, circulares o superenrolladas, de un ADN o ARN monocatenario o bicatenario, derivadas de cualquier fuente, en que varias secuencias de nucleótido se han unido o recombinado en una única construcción que puede introducir en una célula un fragmento de promotor y una secuencia de ADN para un producto génico seleccionado junto con la secuencia no traducida en 3' apropiada.

55 Un “vector de clonación” es un “replicón”, que es una longitud de unidad de un ácido nucleico, preferiblemente ADN, que se replica secuencialmente y que comprende un origen de replicación, tal como un plásmido, fago o cósmido, al que puede unirse otro segmento de ácido nucleico para provocar la replicación del segmento unido. Los vectores de clonación pueden replicarse en un tipo celular y expresarse en otro (“vector lanzadera”)

60 Los vectores pueden introducirse en las células huésped deseadas mediante métodos conocidos en la técnica, por ejemplo transfección, electroporación, microinyección, transducción, fusión celular, DEAE-dextrano, precipitación de fosfato cálcico, lipofección (fusión de lisosomas), uso de una pistola génica o de un transportador de vector de ADN (véase, por ejemplo, Wu *et al.*, 1992, J. Biol. Chem. 267: 963-967; Wu y Wu, 1988, J. Biol. Chem. 263: 14621-14624; y Hartmut *et al.*, solicitud de patente canadiense n.º 2.012.311, presentada el 15 de marzo de 1990).

65 Un polinucleótido según la descripción también puede introducirse *in vivo* mediante lipofección. Durante la última

década, ha habido un uso creciente de liposomas para encapsulación y transfección de ácidos nucleicos *in vitro*. Pueden usarse lípidos catiónicos sintéticos diseñados para limitar las dificultades y peligros encontrados con la transfección mediada por liposomas para preparar liposomas para transfección *in vivo* de un gen que codifica para un marcador (Felgner *et al.*, 1987, PNAS 84:7413; Mackey, *et al.*, 1988. Proc. Natl. Acad. Sci. U.S.A. 85:8027-8031; y Ulmer *et al.*, 1993, Science 259:1745-1748). El uso de lípidos catiónicos puede promover la encapsulación de ácidos nucleicos cargados negativamente y también promover la fusión con membranas celulares cargadas negativamente (Felgner y Ringold, 1989, Science 337: 387-388). Compuestos y composiciones particularmente útiles para transferir ácidos nucleicos se describen en las publicaciones de patente internacional WO 95/18863 y WO 96/17823, y en la patente estadounidense n.º 5.459.127. El uso de lipofección para introducir genes exógenos en órganos específicos *in vivo* tiene determinadas ventajas prácticas. El direccionamiento molecular de liposomas a células específicas representa un área de beneficio. Está claro que dirigir la transfección a tipos celulares particulares sería particularmente preferido en un tejido con heterogeneidad celular, tal como páncreas, hígado y cerebro. Los lípidos pueden acoplarse químicamente a otras moléculas para el fin del direccionamiento (Mackey, *et al.*, 1988, citado anteriormente). Péptidos dirigidos, por ejemplo, hormonas o neurotransmisores, y proteínas tales como anticuerpos, o moléculas no peptídicas, podrían acoplarse químicamente a liposomas.

También son útiles otras moléculas para facilitar la transfección de un ácido nucleico *in vivo*, tal como un oligopéptido catiónico (por ejemplo, documento WO 95/21931), péptidos derivados de proteínas de unión a ADN (por ejemplo, documento WO 96/25508) o un polímero catiónico (por ejemplo, documento WO 95/21931).

También es posible introducir un vector *in vivo* como un plásmido de ADN desnudo (véanse las patentes estadounidenses 5.693.622, 5.589.466 y 5.580.859). También pueden usarse enfoques de suministro de ADN mediado por receptor (Curiel *et al.*, 1992, Hum. Gene Ther. 3: 147-154; y Wu y Wu, 1987, J. Biol. Chem. 262: 4429-4432).

El término "transfección" significa la captación de ARN o ADN exógeno o heterólogo por una célula. Una célula se ha "transfectado" por ARN o ADN exógeno o heterólogo cuando tal ARN o ADN se ha introducido dentro de la célula. Una célula se ha "transformado" mediante ARN o ADN exógeno o heterólogo cuando el ARN o ADN transfectado realiza un cambio fenotípico. El ARN o ADN transformante puede integrarse (unirse covalentemente) en el ADN cromosómico que constituye el genoma de la célula.

"Transformación" se refiere a la transferencia de un fragmento de ácido nucleico en el genoma de un organismo huésped, dando como resultado una herencia genéticamente estable. Los organismos huésped que contienen los fragmentos de ácido nucleico transformados se denominan organismos "transgénicos" o "recombinantes" o "transformados".

El término "región genética" se referirá a una región de una molécula de ácido nucleico o una secuencia de nucleótido que comprende un gen que codifica para un polipéptido.

Además, el vector recombinante que comprende un polinucleótido según la descripción puede incluir uno o más orígenes para la replicación en huéspedes celulares en que se busca su amplificación o su expresión, marcadores o marcadores seleccionables.

El término "marcador seleccionable" significa un factor de identificación, habitualmente un gen de resistencia a antibiótico o producto químico, que puede seleccionarse basándose en el efecto del gen marcador, es decir, la resistencia a un antibiótico, resistencia a un herbicida, marcadores colorimétricos, enzimas, marcadores fluorescentes y similares, en el que el efecto se usa para rastrear la herencia de un ácido nucleico de interés y/o para identificar una célula u organismo que ha heredado el ácido nucleico de interés. Los ejemplos de genes marcadores seleccionables conocidos y usados en la técnica incluyen: genes que proporcionan resistencia a ampicilina, estreptomina, gentamicina, kanamicina, higromicina, herbicida bialafos, sulfonamida, y similares; y genes que se usan como marcadores fenotípicos, es decir, genes reguladores de antocianina, gen de isopentanol transferasa, y similares.

El término "gen indicador" significa un ácido nucleico que codifica para un factor de identificación que puede identificarse basándose en el efecto del gen indicador, en el que el efecto se usa para rastrear la herencia de un ácido nucleico de interés, para identificar una célula u organismo que ha heredado el ácido nucleico de interés y/o para medir la inducción o transcripción de expresión génica. Ejemplos de genes indicadores conocidos y usados en la técnica incluyen: luciferasa (Luc), proteína fluorescente verde (GFP), cloranfenicol acetiltransferasa (CAT),  $\beta$ -galactosidasa (LacZ),  $\beta$ -glucuronidasa (Gus) y similares. Los genes marcadores seleccionables también pueden considerarse genes indicadores.

"Promotor" se refiere a una secuencia de ADN que puede controlar la expresión de una secuencia codificante o ARN funcional. En general, una secuencia codificante está ubicada en 3' con respecto a una secuencia de promotor. Los promotores pueden derivarse en su totalidad de un gen nativo, o pueden estar compuestos por elementos diferentes derivados de promotores diferentes encontrados en la naturaleza, o incluso comprender segmentos sintéticos de ADN. Los expertos en la técnica entienden que diferentes promotores pueden dirigir la expresión de un gen en

diferentes tejidos o tipos celulares, o en diferentes fases de desarrollo, o en respuesta a diferentes condiciones ambientales o fisiológicas. Los promotores que hacen que un gen se exprese en la mayoría de los tipos celulares en la mayoría de los casos, se denominan comúnmente “promotores constitutivos”. Los promotores que hacen que un gen se exprese en un tipo celular específico se denominan comúnmente “promotores específicos de célula” o “promotores específicos de tejido”. Los promotores que hacen que un gen se exprese en una fase específica de desarrollo o diferenciación celular se denominan comúnmente “promotores específicos de desarrollo” o “promotores específicos de diferenciación celular”. Los promotores que se inducen y hacen que un gen se exprese tras la exposición o el tratamiento de la célula con un agente, molécula biológica, producto químico, ligando, luz, o similar que induce el promotor se denominan comúnmente “promotores inducibles” o “promotores regulables”. Se reconoce además que puesto que en la mayoría de los casos no se han definido completamente los límites exactos de las secuencias reguladoras, fragmentos de ADN de diferentes longitudes pueden tener actividad promotora idéntica.

Una “secuencia de promotor” es una región reguladora de ADN que puede unirse a la ARN polimerasa en una célula e iniciar la transcripción de una secuencia codificante aguas abajo (en el sentido de 3'). Para los fines de definir la presente descripción, la secuencia de promotor se une en su extremo terminal 3' por el sitio de inicio de la transcripción y se extiende aguas arriba (en el sentido de 5') para incluir el número mínimo de bases o elementos necesarios para iniciar la transcripción a niveles detectables por encima del fondo. Dentro de la secuencia de promotor se encontrará un sitio de inicio de la transcripción (definido convenientemente por ejemplo, mapeando con nucleasa S1), así como dominios de unión a proteínas (secuencias consenso) responsables de la unión de ARN polimerasa.

Una secuencia codificante está “bajo el control” de secuencias control de la transcripción y la traducción en una célula cuando la ARN polimerasa transcribe la secuencia codificante para dar ARNm, que entonces se somete a corte y empalme trans de ARN (si la secuencia codificante contiene intrones) y se traduce para dar la proteína codificada por la secuencia codificante.

“Secuencias control de la transcripción y la traducción” son secuencias reguladoras de ADN, tales como promotores, potenciadores, terminadores, y similares, que proporcionan la expresión de una secuencia codificante en una célula huésped. En células eucariotas, las señales de poliadenilación son secuencias control.

El término “elemento de respuesta” significa uno o más elementos de ADN que actúan en cis que confieren capacidad de respuesta en un promotor mediada a través de la interacción con los dominios de unión a ADN del primer gen quimérico. Este elemento de ADN puede ser o bien palindrómico (perfecto o imperfecto) en su secuencia o bien puede estar compuesto de por motivos de secuencia o semisitios separados por un número variable de nucleótidos. Los semisitios pueden ser similares o idénticos y estar dispuestos como repeticiones o bien directas o bien invertidas o como un único semisito o múltiplos de semisitios adyacentes en tándem. El elemento de respuesta puede comprender un promotor mínimo aislado de diferentes organismos dependiendo de la naturaleza de la célula u organismo en que se incorporará el elemento de respuesta. El dominio de unión a ADN de la primera proteína híbrida se une, en presencia o ausencia de un ligando, a la secuencia de ADN de un elemento de respuesta para iniciar o suprimir la transcripción de gen o genes aguas abajo bajo la regulación de este elemento de respuesta. Los ejemplos de secuencias de ADN para elementos de respuesta del receptor de ecdisona natural incluyen: RRGG/TTCANTGAC/ACY (véase Cherbas L., *et. al.*, (1991), *Genes Dev.* 5, 120-131); AGGTCAN<sub>(n)</sub>AGGTCA, donde N<sub>(n)</sub> puede ser uno o más nucleótidos espaciadores (véase D'Avino PP., *et. al.*, (1995), *Mol. Cell. Endocrinol.* 113, 1-9); y GGGTTGAATGAATTT (véase Antoniewski C., *et. al.*, (1994). *Mol. Cell Biol.* 14,4465-4474).

El término “unido operativamente” se refiere a la asociación de secuencias de ácido nucleico en un único fragmento de ácido nucleico de modo que la función de una resulta afectada por la otra. Por ejemplo, un promotor está unido operativamente con una secuencia codificante cuando puede afectar a la expresión de esa secuencia codificante (es decir, que la secuencia codificante está bajo el control de transcripción del promotor). Las secuencias codificantes pueden unirse operativamente a secuencias reguladoras en orientación sentido o antisentido.

El término “expresión”, tal como se usa en el presente documento, se refiere a la transcripción y acumulación estable de ARN sentido (ARNm) o antisentido derivado de un ácido nucleico o polinucleótido. Expresión también puede referirse a la traducción del ARNm para dar una proteína o polipéptido.

Los términos “casete”, “casete de expresión” y “casete de expresión génica” se refieren a un segmento de ADN que puede insertarse en un ácido nucleico o polinucleótido en sitios de restricción específicos o mediante recombinación homóloga. El segmento de ADN comprende un polinucleótido que codifica para un polipéptido de interés, y el casete y los sitios de restricción se diseñan para garantizar la inserción del casete en el marco de lectura apropiado para la transcripción y la traducción. “Casete de transformación” se refiere a un vector específico que comprende un polinucleótido que codifica para un polipéptido de interés y que tiene elementos además del polinucleótido que facilitan la transformación de una célula huésped particular. Los casetes, casetes de expresión, casetes de expresión génica y casetes de transformación de la descripción también pueden comprender elementos que permiten la expresión potenciada de un polinucleótido que codifica para un polipéptido de interés en una célula huésped. Estos elementos pueden incluir, pero no se limitan a: un promotor, a promotor mínimo, un potenciador, un elemento de respuesta, una secuencia terminadora, una secuencia de poliadenilación, y similares.

Para los fines de esta invención, el término “interruptor génico” se refiere a la combinación de un elemento de respuesta asociado con un promotor, y un sistema basado en EcR que en presencia de uno o más ligandos, modula la expresión de un gen en el que se incorporan el elemento de respuesta y el promotor.

5 Los términos “modulan” y “modula” significan inducir, reducir o inhibir la expresión génica o de ácido nucleico; dando como resultado la inducción, reducción o inhibición respectivas de la producción de proteínas o polipéptidos.

10 Los plásmidos o vectores según la descripción pueden comprender además al menos un promotor adecuado para dirigir la expresión de un gen en una célula huésped. El término “vector de expresión” significa un vector, plásmido o vehículo diseñado para permitir la expresión de una secuencia de ácido nucleico insertada tras la transformación en el huésped. El gen clonado, es decir, la secuencia de ácido nucleico insertada, se coloca habitualmente bajo el control de elementos de control tales como un promotor, un promotor mínimo, un potenciador, o similares. Los promotores o regiones de control del inicio, que son útiles para dirigir la expresión de un ácido nucleico en la célula huésped deseada son numerosos y resultan familiares para los expertos en la técnica. Prácticamente cualquier promotor que puede dirigir estos genes es adecuado para la presente descripción incluyendo, pero sin limitarse a: promotores virales, promotores bacterianos, promotores de animales, promotores de mamíferos, promotores sintéticos, promotores constitutivos, promotor específico de tejido, promotores específicos de desarrollo, promotores inducibles, promotores regulados por luz; *CYC1*, *HIS3*, *GAL1*, *GAL4*, *GAL10*, *ADH1*, *PGK*, *PHO5*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO*, *TPI*, promotores de fosfatasa alcalina (útiles para la expresión en *Saccharomyces*); promotor *AOX1* (útil para la expresión en *Pichia*); promotores de  $\beta$ -lactamasa, *lac*, *ara*, *tet*, *trp*, *IP<sub>L</sub>*, *IP<sub>R</sub>*, *T7*, *tac* y *trc* (útiles para la expresión en *Escherichia coli*); promotores regulados por luz, específicos de semillas, específicos de polen, específicos de ovario, relacionados con patogenicidad o enfermedad, 35S del virus del mosaico de la coliflor, mínimo 35S de CMV, del virus del mosaico de las nervaduras de mandioca (CsVMV), de la proteína de unión a clorofila a/b, de ribulosa 1, 5-bisfosfato carboxilasa, específicos de brote, específicos de raíz, de quitinasa, inducibles por estrés, del virus baciliforme del tungro del arroz, super-promotor de plantas, de leucina aminopeptidasa de patata, de nitrato reductasa, de manopina sintasa, de nopalina sintasa, de ubiquitina, de proteína zeína y de antocianina (útiles para la expresión en células vegetales); los promotores de animales y mamíferos conocidos en la técnica incluyen, pero no se limitan a, la región promotora temprana del SV40 (SV40e), el promotor contenido en la repetición terminal larga (LTR) en 3' del virus del sarcoma de Rous (RSV), de los genes E1A del promotor tardío mayor (MLP) de adenovirus (Ad), el promotor temprano de citomegalovirus (CMV), el promotor de timidina cinasa (TK) del virus del herpes simple (VHS), un promotor IE1 de baculovirus, un promotor del factor de elongación 1 alfa (EF1), un promotor de fosfoglicerato cinasa (PGK), un promotor de ubiquitina (Ubc), un promotor de albúmina, las secuencias reguladoras del promotor de L-metalotioneína de ratón y regiones de control de la transcripción, los promotores ubicuos (HPRT, vimentina,  $\alpha$ -actina, tubulina y similares), los promotores de los filamentos intermedios (desmina, neurofilamentos, queratina, GFAP, y similares), los promotores de genes terapéuticos (del tipo de MDR, CFTR o factor VIII y similares), promotores relacionados con patogenicidad o enfermedad y promotores que muestran especificidad de tejido y que se han usando en animales transgénicos, tales como la región de control del gen de la elastasa I que es activo en células acinares pancreáticas; la región de control del gen de insulina activo en células beta pancreáticas, la región de control del gen de inmunoglobulina activa en células linfoides, la región de control del virus del cáncer de mama de ratón activa en células testiculares, de mama, linfoides y mastocitos; regiones de control del gen de albúmina, Apo AI y Apo AII activas en hígado, región de control del gen de alfa-fetoproteína activa en hígado, región de control del gen de alfa 1-antitripsina activa en hígado, región de control del gen de beta-globina activa en células mieloides, región de control del gen de la proteína básica de mielina activa en oligodendrocitos en el cerebro, región de control del gen de la cadena ligera 2 de miosina activa en músculo esquelético y región de control del gen de la hormona liberadora de gonadotropina activa en hipotálamo, promotor de piruvato cinasa, promotor de vilina, promotor de la proteína intestinal de unión a ácidos grasos, promotor de  $\alpha$ -actina de células de músculo liso y similares. Además, estas secuencias de expresión pueden modificarse mediante la adición de secuencias potenciadoras o reguladoras y similares.

50 Los potenciadores que pueden usarse en las realizaciones de la descripción incluyen pero no se limitan a: un potenciador de SV40, un potenciador de citomegalovirus (CMV), un factor de elongación 1 (EF1), potenciadores de levaduras, potenciadores de genes virales, y similares.

55 Las regiones de control de la terminación, es decir, secuencias terminadoras o de poliadenilación, también pueden derivarse de diversos genes nativos para los huéspedes preferidos. Opcionalmente, un sitio de terminación puede ser innecesario, sin embargo, es más preferido si está incluido. En una realización preferida de la descripción, la región de control de la terminación puede comprender o derivarse de una secuencia sintética, señal de poliadenilación sintética, una señal de poliadenilación tardía de SV40, una señal de poliadenilación de SV40, una señal de poliadenilación de hormona de crecimiento bovino (BGH), secuencias terminadoras virales, o similares.

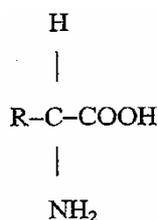
65 Los términos “secuencias no codificantes en 3'” o “región no traducida en 3' (UTR)” se refieren a secuencias de ADN ubicadas aguas abajo (3') de una secuencia codificante y pueden comprender secuencias de reconocimiento de poliadenilación [poli(A)] y otras secuencias que codifican para señales reguladoras que pueden afectar al procesamiento del ARNm o a la expresión génica. La señal de poliadenilación se caracteriza habitualmente por afectar a la adición de tramos de poli(ácido adenílico) al extremo 3' del precursor de ARNm.

“Región reguladora” significa una secuencia de ácido nucleico que regula la expresión de una segunda secuencia de ácido nucleico. Una región reguladora puede incluir secuencias que son responsables de manera natural de la expresión de un ácido nucleico particular (una región homóloga) o pueden incluir secuencias de un origen diferente que son responsables de la expresión de diferentes proteínas o incluso proteínas sintéticas (una región heteróloga). En particular, las secuencias pueden ser secuencias de genes de procariota, eucariota o virales o secuencias derivadas que estimulan o reprimen la transcripción de un gen de una manera específica o no específica y de una manera inducible o no inducible. Las regiones reguladoras incluyen orígenes de replicación, sitios de corte y empalme de ARN, promotores, potenciadores, secuencias de terminación de la transcripción y secuencias señal que dirigen el polipéptido a las rutas secretoras de la célula diana.

Una región reguladora de una “fuente heteróloga” es una región reguladora que no está asociada de manera natural con el ácido nucleico expresado. Incluidas entre las regiones reguladoras heterólogas están regiones reguladoras de una especie diferente, regiones reguladoras de un gen diferente, secuencias reguladoras híbridas y secuencias reguladoras que no se producen en la naturaleza, pero que están diseñadas por un experto en la técnica.

“Transcrito de ARN” se refiere al producto que resulta de la transcripción catalizada por ARN polimerasa de una secuencia de ADN. Cuando el transcrito de ARN es una copia complementaria perfecta de la secuencia de ADN, se denomina el transcrito primario o puede ser una secuencia de ARN derivada de procesamiento post-transcripcional del transcrito primario y se denomina el ARN maduro. “ARN mensajero (ARNm)” se refiere al ARN que está sin intrones y que puede traducirse en una proteína por la célula. “ADNc” se refiere a un ADN bicatenario que es complementario a y se deriva de ARNm. ARN “sentido” se refiere un transcrito de ARN que incluye el ARNm y de este modo puede traducirse en una proteína por la célula. “ARN antisentido” se refiere a un transcrito de ARN que es complementario a todo o parte de un transcrito primario o ARNm diana y que bloquea la expresión de un gen diana. La complementariedad de un ARN antisentido puede ser con cualquier parte del transcrito génico específico, es decir, en la secuencia no codificante en 5', secuencia no codificante en 3' o la secuencia codificante. “ARN funcional” se refiere a ARN antisentido, ARN de ribozimas u otro ARN que no se traduce y aún tiene un efecto en procesos celulares.

Un “polipéptido” es un compuesto polimérico compuesto de residuos de aminoácidos unidos covalentemente. Los aminoácidos tienen la siguiente estructura general:



Los aminoácidos se clasifican en siete grupos basándose en la cadena lateral R: (1) cadenas laterales alifáticas, (2) cadenas laterales que contienen un grupo hidroxílico (OH), (3) cadenas laterales que contienen átomos de azufre, (4) cadenas laterales que contienen un grupo ácido o amida, (5) cadenas laterales que contienen un grupo básico, (6) cadenas laterales que contienen un anillo aromático y (7) prolina, un iminoácido en el que la cadena lateral está fusionada al grupo amino. Un polipéptido de la descripción comprende preferiblemente al menos aproximadamente 14 aminoácidos.

Una “proteína” es un polipéptido que realiza un papel estructural o funcional en una célula viva.

Un “polipéptido aislado” o “proteína aislada” es un polipéptido o proteína que está sustancialmente libre de los compuestos que están asociados normalmente con ellos en su estado natural (por ejemplo, otras proteínas o polipéptidos, ácidos nucleicos, hidratos de carbono, lípidos). “Aislado” no pretende excluir mezclas artificiales o sintéticas con otros compuestos, ni la presencia de impurezas que no interfieren con la actividad biológica, y que pueden estar presentes, por ejemplo, debido a purificación incompleta, adición de estabilizadores o composición dando lugar a una preparación farmacéuticamente aceptable.

Un “polipéptido mutante de sustitución” o un “mutante de sustitución” se entenderá que significa un polipéptido mutante que comprende una sustitución de al menos un (1) aminoácido de tipo natural o que se produce de manera natural con un aminoácido diferente con respecto al polipéptido de tipo natural o que se produce de manera natural. Un polipéptido mutante de sustitución puede comprender sólo una (1) sustitución de aminoácido de tipo natural o que se produce de manera natural y puede denominarse un polipéptido “mutante puntual” o “mutante puntual individual”. Alternativamente, un polipéptido mutante de sustitución puede comprender una sustitución de dos (2) o más aminoácidos de tipo natural o que se producen de manera natural con 2 o más aminoácidos con respecto al polipéptido de tipo natural o que se produce de manera natural. Según la descripción, un polipéptido del dominio de unión a ligando de receptores nucleares de grupo H que comprende una mutación de sustitución comprende una

sustitución de al menos un (1) aminoácido de tipo natural o que se produce de manera natural con un aminoácido diferente con respecto al polipéptido del dominio de unión a ligando de receptores nucleares de grupo H de tipo natural o que se produce de manera natural.

- 5 En el caso en el que el polipéptido mutante de sustitución comprenda una sustitución de dos (2) o más aminoácidos de tipo natural o que se producen de manera natural, esta sustitución puede comprender o bien un número equivalente de aminoácidos de tipo natural o que se producen de manera natural delecionados para la sustitución, es decir, 2 aminoácidos de tipo natural o que se producen de manera natural sustituidos por 2 aminoácidos de tipo no natural o que se no producen de manera natural, o un número no equivalente de aminoácidos de tipo natural delecionados para la sustitución, es decir, 2 aminoácidos de tipo natural sustituidos por 1 aminoácido de tipo no natural (una mutación de sustitución + deleción), o 2 aminoácidos de tipo natural sustituidos por 3 aminoácidos de tipo no natural (una mutación de sustitución + inserción).

15 Los mutantes de sustitución pueden describirse usando un sistema de nomenclatura abreviada para indicar el número y residuo de aminoácido sustituido dentro de la secuencia de polipéptido de referencia y el nuevo residuo de aminoácido sustituido. Por ejemplo, un mutante de sustitución en el que el vigésimo (20<sup>º</sup>) residuo de aminoácido de un polipéptido está sustituido puede abreviarse como "x20z", en el que "x" es el aminoácido que está sustituido, "20" es la posición o el número de residuo de aminoácido dentro del polipéptido y "z" es el nuevo aminoácido sustituido. Por tanto, un mutante de sustitución abreviado de manera intercambiable como "E20A" o "Glu20Ala" indica que el mutante comprende un residuo alanina (abreviado comúnmente en la técnica como "A" o "Ala") en lugar del ácido glutámico (abreviado comúnmente en la técnica como "E" o "Glu") en la posición 20 del polipéptido.

25 Una mutación de sustitución puede realizarse mediante cualquier técnica de mutagénesis conocida en la técnica, incluyendo pero sin limitarse a, mutagénesis dirigida al sitio *in vitro* (Hutchinson, C., *et al.*, 1978, J. Biol. Chem. 253: 6551; Zoller y Smith, 1984, DNA 3: 479-488; Oliphant *et al.*, 1986, Gene 44: 177; Hutchinson *et al.*, 1986, Proc. Natl. Acad. Sci. U.S.A. 83: 710), uso de ligadores TAB<sup>®</sup> (Pharmacia), deleción y sustitución de fragmento/digestión con endonucleasa de restricción, mutagénesis mediada por PCR/dirigida por oligonucleótido, y similares. Se prefieren las técnicas basadas en PCR para la mutagénesis dirigida al sitio (véase Higuchi, 1989, "Using PCR to Engineer DNA", en PCR Technology: Principles and Applications for DNA Amplification, H. Erlich, ed., Stockton Press, capítulo 6, pág. 61-70).

35 "Fragmento" de un polipéptido según la descripción se entenderá que significa un polipéptido cuya secuencia de aminoácidos es más corta que la del polipéptido de referencia y que comprende, con respecto a la parte completa con estos polipéptidos de referencia, una secuencia de aminoácidos idéntica. Tales fragmentos pueden incluirse, cuando sea apropiado, en un polipéptido más grande del que forman parte. Tales fragmentos de un polipéptido según la descripción pueden tener una longitud de al menos 2, 3, 4, 5, 6, 8, 10, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 25, 26, 30, 35, 40, 45, 50, 100, 200, 240 ó 300 aminoácidos.

40 Una "variante" de un polipéptido o proteína es cualquier análogo, fragmento, derivado o mutante que se deriva de un polipéptido o proteína y que conserva al menos una propiedad biológica del polipéptido o proteína. En la naturaleza pueden existir diferentes variantes del polipéptido o proteína. Estas variantes pueden ser variaciones alélicas caracterizadas por diferencias en la secuencia de nucleótidos del gen estructural que codifica para la proteína, o puede implicar una modificación postraduccional o de corte y empalme diferencial. El experto en la técnica puede producir variantes que tienen sustituciones, deleciones, adiciones o remplazos de aminoácido individuales o múltiples. Estas variantes pueden incluir, entre otras: (a) variantes en que uno o más residuos de aminoácido están sustituidos con aminoácidos conservativos o no conservativos, (b) variantes en que se añaden uno o más aminoácidos al polipéptido o proteína, (c) variantes en que uno o más de los aminoácidos incluyen un grupo sustituyente y (d) variantes en que el polipéptido o proteína está fusionado con otro polipéptido tal como albúmina sérica. El experto habitual en la técnica conoce las técnicas para obtener estas variantes, incluyendo técnicas genéticas (supresiones, deleciones, mutaciones, etc.), químicas y enzimáticas. Un polipéptido variante comprende preferiblemente al menos aproximadamente 14 aminoácidos.

Una "proteína heteróloga" se refiere a una proteína que no se produce de manera natural en la célula.

55 Una "proteína madura" se refiere a un polipéptido que se procesa postraduccionalmente; es decir, uno del que se ha eliminado cualquier pre o propéptido presente en el producto de traducción primario. Proteína "precursora" se refiere al producto primario de traducción del ARNm; es decir, con pre y propéptidos todavía presentes. Los pre y propéptidos pueden ser, pero no se limitan a señales de localización intracelulares.

60 El término "péptido señal" se refiere a un polipéptido amino terminal que precede a la proteína madura secretada. El péptido señal se escinde de y por tanto no está presente en la proteína madura. Los péptidos señal tiene la función de dirigir y traslocar proteínas secretadas a través de las membranas celulares. El péptido señal también se denomina proteína señal.

65 Una "secuencia señal" está incluida al comienzo de la secuencia codificante de una proteína que va a expresarse en la superficie de una célula. Esta secuencia codifica para un péptido señal, en el extremo N-terminal con respecto al

polipéptido maduro, que dirige la célula huésped para traslocar el polipéptido. El término “secuencia señal de traslocación” se usa en el presente documento para referirse a este tipo de secuencia señal. Las secuencias señal de traslocación pueden encontrarse asociadas con una variedad de proteínas nativas a eucariotas y procariotas, y a menudo son funcionales en ambos tipos de organismos.

5 El término “homología” se refiere al porcentaje de identidad entre restos de dos polinucleótidos o dos polipéptidos. La correspondencia entre la secuencia de un resto a otro puede determinarse mediante técnicas conocidas en la técnica. Por ejemplo, puede determinarse la homología mediante una comparación directa de la información de secuencia entre dos moléculas de polipéptido alineando la información de secuencia y usando programas informáticos fácilmente disponibles. Alternativamente, puede determinarse la homología mediante hibridación de polinucleótidos en condiciones que forma dúplex estables entre regiones homólogas, seguido por digestión con nucleasa(s) específica(s) monocatenaria(s) y determinación del tamaño de los fragmentos digeridos.

15 Tal como se usa en el presente documento, el término “homólogo” en todas sus formas gramaticales y variaciones ortográficas se refiere a la relación entre proteínas que presentan un “origen evolutivo común”, incluyendo proteínas de superfamilias (por ejemplo, la superfamilia de inmunoglobulinas) y proteínas homólogas de diferentes especies (por ejemplo, cadena ligera de miosina, etc.) (Reeck *et al.*, 1987, Cell 50:667.). Tales proteínas (y sus genes codificantes) tienen homología de secuencia, tal como se refleja por su alto grado de similitud de secuencia. Sin embargo, en el uso común y en la presente solicitud, el término “homólogo,” cuando está modificado con un adverbio tal como “altamente”, puede referirse a similitud de secuencia y no a un origen evolutivo común.

20 Por consiguiente, el término “similitud de secuencia” en todas sus formas gramaticales se refiere al grado de identidad o correspondencia entre secuencias de ácidos nucleicos o aminoácidos de proteínas que pueden compartir o no un origen evolutivo común (véase Reeck *et al.*, 1987, Cell 50: 667).

25 En una realización específica, dos secuencia de ADN son “sustancialmente homólogas” o “sustancialmente similares” cuando al menos aproximadamente el 50% (preferiblemente al menos aproximadamente el 75%, y lo más preferiblemente al menos aproximadamente el 90 o el 95%) de los nucleótidos coincide a lo largo de la longitud definida de las secuencias de ADN. Las secuencias que son sustancialmente homólogas pueden identificarse comparando las secuencias usando software convencional disponible en bancos de datos de secuencias, o en un experimento de hibridación de tipo Southern en, por ejemplo, condiciones rigurosas definidas para ese sistema particular. Definir las condiciones de hibridación apropiadas está dentro del conocimiento de la técnica. Véase, por ejemplo, Sambrook *et al.*, 1989, citado anteriormente.

35 Tal como se usa en el presente documento, “sustancialmente similar” se refiere a fragmentos de ácido nucleico en los que cambios en una o más bases nucleotídicas dan como resultado la sustitución de uno o más aminoácidos, pero no afecta a las propiedades funcionales de la proteína codificada por la secuencia de ADN. “Sustancialmente similar” también se refiere a fragmentos de ácido nucleico en los que cambios en una o más bases nucleotídicas no afectan a la capacidad del fragmento de ácido nucleico para mediar la alteración de la expresión génica mediante tecnología antisentido o cosupresión. “Sustancialmente similar” también se refiere a modificaciones de los fragmentos de ácido nucleico de la presente descripción tal como delección o inserción de una o más bases nucleotídicas que no afectan sustancialmente a las propiedades funcionales del transcrito resultante. Por tanto, se entiende que la descripción abarca más que las secuencias a modo de ejemplo específicas. Cada una de las modificaciones propuestas está completamente dentro del conocimiento de rutina en la técnica, como es la determinación de la retención de la actividad biológica de los productos codificados.

45 Además, el experto en la técnica reconoce que las secuencias sustancialmente similares abarcadas por esta descripción también se definen por su capacidad para hibridarse, en condiciones rigurosas (0,1X SSC, SDS al 0,1%, 65°C y lavado con 2X SSC, SDS al 0,1% seguido por 0,1X SSC, SDS al 0,1%), con las secuencias a modo de ejemplo en el presente documento. Los fragmentos de ácido nucleico sustancialmente similares de la presente descripción son los fragmentos de ácido nucleico cuyas secuencias de ADN son idénticas al menos en el 70% a la secuencia de ADN de los fragmentos de ácido nucleico notificados en el presente documento. Los fragmentos de ácido nucleico sustancialmente preferidos de la presente descripción son los fragmentos de ácido nucleico cuyas secuencias de ADN son idénticas al menos en el 80% a la secuencia de ADN de los fragmentos de ácido nucleico notificados en el presente documento. Fragmentos de ácido nucleico más preferidos son idénticos al menos en el 90% a la secuencia de ADN de los fragmentos de ácido nucleico notificados en el presente documento. Incluso más preferidos son fragmentos de ácido nucleico que son idénticos al menos en el 95% a la secuencia de ADN de los fragmentos de ácido nucleico notificados en el presente documento.

60 Dos secuencias de aminoácidos son “sustancialmente homólogas” o “sustancialmente similares” cuando más de aproximadamente el 40% de los aminoácidos son idénticos, o más del 60% son similares (funcionalmente idénticos). Preferiblemente, las secuencias similares u homólogas se identifican mediante alineación usando, por ejemplo, el programa Pileup de GCG (Genetics Computer Group, Program Manual for the GCG Package, versión 7, Madison, Wisconsin).

65 El término “se corresponde con” se usa en el presente documento para referirse a secuencias similares u

homólogas, ya sea la posición exacta idéntica o diferente de la molécula de la que se mide la similitud u homología. Una alineación de secuencia de aminoácidos o ácido nucleico puede incluir espacios. Por tanto, el término “se corresponde con” se refiere a la similitud de secuencia y no a la numeración de los residuos de aminoácido o bases nucleotídicas.

5 Una “parte sustancial” de una secuencia de aminoácidos o nucleótidos comprende suficiente cantidad de la secuencia de aminoácidos de un polipéptido o la secuencia de nucleótidos de un gen para identificar supuestamente ese polipéptido o gen, o bien mediante evaluación manual de la secuencia por un experto en la técnica, o bien mediante comparación e identificación de secuencia automatizadas por ordenador usando algoritmos tales como  
 10 BLAST (Basic Local Alignment Search Tool; Altschul, S. F., *et al.*, (1993) J. Mol. Biol. 215: 403-410; véase también [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)). En general, se necesita una secuencia de diez o más aminoácidos contiguos o treinta o más nucleótidos con el fin de identificar supuestamente una secuencia de polipéptido o ácido nucleico como homóloga a una proteína o gen conocidos. Además, con respecto a las secuencias de nucleótidos, pueden usarse sondas de oligonucleótido específicas de genes que comprenden 20-30 nucleótidos contiguos en métodos de  
 15 identificación de genes dependientes de secuencia (por ejemplo, hibridación de tipo Southern) y aislamiento (por ejemplo, hibridación *in situ* de colonias bacterianas o placas de bacteriófago). Además, pueden usarse oligonucleótidos cortos de 12-15 bases como cebadores de amplificación en PCR con el fin de obtener un fragmento de ácido nucleico particular que comprende los cebadores. Por consiguiente, una “parte sustancial” de una secuencia de nucleótido comprende suficiente cantidad de la secuencia para identificar específicamente y/o aislar un  
 20 fragmento de ácido nucleico que comprende la secuencia.

El término “porcentaje de identidad”, tal como se conoce en la técnica, es una relación entre dos o más secuencias de polipéptidos o dos o más secuencias de polinucleótidos, tal como se determina comparando las secuencias. En la  
 25 técnica, “identidad” también significa el grado de relación de secuencia entre secuencias de polipéptidos o polinucleótidos, según el caso, tal como se determina mediante la coincidencia entre sucesiones de tales secuencias. “Identidad” y “similitud” pueden calcularse fácilmente mediante métodos conocidos, incluyendo pero sin limitarse a los descritos en: Computational Molecular Biology (Lesk, A. M., ed.) Oxford University Press, Nueva York (1988); Biocomputing: Informatics and Genome Projects (Smith, D. W., ed.) Academic Press, Nueva York (1993); Computer Analysis of Sequence Data, Part I (Griffin, A. M., y Griffin, H. G., eds.) Humana Press, Nueva Jersey (1994); Sequence Analysis in Molecular Biology (von Heinje, G., ed.) Academic Press (1987); y Sequence Analyses  
 30 Primer (Gribskov, M. y Devereux, J., eds.) Stockton Press, Nueva York (1991). Los métodos preferidos para determinar la identidad se diseñan para dar la mejor coincidencia entre las secuencias sometidas a prueba. Los métodos para determinar la identidad y similitud están codificados en programas informáticos disponibles públicamente. Las alineaciones de secuencia y cálculos del porcentaje de identidad pueden realizarse usando el programa Megalign del conjunto de computación bioinformático LASERGENE (DNASTAR Inc., Madison, WI). Puede realizarse la alineación múltiple de las secuencias usando el método Clustal de alineación (Higgins y Sharp (1989) CABIOS. 5:151-153) con los parámetros por defecto (PENALIZACIÓN POR HUECO=10, PENALIZACIÓN POR LONGITUD DE HUECO=10). Pueden seleccionarse los parámetros por defecto para alineaciones en parejas usando el método Clustal: KTUPLE 1, PENALIZACIÓN POR HUECO=3, VENTANA=5 y DIAGONALES GUARDADAS=5.

40 El término “software de análisis de secuencia” se refiere a cualquier programa de software o algoritmo informático que es útil para el análisis de secuencias de nucleótidos o aminoácidos. El “software de análisis de secuencia” puede estar disponible comercialmente o puede desarrollarse independientemente. El software de análisis de secuencia típico incluirá pero no se limitará al conjunto de programas de GCG (Wisconsin Package Versión 9.0, Genetics Computer Group (GCG), Madison, WI), BLASTP, BLASTN, BLASTX (Altschul *et al.*, J. Mol. Biol. 215:403-410 (1990) y DNASTAR (DNASTAR, Inc. 1228 S. Park St. Madison, WI 53715 EE.UU). Dentro del contexto de esta solicitud se entenderá que cuando se usa software de análisis de secuencia para el análisis, los resultados del análisis se basarán en los “valores por defecto” del programa de referencia, a menos que se especifique lo contrario. Tal como se usa en el presente documento “valores por defecto” significará cualquier conjunto de valores o  
 50 parámetros que originalmente se cargan con el software cuando se inicia por primera vez.

Los “genes sintéticos” pueden ensamblarse a partir de elementos estructurales de oligonucleótidos que se sintetizan químicamente usando procedimientos conocidos por los expertos en la técnica. Estos elementos estructurales se ligan y se aparean para formar segmentos de genes que entonces se ensamblan enzimáticamente para construir el  
 55 gen completo. “Sintetizado químicamente”, en lo que se refiere a una secuencia de ADN, significa que los nucleótidos componentes se ensamblaron *in vitro*. La síntesis química manual de ADN puede llevarse a cabo usando procedimientos bien establecidos, o puede realizarse la síntesis química automatizada usando una de varias de máquinas disponibles comercialmente. Por consiguiente, los genes pueden adaptarse para la expresión génica óptima basándose en la optimización de la secuencia de nucleótidos para reflejar el desplazamiento del codón de la célula huésped. El experto en la técnica aprecia la posibilidad de expresión génica satisfactoria si el uso del codón se desplaza hacia los codones favorecidos por el huésped. La determinación de codones preferidos puede basarse en una supervivencia de genes derivados de la célula huésped en la que está disponible la información de  
 60 secuencia.

65 Tal como se usa en el presente documento, se dice que dos o más sistemas de regulación génica operables individualmente son “ortogonales” cuando; a) la modulación de cada uno de los sistemas dados por su ligando

respectivo, a una concentración elegida, da como resultado un cambio medible en la magnitud de expresión del gen de ese sistema, y b) el cambio es diferente de manera estadísticamente significativa del cambio en la expresión del resto de los sistemas operables simultáneamente en la célula, tejido u organismo, independientemente de la simultaneidad o secuencialidad de la modulación real. Preferiblemente, la modulación de cada sistema de regulación génica operable individualmente efectúa un cambio en la expresión génica al menos 2 veces mayor que el resto de los sistemas operables en la célula, tejido u organismo. Más preferiblemente, el cambio es al menos 5 veces mayor. Incluso más preferiblemente, el cambio es al menos 10 veces mayor. Todavía más preferiblemente, el cambio es al menos 100 veces mayor. Incluso todavía más preferiblemente, el cambio es al menos 500 veces mayor. Idealmente, la modulación de cada uno de los sistemas dados por su ligando respectivo a una concentración elegida da como resultado un cambio medible en la magnitud de expresión del gen de ese sistema y ningún cambio medible en la expresión del resto de los sistemas operables en la célula, tejido u organismo. En tales casos se dice que los múltiples sistemas de regulación génica inducibles son “completamente ortogonales”. La presente descripción es útil para buscar ligandos ortogonales y sistemas de expresión génica basados en receptores ortogonales tales como los descritos en la solicitud estadounidense de patente en tramitación junto con la presente 09/965.697.

El término “modular” significa la capacidad de un complejo ligando/receptor dado para inducir o suprimir la transactivación de un gen exógeno.

El término “gen exógeno” significa un gen foráneo para el sujeto, es decir, un gen que se introduce en el sujeto a través de un procedimiento de transformación, una versión no mutada de un gen mutado endógeno o una versión mutada de un gen no mutado endógeno. El método de transformación no es crítico para esta descripción y puede ser cualquier método adecuado para el objeto conocido en la técnica. Por ejemplo, se obtienen plantas transgénicas mediante regeneración de las células transformadas. A partir de la bibliografía se conocen numerosos procedimientos de transformación tales como agroinfección usando *Agrobacterium tumefaciens* o su plásmido T<sub>1</sub>, electroporación, microinyección de protoplastos y células vegetales, y transformación mediante microproyectiles. Se conocen técnicas complementarias para la transformación de células animales y la regeneración de tales células transformadas en animales transgénicos. Los genes exógenos pueden ser genes o bien naturales o bien sintéticos y genes terapéuticos que se introducen en el sujeto en forma de ADN o ARN que pueden funcionar a través de un producto intermedio de ADN tal como mediante transcriptasa inversa. Tales genes pueden introducirse en células diana, introducirse directamente en el sujeto o introducirse indirectamente mediante la transferencia de células transformadas en el sujeto. El término “gen terapéutico” significa un gen que confiere una función beneficiosa a la célula huésped en la que se expresa tal gen. Los genes terapéuticos no se encuentran de manera natural en células huésped.

El término “complejo de receptor de ecdisona” se refiere en general a un complejo de proteína heterodimérico que consiste en dos miembros de la familia de receptores de esteroides, receptor de ecdisona (“EcR”) y proteínas ultraespiráculo (“USP”) (véase Yao, T.P., *et. al.* (1993) *Nature* 366, 476-479; Yao, T.-P., *et. al.*, (1992) *Cell* 71, 63-72). El complejo de receptor ecdiesteroide funcional también puede incluir proteína(s) adicional(es) tal(es) como inmunofilinas. Los miembros adicionales de la familia de receptores de esteroides de proteínas, conocidos como factores de transcripción (tal como DHR38, *betaFTZ-1* u otros homólogos de insecto), también pueden ser parejas dependientes o independientes de ligando para EcR y/o USP. El complejo de receptor de ecdisona también puede ser un heterodímero de la proteína de receptor de ecdisona y el homólogo de vertebrados de proteína ultraespiráculo, proteína de receptor X de ácido retinoico (“RXR”). Los complejos homodiméricos de la proteína de receptor de ecdisona o USP también pueden ser funcionales en algunas circunstancias.

Un complejo de receptor ecdiesteroide puede activarse mediante un ecdiesteroide o ligando no esteroideo activo unido a una de las proteínas del complejo, incluso de EcR, pero sin excluir otras proteínas del complejo.

El complejo de receptor de ecdisona incluye proteínas que son miembros de la superfamilia de receptores esteroides en el que todos los miembros se caracterizan por la presencia de un dominio de transactivación aminoterminal, un dominio de unión a ADN (“DBD”) y un dominio de unión a ligando (“LBD”) separados por una región bisagra. Algunos miembros de la familia también pueden tener otro dominio de transactivación en el lado carboxiloterminal del LBD. El DBD se caracteriza por la presencia de dos dedos de zinc-cisteína entre los que hay dos motivos de aminoácido, la caja P y la caja D, que confieren especificidad para elementos de respuesta a ecdisona. Estos dominios pueden ser o bien nativos o bien modificados o bien quimeras de diferentes dominios de proteínas de receptor heterólogos.

Las secuencias de ADN que componen el gen exógeno, el elemento de respuesta y el complejo de receptor de ecdisona pueden incorporarse en archaeobacterias, células procariotas tales como *Escherichia coli*, *Bacillus subtilis* u otras enterobacterias, o células eucariotas tales como células vegetales o animales. Sin embargo, debido a que muchas de las proteínas expresadas por el gen se procesan incorrectamente en bacterias, se prefieren las células eucariotas. Las células pueden estar en forma de células individuales u organismos multicelulares. Las secuencias de nucleótidos para el gen exógeno, el elemento de respuesta y el complejo de receptor también pueden incorporarse como moléculas de ARN, preferiblemente en forma de ARN virales funcionales tales como el virus del mosaico de tabaco. De la células eucariotas, se prefieren las células de vertebrados debido a que carecen de manera natural de las moléculas que confieren respuestas a los ligandos de esta descripción para el receptor de

ecdisona. Como resultado, son insensibles a los ligandos de esta descripción. Por tanto, los ligandos de esta descripción tendrán efectos fisiológicos u otros efectos desdeñables sobre las células transformadas o el organismo completo. Por tanto, las células pueden crecer y expresar el producto deseado, sustancialmente sin verse afectadas por la presencia del propio ligando.

5 El término "sujeto" significa una planta o un animal intactos o una célula de una planta o un animal. También se prevé que los ligandos funcionarán igualmente bien cuando el sujeto es un hongo o una levadura. Cuando el sujeto es un animal intacto, preferiblemente el animal es un vertebrado, lo más preferiblemente un mamífero.

10 Los ligandos de la presente descripción, cuando se usan con el complejo de receptor de ecdisona que a su vez está unido al elemento de respuesta unido a un gen exógeno, proporcionan los medios para la regulación temporal externa de la expresión del gen exógeno. El orden en el que los diversos componentes se unen entre sí, es decir, ligando a complejo de receptor y complejo de receptor a elemento de respuesta, no es crítico. Normalmente, la modulación de la expresión del gen exógeno es en respuesta a la unión del complejo de receptor de ecdisona a un elemento de ADN de control o regulador específico. La proteína de receptor de ecdisona, al igual que otros miembros de la familia de receptores de esteroides, presenta al menos tres dominios, un dominio de transactivación, un dominio de unión a ADN y un dominio de unión a ligando. Este receptor, al igual que un subconjunto de la familia de receptores de esteroides, también presenta menos regiones bien definidas responsables de las propiedades de heterodimerización. La unión del ligando al dominio de unión a ligando de la proteína de receptor de ecdisona, tras la heterodimerización con la proteína USP o RXR, permite que los dominios de unión a ADN de las proteínas heterodiméricas se unan al elemento de respuesta en forma activada, dando como resultado por tanto la expresión o supresión del gen exógeno. Este mecanismo no excluye la posibilidad de unión del ligando a o bien EcR o bien USP, y la formación resultante de complejos homodiméricos activos (por ejemplo EcR+EcR o USP+USP). Preferiblemente, uno o más de los dominios de receptor pueden variarse produciendo un cambio de gen quimérico. Normalmente, pueden elegirse uno o más de los tres dominios de una fuente diferente de la fuente de los otros dominios, de modo que el receptor quimérico se optimiza en la célula u organismo huésped elegido para la actividad de transactivación, unión complementaria del ligando y reconocimiento de un elemento de respuesta específico. Además, el propio elemento de respuesta puede estar modificado o sustituido con elementos de respuesta para otros dominios de proteína de unión a ADN tales como la proteína GAL-4 de levadura (véase Sadowski, *et. al.* (1988) *Nature*, 335, 563-564) o la proteína LexA de *E. coli* (véase Brent y Ptashne (1985), *Cell*, 43, 729-736) para acomodar complejos de receptor de ecdisona quiméricos. Otra ventaja de los sistemas quiméricos es que permiten la elección de un promotor usado para dirigir el gen exógeno según un resultado final deseado. Un doble control de este tipo puede ser particularmente importante en áreas de terapia génica, especialmente cuando se producen proteínas citotóxicas, debido a que puede controlarse tanto el tiempo de expresión así como las células en las que se produce la expresión. El término "promotor" significa una secuencia de nucleótidos específica reconocida por ARN polimerasa. La secuencia es el sitio en el que puede iniciarse específicamente la transcripción en condiciones apropiadas. Cuando genes exógenos, unidos operativamente a un promotor adecuado, se introducen en las células del sujeto, la expresión de los genes exógenos está controlada por la presencia del ligando de esta descripción. Los promotores pueden regularse de manera constitutiva o inducible o pueden ser específicos de tejido (es decir, expresados sólo en un tipo particular de célula) o específicos de determinadas etapas de desarrollo del organismo.

Otro aspecto de esta descripción es un método para modular la expresión de uno o más genes exógenos en un sujeto, que comprende administrar al sujeto una cantidad eficaz, es decir, la cantidad requerida para provocar la expresión o supresión génica deseada, de un ligando que comprende un compuesto de la presente invención y en el que las células del sujeto contienen:

a) un complejo de receptor de ecdisona que comprende: 1) un dominio de unión a ADN;

2) un dominio de unión para el ligando; y

3) un dominio de transactivación; y

b) un constructo de ADN que comprende: 1) el gen exógeno; y

2) un elemento de respuesta;

en el que el gen exógeno está bajo el control del elemento de respuesta; y la unión del dominio de unión a ADN al elemento de respuesta en presencia del ligando da como resultado la activación o supresión del gen.

60 Un aspecto relacionado de esta descripción es un método para regular la expresión génica endógena o heteróloga en un sujeto transgénico no humano que comprende poner en contacto un ligando que comprende un compuesto de la presente invención con un receptor de ecdisona dentro de las células del sujeto en el que las células contienen una secuencia de unión a ADN para el receptor de ecdisona y en el que la formación de un complejo de receptor de ecdisona-ligando-secuencia de unión a ADN induce la expresión del gen.

65 Un cuarto aspecto de la presente descripción es un método para producir un polipéptido que comprende las etapas

de:

a) seleccionar una célula que es sustancialmente insensible a la exposición a un ligando que comprende un compuesto de la presente invención;

5

b) introducir en la célula:

1) un constructo de ADN que comprende:

10

i) un gen exógeno que codifica para el polipéptido; y

ii) un elemento de respuesta;

en el que el gen está bajo el control del elemento de respuesta; y

15

2) un complejo de receptor de ecdisona que comprende:

i) un dominio de unión a ADN;

20

ii) un dominio de unión para el ligando; y

iii) un dominio de transactivación; y

c) exponer la célula al ligando.

25

Además de la ventaja de controlar temporalmente la producción de polipéptidos por la célula, este aspecto de la descripción proporciona una ventaja adicional, en los casos en que la acumulación de un polipéptido de este tipo puede dañar la célula, porque la expresión del polipéptido puede limitarse a periodos cortos. Tal control es particularmente importante cuando el gen exógeno es un gen terapéutico. Puede recurrirse a genes terapéuticos para producir polipéptidos que controlan funciones necesarias, tales como la producción de insulina en pacientes diabéticos. También pueden usarse para producir proteínas dañinas o incluso letales, tales como las letales para células cancerosas. Tal control también puede ser importante cuando los niveles de proteínas producidos pueden constituir un desgaste metabólico sobre el crecimiento o la reproducción, tal como en plantas transgénicas.

30

35

En la técnica se conocen bien numerosas secuencias de ácidos nucleicos de ADNc y genómicas que codifican para una variedad de polipéptidos. El material genético exógeno útil con los ligandos de esta descripción incluye genes que codifican para proteínas de interés biológicamente activas, tales como, por ejemplo, proteínas secretoras que pueden liberarse de una célula; enzimas que pueden metabolizar un sustrato de una sustancia tóxica a una sustancia no tóxica, o de una sustancia inactiva a una sustancia activa; proteínas reguladoras; receptores de superficie celular; y similares. Los genes útiles también incluyen genes que codifican factores de coagulación sanguíneos, hormonas tales como insulina, hormona paratiroidea, factor liberador de hormona luteinizante, inhibinas seminales alfa y beta y hormona de crecimiento humana; genes que codifican para proteínas tales como enzimas, cuya ausencia conduce a la aparición de un estado anómalo; genes que codifican citocinas o linfocinas tales como interferones, factor estimulante de colonias de macrófagos y granulocitos, factor estimulante de colonias 1, factor de necrosis tumoral y eritropoyetina; genes que codifican para sustancias inhibitorias tales como alfa<sub>1</sub>-antitripsina, genes que codifican para sustancias que funcionan como fármacos tales como toxinas diftérica y colérica; y similares. Los genes útiles también incluyen los útiles para terapias contra el cáncer y para tratar trastornos genéticos. Los expertos en la técnica tienen acceso a información de secuencias de ácidos nucleicos para prácticamente todos los genes conocidos y o bien pueden obtener la molécula de ácido nucleico directamente de un depósito público, la institución que publicó la secuencia, o bien pueden emplear métodos de rutina para preparar la molécula.

40

45

50

Para su uso en terapia génica, los ligandos descritos en el presente documento pueden llevarse en portadores farmacéuticamente aceptables, tales como, por ejemplo, disoluciones, suspensiones, comprimidos, cápsulas, pomadas, elixires y composiciones inyectables. Las preparaciones farmacéuticas pueden contener desde el 0,01% hasta el 99% en peso del ligando. Las preparaciones pueden estar en forma de dosis o bien individual o bien múltiple. La cantidad de ligando en cualquier preparación farmacéutica particular dependerá de la dosis eficaz, es decir, la dosis requerida para provocar la expresión o supresión génica deseada.

55

60

Las vías de administración adecuadas de las preparaciones farmacéuticas incluyen oral, rectal, tópica (incluyendo dérmica, bucal y sublingual), vaginal, parenteral (incluyendo subcutánea, intramuscular, intravenosa, intradérmica, intratecal y epidural) y mediante sonda nasogástrica. Los expertos en la técnica entenderán que la vía de administración preferida dependerá del estado que va a tratarse y puede variar con factores tales como el estado del receptor.

65

Los ligandos descritos en el presente documento también pueden administrarse conjuntamente con otros

compuestos farmacéuticamente activos. Los expertos en la técnica entenderán que los compuestos farmacéuticamente activos que van a usarse en combinación con los ligandos descritos en el presente documento se seleccionarán con el fin de evitar efectos adversos en el receptor o interacciones indeseables entre los compuestos. Los ejemplos de otros compuestos farmacéuticamente activos que pueden usarse en combinación con los ligandos incluyen, por ejemplo, agentes quimioterápicos contra el SIDA, derivados de aminoácidos, analgésicos, anestésicos, productos anorrectales, antiácidos y antiflatulentos, antibióticos, anticoagulantes, antidotos, agentes antifibrinolíticos, antihistamínicos, agentes antiinflamatorios, antineoplásicos, antiparasitarios, antiprotozoarios, antipiréticos, antisépticos, antiespasmódicos y anticolinérgicos, antivirales, supresores del apetito, medicamentos para la artritis, modificadores de la respuesta biológica, reguladores del metabolismo óseo, purgantes del intestino, agentes cardiovasculares, estimulantes del sistema nervioso central, potenciadores metabólicos cerebrales, cerumenolíticos, inhibidores de la colinesterasa, preparaciones para los resfriados y la tos, factores estimulantes de colonias, anticonceptivos, agentes citoprotectores, preparaciones dentales, desodorantes, productos dermatológicos, agentes detoxificantes, agentes para la diabetes, agentes de diagnóstico, medicamentos contra la diarrea, agonistas del receptor de dopamina, electrolitos, enzimas y productos digestivos, preparaciones ergóticas, agentes para la fertilidad, complementos de fibra, agentes antifúngicos, inhibidores de la galactorrea, inhibidores de la secreción de ácido gástrico, agentes procinéticos gastrointestinales, inhibidores de gonadotropina, estimulantes del crecimiento del pelo, hematínicos, agentes hemorreológicos, hemostáticos, antagonistas del receptor de histamina H<sub>2</sub>, hormonas, agentes hiperglucémicos, hipolipemiantes, inmunosupresores, laxantes, leprotáticos, adyuvantes de leucaféresis, tensioactivos pulmonares, preparaciones contra la migraña, mucolíticos, antagonistas de relajantes musculares, relajantes musculares, antagonistas de narcóticos, pulverizaciones nasales, análogos de nucleósidos para medicamentos contra las náuseas, complementos nutricionales, preparaciones contra la osteoporosis, oxióticos, parasimpaticolíticos, parasimpaticomiméticos, fármacos para el parkinsonismo, adyuvantes de penicilina, fosfolípidos, inhibidores de plaquetas, agentes contra la porfiria, análogos de prostaglandinas, prostaglandinas, inhibidores de la bomba de protones, medicamentos para el prurito, psicofármacos, quinolonas, estimulantes respiratorios, estimulantes de la saliva, sustitutos de sal, agentes esclerosantes, preparaciones para heridas cutáneas, ayudas para dejar de fumar, sulfonamidas, simpaticolíticos, trombolíticos, agentes contra el síndrome de Tourett, preparaciones contra temblores, preparaciones contra la tuberculosis, agentes uricosúricos, agentes de las vías urinarias, contractores del útero, relajantes del útero, preparaciones vaginales, agentes antivertiginosos, análogos de vitamina D, vitaminas y medios de contraste para la obtención de imágenes médicas. En algunos casos los ligandos pueden ser útiles como adyuvantes para la terapia farmacológica, por ejemplo, para “apagar” un gen que produce una enzima que metaboliza un fármaco particular.

Para las aplicaciones de agricultura, además de las aplicaciones descritas anteriormente, los ligandos de esta descripción también pueden usarse para controlar la expresión de proteínas pesticidas tales como la toxina de *Bacillus thuringiensis* (Bt). Tal expresión puede ser específica de tejido o de planta. Además, particularmente cuando también se necesite el control de plagas en plantas, pueden combinarse uno o más pesticidas con los ligandos descritos en el presente documento, proporcionando así ventajas y eficacia adicionales, incluyendo menos aplicaciones totales, que si se aplicaran los pesticidas por separado. Cuando se emplean mezclas con pesticidas, las proporciones relativas de cada componente en la composición dependerán de la eficacia relativa y de la tasa de aplicación deseada de cada pesticida con respecto a las cosechas, plagas y/o malas hierbas que van a tratarse. Los expertos en la técnica reconocerán que las mezclas de pesticidas pueden proporcionar ventajas tales como un espectro de actividad más amplio que un pesticida usado sólo. Los ejemplos de pesticidas que pueden combinarse en composiciones con los ligandos descritos en el presente documento incluyen fungicidas, herbicidas, insecticidas, miticidas y microbicidas.

Los ligandos descritos en el presente documento pueden aplicarse al follaje de una planta como pulverizaciones acuosas mediante métodos empleados comúnmente, tales como pulverizaciones hidráulicas convencionales de gran volumen, pulverizaciones de bajo volumen, pulverizaciones de choro de aire y aéreas. La dilución y la tasa de aplicación dependerán del tipo de equipo empleado, del método y de la frecuencia de aplicación deseados, y de la tasa de aplicación del ligando. Puede ser deseable incluir adyuvantes adicionales en el tanque de pulverización. Tales adyuvantes incluyen tensioactivos, dispersantes, propagantes, adhesivos, agentes antiespumantes, emulsionantes y otros materiales similares descritos en *McCutcheon's Emulsifiers and Detergents*, *McCutcheon's Emulsifiers and Detergents/Functional Materials*, y *McCutcheon's Functional Materials*, todos publicados anualmente por McCutcheon Division de MC Publishing Company (Nueva Jersey). Los ligandos también pueden mezclarse con fertilizantes o materiales fertilizantes antes de su aplicación. Los ligandos y el material fertilizante sólido también puede mezclarse en un equipo de mezclado o combinado, o pueden incorporarse con fertilizantes en formulaciones granulares. Puede usarse cualquier proporción relativa de fertilizante que sea adecuada para las cosechas y malas hierbas que van a tratarse. Los ligandos descritos en el presente documento comprenderán comúnmente desde el 5% hasta el 50% de la composición fertilizante. Estas composiciones proporcionan materiales fertilizantes que promueven el crecimiento rápido de las plantas deseadas, y al mismo tiempo controlan la expresión génica.

#### Células huésped y organismos no humanos de la invención

Tal como se describió anteriormente, los ligandos para modular el sistema de expresión génica de la presente descripción pueden usarse para modular la expresión génica en una célula huésped. La expresión en células huésped transgénicas no humanas puede ser útil para la expresión de diversos genes de interés. La presente

- descripción proporciona ligandos para la modulación de la expresión génica en células huésped procariotas y eucariotas. La expresión en células huésped transgénicas no humanas es útil para la expresión de diversos polipéptidos de interés incluyendo pero sin limitarse a antígenos producidos en plantas como vacunas, enzimas como alfa-amilasa, fitasa, glucanos y xilanasas, genes para la resistencia frente a insectos, nematodos, hongos, bacterias, virus y estrés abiótico, antígenos, productos nutraceuticos, productos farmacéuticos, vitaminas, genes para modificar el contenido en aminoácidos, resistencia a herbicidas, tolerancia al frío, las sequías y el calor, productos industriales, aceites, proteínas, hidratos de carbono, antioxidantes, plantas macho estériles, flores, combustibles, otras características de producción, polipéptidos terapéuticos, productos intermedios de rutas; para la modulación de rutas ya existentes en el huésped para la síntesis de nuevos productos hasta ahora imposibles usando el huésped; ensayos basados en células; ensayos genómicos funcionales, producción de proteínas bioterapéuticas, ensayos proteómicos, y similares. Adicionalmente, los productos génicos pueden ser útiles para conferir rendimientos de crecimiento superiores del huésped o para permitir que se utilice un modo de crecimiento alternativo.
- 15 Por tanto, la presente descripción proporciona ligandos para modular la expresión génica en una célula huésped embrionaria no humana, aislada, según la descripción. La célula huésped puede ser una célula bacteriana, una célula fúngica, una célula de nematodo, una célula de insecto, una célula de pez, una célula vegetal, una célula de ave, una célula animal o una célula de mamífero. Todavía en otra realización, la descripción se refiere a ligandos para modular la expresión génica en una célula huésped, en los que el método comprende cultivar la célula huésped tal como se describió anteriormente en medio de cultivo en condiciones que permiten la expresión de un polinucleótido que codifica para el dominio de unión a ligando de receptores nucleares que comprende una mutación de sustitución, y aislar el dominio de unión a ligando de receptores nucleares que comprende una mutación de sustitución del cultivo.
- 25 En una realización específica, la célula huésped aislada es una célula huésped procariota o una célula huésped eucariota. En otra realización específica, la célula huésped aislada es una célula huésped de invertebrados o una célula huésped embrionaria no humana, de vertebrados. Preferiblemente, la célula huésped se selecciona del grupo que consiste en una célula bacteriana, una célula fúngica, una célula de levadura, una célula de nematodo, una célula de insecto, una célula de pez, una célula vegetal, una célula de ave, una célula de animal embrionaria no humana, y una célula de mamífero embrionaria no humana. Más preferiblemente, la célula huésped es una célula de levadura, una célula de nematodo, una célula de insecto, una célula vegetal, una célula de pez cebra, una célula de pollo, una célula de hámster, una célula de ratón, una célula de rata, una célula de conejo, una célula de gato, una célula de perro, una célula de animal bovino, una célula de cabra, una célula de vaca, una célula de cerdo, una célula de caballo, una célula de oveja, una célula de simio, una célula de mono, una célula de chimpancé o una célula no embrionaria humana. Los ejemplos de células huésped preferidas incluyen, pero no se limitan a, células huésped de especies de hongos o levadura tales como *Aspergillus*, *Trichoderma*, *Saccharomyces*, *Pichia*, *Candida*, *Hansenula*, o especies bacterianas tales como las de los géneros *Synechocystis*, *Synechococcus*, *Salmonella*, *Bacillus*, *Acinetobacter*, *Rhodococcus*, *Streptomyces*, *Escherichia*, *Pseudomonas*, *Metilomonas*, *Metilobacter*, *Alcaligenes*, *Synechocystis*, *Anabaena*, *Tiobacillus*, *Metanobacterium* y *Klebsiella*; de especies vegetales seleccionadas del grupo que consiste en manzana, *Arabidopsis*, mijo perla, plátano, cebada, alubias, remolacha, lenteja negra, garbanzo, chile, pepino, berenjena, habas, maíz, melón, mijo, soja verde, avena, oca, *Panicum*, papaya, cacahuete, guisante, pimienta, guandul, piña, *Phaseolus*, patata, calabaza, arroz, sorgo, soja, calabacín, caña de azúcar, remolacha azucarera, girasol, boniato, té, tomate, tabaco, sandía y trigo; de animales no humanos embrionarios; y de mamífero no humano embrionario.
- 45 En una realización específica, la célula huésped es una célula de levadura seleccionada del grupo que consiste en una célula huésped de *Saccharomyces*, *Pichia* y *Candida*.
- 50 En otra realización específica, la célula huésped es una célula de nematodo *Caenorhabdus elegans*.
- En otra realización específica, la célula huésped es una célula de insecto.
- 55 En otra realización específica, la célula huésped es una célula vegetal seleccionada del grupo que consiste en una célula de manzana, *Arabidopsis*, mijo perla, plátano, cebada, alubias, remolacha, lenteja negra, garbanzo, chile, pepino, berenjena, habas, maíz, melón, mijo, soja verde, avena, oca, *Panicum*, papaya, cacahuete, guisante, pimienta, guandul, piña, *Phaseolus*, patata, calabaza, arroz, sorgo, soja, calabacín, caña de azúcar, remolacha azucarera, girasol, boniato, té, tomate, tabaco, sandía y trigo.
- 60 En otra realización específica, la célula huésped es una célula de pez cebra.
- 65 En otra realización específica, la célula huésped es una célula de mamífero seleccionada del grupo que consiste en una célula de hámster, una célula de ratón, una célula de rata, una célula de conejo, una célula de gato, una célula de perro, una célula de animal bovino, una célula de cabra, una célula de vaca, una célula de cerdo, una célula de caballo, una célula de oveja, una célula de mono, una célula de chimpancé y una célula de ser humano no humana

embrionaria.

En la técnica se conoce bien la transformación de células huésped y puede lograrse mediante una variedad de métodos incluyendo pero sin limitarse a electroporación, infección viral, transfección mediante plásmido/vector, transfección mediada por vector no viral, transformación mediada por *Agrobacterium*, bombardeo de partículas, y similares. La expresión de productos génicos deseados implica cultivar las células huésped transformadas en condiciones adecuadas e inducir la expresión del gen transformado. En la técnica se conocen bien las condiciones de cultivo y los protocolos de expresión génica en células procariontas y eucariotas (véase la sección de métodos generales de los ejemplos). Las células pueden recogerse y los productos génicos aislarse según protocolos específicos para el producto génico.

Además, puede elegirse una célula huésped que modula la expresión del polinucleótido insertado, o modifica y procesa el producto de polipéptido del modo específico deseado. Las diferentes células huésped tienen mecanismos característicos y específicos para la modificación y el procesamiento traduccional y postraduccional [por ejemplo, glicosilación, escisión (por ejemplo, de secuencia señal)] de proteínas. Las líneas celulares o sistemas huésped apropiados pueden elegirse para garantizar la modificación y el procesamiento deseados de la proteína foránea expresada. Por ejemplo, la expresión en un sistema bacteriano puede usarse para producir un producto proteico de núcleo no glicosilado. Sin embargo, un polipéptido expresado en bacterias puede no estar plegado apropiadamente. La expresión en levaduras puede producir un producto glicosilado. La expresión en células eucariotas puede aumentar la posibilidad de plegamiento y glicosilación "nativos" de una proteína heteróloga. Además, la expresión en células de mamífero puede proporcionar una herramienta para reconstituir, o constituir, la actividad del polipéptido. Además, diferentes sistemas de expresión de vector/huésped pueden afectar a reacciones de procesamiento, tales como escisiones proteolíticas, en un grado diferente. La presente descripción también se refiere a un organismo no humano que comprende una célula huésped aislada según la descripción. En una realización específica, el organismo no humano es un organismo procarionta o un organismo eucariota. En otra realización específica, el organismo no humano es un organismo invertebrado o un organismo vertebrado.

Preferiblemente, el organismo no humano se selecciona del grupo que consiste en una bacteria, un hongo, una levadura, un nematodo, un insecto, un pez, una planta, un ave, un animal y un mamífero. Más preferiblemente, el organismo no humano es una levadura, un nematodo, un insecto, una planta, un pez cebra, un pollo, un hámster, un ratón, una rata, un conejo, un gato, un perro, un animal bovino, una cabra, una vaca, un cerdo, un caballo, una oveja, un simio, un mono o un chimpancé.

En una realización específica, el organismo no humano es una levadura seleccionada del grupo que consiste en *Saccharomyces*, *Pichia* y *Candida*.

En otra realización específica, el organismo no humano es un nematodo *Caenorhabdus elegans*.

En otra realización específica, el organismo no humano es una planta seleccionada del grupo que consiste en manzana, *Arabidopsis*, mijo perla, plátano, cebada, alubias, remolacha, lenteja negra, garbanzo, chile, pepino, berenjena, habas, maíz, melón, mijo, soja verde, avena, oca, *Panicum*, papaya, cacahuete, guisante, pimiento, guandul, piña, *Phaseolus*, patata, calabaza, arroz, sorgo, soja, calabacín, caña de azúcar, remolacha azucarera, girasol, boniato, té, tomate, tabaco, sandía y trigo.

En otra realización específica, el organismo no humano es un ratón *Mus musculus*.

#### Sistema de modulación de la expresión génica de la invención

La presente descripción se refiere a un grupo de ligandos que son útiles en un sistema de expresión génica inducible basado en el receptor de ecdisona. Tal como se presenta en el presente documento, un grupo novedoso de ligandos proporciona un sistema de expresión génica inducible mejorado en células huésped tanto procariontas como eucariotas. Por tanto, la presente descripción se refiere a ligandos que son útiles para modular la expresión de genes. En particular, la presente descripción se refiere a ligandos que tienen la capacidad de transactivar un sistema de modulación de la expresión génica que comprende al menos un casete de expresión génica que puede expresarse en una célula huésped que comprende un polinucleótido que codifica para un polipéptido que comprende un dominio de unión a ligando de receptores nucleares de grupo H. Preferiblemente, la unión a ligando de receptores nucleares de grupo H es de un receptor de ecdisona, un receptor ubicuo, un receptor huérfano 1, un NER-1, un receptor nuclear de hormonas esteroideas 1, un receptor X retinoide que interacciona con la proteína-15, un receptor X hepático  $\beta$ , una proteína similar al receptor de hormonas esteroideas, un receptor X hepático, un receptor X hepático  $\alpha$ , un receptor X farnesoide, un receptor que interacciona con la proteína 14 y un receptor de farnesol. Más preferiblemente, el dominio de unión a ligando de receptores nucleares de grupo H es de un receptor de ecdisona.

En una realización específica, el sistema de modulación de la expresión génica comprende un casete de expresión génica que comprende un polinucleótido que codifica para un polipéptido que comprende un dominio de transactivación, un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya

expresión va a modularse; y un dominio de unión a ligando de receptores nucleares de grupo H que comprende una mutación de sustitución. El sistema de modulación de la expresión génica puede comprender además un segundo casete de expresión génica que comprende: i) un elemento de respuesta reconocido por el dominio de unión a ADN del polipéptido codificado del primer casete de expresión génica; ii) un promotor que se activa mediante el dominio de transactivación del polipéptido codificado del primer casete de expresión génica; y iii) un gen cuya expresión va a modularse.

En otra realización específica, el sistema de modulación de la expresión génica comprende un casete de expresión génica que comprende a) un polinucleótido que codifica para un polipéptido que comprende un dominio de transactivación, un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y un dominio de unión a ligando de receptores nucleares de grupo H que comprende una mutación de sustitución, y b) un segundo dominio de unión a ligando de receptores nucleares seleccionado del grupo que consiste en un dominio de unión a ligando del receptor X retinoide de vertebrados, un dominio de unión a ligando del receptor X retinoide de invertebrados, un dominio de unión a ligando de proteína ultraespiráculo y un dominio de unión a ligando químérico que comprende dos fragmentos de polipéptido, en el que el primer fragmento de polipéptido es de un dominio de unión a ligando del receptor X retinoide de vertebrados, un dominio de unión a ligando del receptor X retinoide de invertebrados o un dominio de unión a ligando de proteína ultraespiráculo, y el segundo fragmento de polipéptido es de un dominio de unión a ligando del receptor X retinoide de vertebrados, dominio de unión a ligando del receptor X retinoide de invertebrados o dominio de unión a ligando de proteína ultraespiráculo diferente. El sistema de modulación de la expresión génica puede comprender además un segundo casete de expresión génica que comprende: i) un elemento de respuesta reconocido por el dominio de unión a ADN del polipéptido codificado del primer casete de expresión génica; ii) un promotor que se activa mediante el dominio de transactivación del polipéptido codificado del primer casete de expresión génica; y iii) un gen cuya expresión va a modularse.

En otra realización específica, el sistema de modulación de la expresión génica comprende un primer casete de expresión génica que comprende un polinucleótido que codifica para un primer polipéptido que comprende un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse y un dominio de unión a ligando de receptores nucleares, y un segundo casete de expresión génica que comprende un polinucleótido que codifica para un segundo polipéptido que comprende un dominio de transactivación y un dominio de unión a ligando de receptores nucleares, en el que uno de los dominios de unión a ligando de receptores nucleares es un dominio de unión a ligando de receptores nucleares de grupo H que comprende una mutación de sustitución. En una realización preferida, el primer polipéptido está sustancialmente libre de un dominio de transactivación y el segundo polipéptido está sustancialmente libre de un dominio de unión a ADN. Para los fines de la descripción, “sustancialmente libre” significa que la proteína en cuestión no contiene suficiente secuencia del dominio en cuestión para proporcionar actividad de activación o unión. El sistema de modulación de la expresión génica puede comprender además un tercer casete de expresión génica que comprende: i) un elemento de respuesta reconocido por el dominio de unión a ADN del primer polipéptido del primer casete de expresión génica; ii) un promotor que se activa mediante el dominio de transactivación del segundo polipéptido del segundo casete de expresión génica; y iii) un gen cuya expresión va a modularse.

El que caso en el que sólo un dominio de unión a ligando de receptores nucleares sea un dominio de unión a ligando de grupo H que comprende una mutación de sustitución, el otro dominio de unión a ligando de receptores nucleares puede ser de cualquier otro receptor nuclear que forme un dímero con el dominio de unión a ligando de grupo H que comprende la mutación de sustitución. Por ejemplo, cuando el dominio de unión a ligando de receptores nucleares de grupo H que comprende una mutación de sustitución es un dominio de unión a ligando del receptor de ecdisona que comprende una mutación de sustitución, el otro dominio de unión a ligando de receptores nucleares (“pareja”) puede ser de un receptor de ecdisona, un receptor X retinoide (RXR) de vertebrados, un RXR de invertebrados, una proteína ultraespiráculo (USP) o un receptor nuclear químérico que comprende al menos dos fragmentos de polipéptido de dominio de unión a ligando de receptores nucleares diferentes seleccionados del grupo que consiste en un RXR de vertebrados, un RXR de invertebrados y una USP (véanse las solicitudes de patente en tramitación junto con la presente PCT/US01/09050, PCT/US02/05235 y PCT/US02/05706). El dominio de unión a ligando de receptores nucleares “pareja” puede comprender además una mutación de truncamiento, una mutación de delección, una mutación de sustitución u otra modificación.

Preferiblemente, el dominio de unión a ligando de RXR de vertebrados es de un RXR de ser humano *Homo sapiens*, ratón *Mus musculus*, rata *Rattus norvegicus*, pollo *Gallus gallus*, cerdo *Sus scrofa domestica*, rana *Xenopus laevis*, pez cebra *Danio rerio*, tunicado *Polyandrocampa misakiensis* o medusa *Tripedalia cysophora*.

Preferiblemente, el dominio de unión a ligando de RXR de invertebrados es de un polipéptido ultraespiráculo de langosta *Locusta migratoria* (“LmUSP”), un homólogo de RXR 1 de garrapata *Amblyomma americanum* (“AmaRXR1”), un homólogo de RXR 2 de garrapata *Amblyomma americanum* (“AmaRXR2”), un homólogo de RXR de cangrejo violinista *Celca pugilator* (“CpRXR”), un homólogo de RXR de escarabajo *Tenebrio molitor* (“TmRXR”), un homólogo de RXR de una abeja *Apis mellifera* (“AmRXR”), un homólogo de RXR de áfido *Myzus persicae* (“MpRXR”) o un homólogo de RXR de no díptero/no lepidóptero.

- Preferiblemente, el dominio de unión a ligando de RXR quimérico comprende al menos dos fragmentos de polipéptido seleccionados del grupo que consiste en un fragmento de polipéptido de RXR de especie de vertebrado, un fragmento de polipéptido de RXR de especie de invertebrado y un fragmento de polipéptido homólogo de RXR de especie de invertebrado no díptero/no lepidóptero. Un dominio de unión a ligando de RXR quimérico para su uso en la presente invención puede comprender al menos dos fragmentos de polipéptido de RXR de especies diferentes, o cuando la especies son iguales, los dos o más fragmentos de polipéptido pueden ser de dos o más isoformas diferentes del fragmento de polipéptido de RXR de la especie.
- En una realización preferida, el dominio de unión a ligando de RXR quimérico comprende al menos un fragmento de polipéptido de RXR de especie de vertebrado y un fragmento de polipéptido de RXR de especie de invertebrado.
- En una realización más preferida, el dominio de unión a ligando de RXR quimérico comprende al menos un fragmento de polipéptido de RXR de especie de vertebrado y un fragmento de polipéptido homólogo de RXR de especie de invertebrado no díptero/no lepidóptero.
- En una realización específica, el gen cuya expresión va a modularse es un gen homólogo con respecto a la célula huésped. En otra realización específica, el gen cuya expresión va a modularse es un gen heterólogo con respecto a la célula huésped.
- Los ligandos para su uso en la presente invención tal como se describe a continuación, cuando se combinan con el dominio de unión a ligando del/de los receptor(es) nuclear(es), que a su vez se une(n) al elemento de respuesta unido a un gen, proporcionan los medios para la regulación temporal externa de expresión del gen. El mecanismo de unión o el orden en el que los diversos componentes de esta descripción se unen entre sí, es decir, por ejemplo, ligando a dominio de unión a ligando, dominio de unión a ADN a elemento de respuesta, dominio de transactivación a promotor, etc., no es crítico.
- En un ejemplo específico, la unión del ligando al dominio de unión a ligando de un receptor nuclear de grupo H y su pareja de dominio de unión a ligando de receptores nucleares permite la expresión o supresión del gen. Este mecanismo no excluye la posibilidad de unión del ligando al receptor nuclear de grupo H (GHNR) o su pareja, y la formación resultante de complejos homodiméricos activos (por ejemplo GHNR + GHNR o pareja + pareja). Preferiblemente, uno o más de los dominios de receptor varían produciendo un cambio de gen híbrido. Normalmente, pueden elegirse uno o más de los tres dominios, DBD, LBD y dominio de transactivación, de una fuente diferente de la fuente de los otros dominios de modo que los genes híbridos y las proteínas híbridas resultantes se optimizan en la célula u organismo huésped elegido para la actividad de transactivación, unión complementaria del ligando y reconocimiento de un elemento de respuesta específico. Además, el propio elemento de respuesta puede estar modificado o sustituido con elementos de respuesta para otros dominios de proteína de unión a ADN tales como la proteína GAL-4 de levadura (véase Sadowski, *et al.* (1988) *Nature*, 335: 563-564) o la proteína LexA de *Escherichia coli* (véase Brent y Ptashne (1985), *Cell*, 43: 729-736), o elementos de respuesta sintéticos específicos para interacciones dirigidas con proteínas diseñadas, modificadas y seleccionadas para tales interacciones específicas (véase, por ejemplo, Kim, *et al.* (1997), *Proc. Natl. Acad. Sci., USA*, 94: 3616-3620) para acomodar receptores híbridos. Otra ventaja de los sistemas de dos híbridos es que permiten la elección de un promotor usado para dirigir la expresión del gen según un resultado final deseado. Un doble control de este tipo puede ser particularmente importante en áreas de terapia génica, especialmente cuando se producen proteínas citotóxicas, debido a que puede controlarse tanto el tiempo de expresión así como las células en las que se produce la expresión. Cuando los genes, unidos operativamente a un promotor adecuado, se introducen en las células del sujeto, la expresión de los genes exógenos está controlada por la presencia del sistema de esta descripción. Los promotores pueden regularse de manera constitutiva o inducible o pueden ser específicos de tejido (es decir, expresados sólo en un tipo particular de célula) o específicos de determinadas etapas de desarrollo del organismo.
- El receptor de ecdisona es un miembro de la superfamilia de receptores nucleares y se clasifica en la subfamilia 1, grupo H (denominado en el presente documento "receptores nucleares de grupo H"). Los miembros de cada grupo comparten un 40-60% de identidad de aminoácidos en el dominio E (unión a ligando) (Laudet *et al.*, A Unified Nomenclature System for the Nuclear Receptor Subfamily, 1999; *Cell* 97: 161-163). Además del receptor de ecdisona, otros miembros de esta subfamilia 1 de receptores nucleares, grupo H, incluyen: receptor ubicuo (UR), receptor huérfano 1 (O-1), receptor nuclear de hormonas esteroideas 1 (NER-1), receptor X retinoide que interacciona con la proteína-15 (RIP-15), receptor X hepático  $\beta$  (LXR $\beta$ ), proteína similar al receptor de hormonas esteroideas (RLD-1), receptor X hepático (LXR), receptor X hepático  $\alpha$  (LXR $\alpha$ ), receptor X farnesoide (FXR), receptor que interacciona con proteína 14 (RIP-14) y receptor de farnesol (HRR-1).
- En particular, en el presente documento se describen ligandos novedoso útiles en un sistema de modulación de la expresión génica que comprende un dominio de unión a ligando de receptores nucleares de grupo H que comprende una mutación de sustitución. Este sistema de expresión génica puede ser un sistema de expresión génica basado en "cambio individual" en el que el dominio de transactivación, dominio de unión a ADN y dominio de unión a ligando están en un polipéptido codificado. Alternativamente, el sistema de modulación de la expresión génica puede ser un sistema de modulación de la expresión génica basado en "cambio doble" o "en dos híbridos" en el que el dominio de transactivación y dominio de unión a ADN se ubican en dos polipéptidos codificados diferentes.

Un sistema de modulación de la expresión génica basado en el receptor de ecdisona de la presente descripción puede ser o bien heterodimérico o bien homodimérico. Un complejo de EcR funcional se refiere en general a un complejo de proteína heterodimérico que consiste en dos miembros de la familia de receptores de esteroides, una proteína de receptor de ecdisona obtenida de diversos insectos y una proteína ultraespiráculo (USP) o el homólogo de vertebrados de USP, proteína de receptor X retinoide (véase Yao, *et al.* (1993) *Nature* 366, 476-479; Yao, *et al.*, (1992) *Cell* 71, 63-72). Sin embargo, el complejo también puede ser un homodímero tal como se detalla a continuación. El complejo de receptor ecdiesteroide funcional también puede incluir proteína(s) adicional(es) tal(es) como inmunofilinas. Los miembros adicionales de la familia de receptores de esteroides de proteínas, conocidos como factores de transcripción (tales como DHR38 o *betaFTZ-1*), también pueden ser parejas dependientes o independientes de ligando para EcR, USP y/o RXR. Adicionalmente, pueden requerirse otros cofactores tales como proteínas conocidas en general como coactivadores (también denominadas adaptadores o mediadores). Estas proteínas no se unen de manera específica de secuencia a ADN y no están implicadas en la transcripción basal. Pueden ejercer su efecto sobre la activación de la transcripción a través de diversos mecanismos, incluyendo estimulación de la unión a ADN de activadores, afectando a la estructura de la cromatina o mediando en las interacciones de activador-complejo de iniciación. Los ejemplos de tales coactivadores incluyen RIP140, TIF1, RAP46Bag-1, ARA70, SRC-1/NCoA-1, TIF2/GRIP/NCoA-2, ACTR/AIB1/RAC3/pCIP así como la proteína de unión al elemento de respuesta B del coactivador C promiscuo, CBP/p300 (para revisión véase Glass *et al.*, *Curr. Opin. Cell Biol.* 9:222-232, 1997). Además, pueden requerirse cofactores proteicos conocidos en general como correpresores (también conocidos como represores, silenciadores o mediadores del silenciamiento) para inhibir eficazmente la activación transcripcional en ausencia de ligando. Estos correpresores pueden interactuar con el receptor de ecdisona no ligado para silenciar la actividad en el elemento de respuesta. Las pruebas actuales sugieren que la unión del ligando cambia la conformación del receptor, lo que da como resultado la liberación del correpresor y el reclutamiento de los coactivadores descritos anteriormente, suprimiendo así su actividad de silenciamiento. Los ejemplos de correpresores incluyen N-CoR y SMRT (para revisión, véase Horwitz *et al.* *Mol Endocrinol.* 10: 1167-1177, 1996). Estos cofactores o bien pueden ser endógenos dentro de la célula u organismo o bien pueden añadirse exógenamente como transgenes para expresarse de modo o bien regulado o bien no regulado. Los complejos homodiméricos de la proteína de receptor de ecdisona, USP o RXR también pueden ser funcionales en ciertas circunstancias.

El complejo de receptor de ecdisona normalmente incluye proteínas que son miembros de la superfamilia de receptores nucleares en el que todos los miembros se caracterizan generalmente por la presencia de un dominio de transactivación amino-terminal, un dominio de unión a ADN ("DBD") y un dominio de unión a ligando ("LBD") separado del DBD por una región bisagra. Tal como se usa en el presente documento, el término "dominio de unión a ADN" comprende una secuencia de polipéptido mínima de una proteína de unión a ADN, hasta la longitud completa de una proteína de unión a ADN, siempre que el dominio de unión a ADN funcione para asociarse con un elemento de respuesta particular. Los miembros de la superfamilia de receptores nucleares también se caracterizan por la presencia de cuatro o cinco dominios: A/B, C, D, E, y en algunos miembros F (véase la patente estadounidense 4.981.784 y Evans, *Science* 240:889-895 (1988)). El dominio "AB" corresponde al dominio de transactivación, "C" corresponde al dominio de unión a ADN, "D" corresponde a la región bisagra y "E" corresponde al dominio de unión a ligando. Algunos miembros de la familia también pueden tener otro dominio de transactivación en el lado carboxilo-terminal del LBD que corresponde a "F".

El DBD se caracteriza por la presencia de dos dedos de zinc-cisteína entre los que hay dos motivos de amino ácido, la caja P y la caja D, que confieren especificidad para elementos de respuesta a ecdisona. Estos dominios pueden ser o bien nativos o bien modificados o bien quimeras de diferentes dominios de proteínas de receptor heterólogas. El receptor de EcR, como un subconjunto de la familia de receptores de esteroides, también presenta menos regiones bien definidas responsables de las propiedades de heterodimerización. Debido a que los dominios de receptores nucleares son de naturaleza modular, pueden intercambiarse los LBD, DBD y dominios de transactivación.

Se sabe que los sistemas de cambio génico incorporan componentes del complejo de receptor de ecdisona. Sin embargo, en estos sistemas conocidos, siempre que se usa EcR, está asociado con dominios de unión a ADN nativos o modificados y dominios de transactivación en la misma molécula. USP o RXR se usan normalmente como parejas silenciosas. Anteriormente se ha mostrado que cuando los dominios de unión a ADN y los dominios de transactivación están en la misma molécula, la actividad de fondo en ausencia de ligando es alta y que tal actividad se reduce drásticamente cuando los dominios de unión a ADN y los dominios de transactivación están en moléculas diferentes, es decir, en cada una de las dos parejas de un complejo heterodimérico u homodimérico (véase el documento PCT/US01/09050).

#### Método de modulación de la expresión génica de la invención

La presente descripción también se refiere a métodos de modulación de la expresión génica en una célula huésped usando un sistema de modulación de la expresión génica según la descripción. Específicamente, la presente descripción proporciona un método de modulación de la expresión de un gen en una célula huésped que comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica según la

descripción; y b) introducir en la célula huésped a ligando; en el que el gen que va a modularse es un componente de un casete de expresión génica que comprende: i) un elemento de respuesta que comprende un dominio reconocido por dominio de unión a ADN del sistema de expresión génica; ii) un promotor que se activa mediante el dominio de transactivación del sistema de expresión génica; y iii) un gen cuya expresión va a modularse, mediante lo cual tras la introducción del ligando en la célula huésped, se modula la expresión del gen.

La descripción también proporciona un método de modulación de la expresión de un gen en una célula huésped que comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica según la descripción; b) introducir en la célula huésped un casete de expresión génica según la descripción, en el que el casete de expresión génica comprende i) un elemento de respuesta que comprende un dominio reconocido por el dominio de unión a ADN del sistema de expresión génica; ii) un promotor que se activa mediante el dominio de transactivación del sistema de expresión génica; y iii) un gen cuya expresión va a modularse; y c) introducir en la célula huésped un ligando; mediante lo cual tras la introducción del ligando en la célula huésped, se modula la expresión del gen.

La presente descripción también proporciona un método de modulación de la expresión de un gen en una célula huésped que comprende un casete de expresión génica que comprende un elemento de respuesta que comprende un dominio al que se une el dominio de unión a ADN del primer polipéptido híbrido del sistema de modulación de la expresión génica; un promotor que se activa mediante el dominio de transactivación del segundo polipéptido híbrido del sistema de modulación de la expresión génica; y un gen cuya expresión va a modularse; en el que el método comprende las etapas de: a) introducir en la célula huésped un sistema de modulación de la expresión génica según la descripción; y b) introducir en la célula huésped un ligando; mediante lo cual tras la introducción del ligando en el huésped, se modula la expresión del gen.

Los genes de interés para la expresión en una célula huésped usando los métodos dados a conocer en el presente documento pueden ser genes endógenos o genes heterólogos. La información de la secuencia de ácido nucleico o aminoácidos para un gen o una proteína deseada puede estar ubicada en una de las muchas bases de datos de acceso público, por ejemplo, GENBANK, EMBL, Swiss-Prot y PIR, o en muchas publicaciones de revistas relacionadas con la biología. Por tanto, los expertos en la técnica tienen acceso a la información de secuencias de ácidos nucleicos para prácticamente todos los genes conocidos. Tal información puede usarse entonces para construir los constructos deseados para la inserción del gen de interés dentro de los casetes de expresión génica usados en los métodos descritos en el presente documento.

Los ejemplos de genes de interés para la expresión en una célula huésped usando métodos expuestos en el presente documento incluyen, pero no se limitan a: antígenos producidos en plantas como vacunas, enzimas como alfa-amilasa, fitasa, glucanosa y xilanasas, genes para la resistencia frente a insectos, nematodos, hongos, bacterias, virus y estrés abiótico, productos nutracéuticos, productos farmacéuticos, vitaminas, genes para modificar el contenido en aminoácidos, resistencia a herbicidas, tolerancia al frío, las sequías y el calor, productos industriales, aceites, proteínas, hidratos de carbono, antioxidantes, plantas macho estériles, flores, combustibles, otras características de producción, genes que codifican para polipéptidos terapéuticamente deseables o productos que pueden usarse para tratar un estado, una enfermedad, un trastorno, una disfunción, un defecto genético, tales como anticuerpos monoclonales, enzimas, proteasas, citocinas, interferones, insulina, eritropoyetina, factores de coagulación, otros factores o componentes sanguíneos, vectores virales para terapia génica, virus para vacunas, dianas para el descubrimiento de fármacos, análisis proteómicos y genómicos funcionales y aplicaciones, y similares.

#### Medición de la expresión/transcripción génica

Una medición útil de los métodos de la descripción es la del estado transcripcional de la célula que incluye las identidades y abundancias de ARN, preferiblemente especies de ARNm. Tales mediciones se realizan convenientemente midiendo las abundancias de ADNc mediante cualquiera de las diversas tecnologías de expresión génica existentes.

La tecnología de matriz de ácidos nucleicos es una técnica útil para determinar la expresión de ARNm diferencial. Una tecnología de este tipo incluye, por ejemplo, chips de oligonucleótidos y micromatrices de ADN. Estas técnicas se basan en fragmentos de ADN u oligonucleótidos que se corresponden con diferentes genes o ADNc que están inmovilizados sobre un soporte sólido y se hibridan con sondas preparadas a partir de conjuntos de ARNm total extraídos de células, tejidos u organismos completos y se convierten en ADNc. Los chips de oligonucleótidos son matrices de oligonucleótidos sintetizados en un sustrato usando técnicas fotolitográficas. Se han producido chips que pueden analizar hasta 1700 genes. Las micromatrices de ADN son matrices de muestras de ADN, normalmente productos de PCR, que se imprimen robóticamente sobre un portaobjetos microscópico. Cada gen se analiza mediante una secuencia de ADN diana de longitud parcial o completa. Actualmente se preparan comercialmente de manera rutinaria micromatrices con hasta 10.000 genes. La diferencia principal entre estas dos técnicas es que los chips de oligonucleótidos normalmente utilizan oligonucleótidos de 25 meros que permiten el fraccionamiento de moléculas de ADN cortas mientras que las dianas de ADN más largas de las micromatrices, de aproximadamente 1000 pares de bases, pueden proporcionar más sensibilidad en el fraccionamiento de mezclas de ADN complejas.

Otra medición útil de los métodos de la descripción es la determinación del estado de traducción de la célula midiendo las abundancias de la especie de proteína constituyente presente en la célula usando procedimientos bien conocidos en la técnica.

5 Cuando se desea la identificación de genes asociados con diversas funciones fisiológicas, puede emplearse un ensayo en el que se miden cambios en tales funciones como crecimiento celular, apoptosis, senescencia, diferenciación, adhesión, unión a una molécula específica, unión a otra célula, organización celular, organogénesis, transporte intracelular, facilitación del transporte, conversión de energía, metabolismo, miogénesis, neurogénesis, y/o hematopoyesis.

Además, puede usarse la expresión génica de marcador o indicador seleccionable para medir la modulación de la expresión génica usando la presente invención.

15 En la técnica se conocen bien otros métodos para detectar los productos de expresión génica e incluyen análisis de transferencias de tipo Southern (detección de ADN), transferencias de punto o de ranura (ADN, ARN), transferencias de tipo Northern (ARN), RT-PCR (ARN), inmunotransferencias de tipo Western (detección de polipéptidos) y ELISA (polipéptidos). Aunque es menos preferido, pueden usarse proteínas marcadas para detectar una secuencia de ácido nucleico particular a la que se hibrida.

20 En algunos casos es necesario amplificar la cantidad de una secuencia de ácido nucleico. Esto puede llevarse a cabo usando uno o más de varios métodos adecuados que incluyen, por ejemplo, reacción en cadena de la polimerasa ("PCR"), reacción en cadena de la ligasa ("LCR"), amplificación por desplazamiento de cadena ("SDA"), amplificación basada en la transcripción, y similares. La PCR se lleva a cabo de acuerdo con técnicas conocidas en las que, por ejemplo, se trata una muestra de ácido nucleico en presencia de una ADN polimerasa termoestable, en condiciones de hibridación, con un par de cebadores oligonucleotídicos, con un cebador que se hibrida con una cadena (molde) de la secuencia específica que va a detectarse. Los cebadores son suficientemente complementarios a cada cadena de molde de la secuencia específica para hibridarse con la misma. Un producto de extensión de cada cebador se sintetiza y es complementario a la cadena de molde de ácido nucleico a la que se hibrida. El producto de extensión sintetizado de cada cebador también puede servir como molde para la síntesis adicional de productos de extensión usando los mismos cebadores. Tras un número suficiente de rondas de síntesis de productos de extensión, puede analizarse la muestra tal como se describió anteriormente para evaluar si la secuencia o secuencias que van a detectarse están presentes.

35 La presente invención puede entenderse mejor con referencia a los siguientes ejemplos, que se proporcionan para ilustrar la invención.

### Ejemplos

#### 40 MÉTODOS GENERALES

Las técnicas convencionales de clonación molecular y ADN recombinante usadas en el presente documento se conocen bien en la técnica y se describen por Sambrook, J., Fritsch, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y. (1989) (Maniatis) y por T. J. Silhavy, M. L. Bannan, y L. W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1984) y por Ausubel, F. M. *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley-Interscience (1987).

50 Se conocen bien en la técnica materiales y métodos adecuados para el mantenimiento y crecimiento de cultivos bacterianos. Pueden encontrarse técnicas adecuadas para su uso en los siguientes ejemplos tal como se expone en *Manual of Methods for General Bacteriology* (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg y G. Briggs Phillips, eds), American Society for Microbiology, Washington, DC. (1994) o por Thomas D. Brock en *Biotechnology: A Textbook of Industrial Microbiology*, segunda edición, Sinauer Associates, Inc., Sunderland, MA (1989). Todos los reactivos, enzimas de restricción y materiales usados para el crecimiento y mantenimiento de células huésped se obtuvieron de Aldrich Chemicals (Milwaukee, WI), DIFCO Laboratories (Detroit, MI), GIBCO/BRL (Gaithersburg, MD) o Sigma Chemical Company (St. Louis, MO) a menos que se especifique lo contrario.

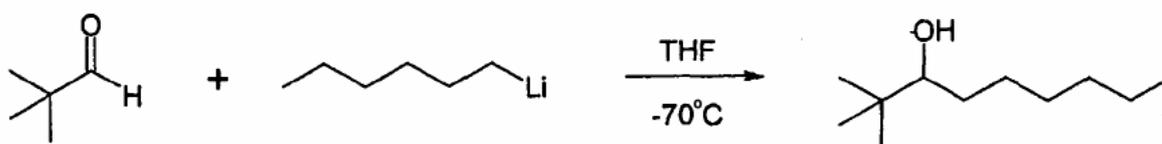
60 Pueden lograrse manipulaciones de secuencias genéticas usando el juego de programas disponible del Genetics Computer Group Inc. (Wisconsin Package versión 9.0, Genetics Computer Group (GCG), Madison, WI). Cuando se usa el programa de GCG "Pileup", puede usarse el valor por defecto de creación de huecos de 12 y el valor por defecto de extensión de huecos de 4. Cuando se usa el programa de GCG "Gap" o "Bestfit", puede usarse la penalización por creación de huecos por defecto de 50 y la penalización por extensión de huecos por defecto de 3. En cualquier caso en el que no se soliciten parámetros del programa de GCG, en estos o en cualquier otro programa de GCG pueden usarse valores por defecto.

El significado de las abreviaturas es tal como sigue: "h" significa hora(s), "min." significa minuto(s), "s" significa segundo(s), "d" significa día(s), "μl" significa microlitro(s), "ml" significa mililitro(s), "l" significa litro(s), "μM" significa micromolar, "mM" significa milimolar, "M" significa molar, "mmol" significa milimoles, "μg" significa microgramo(s), "mg" significa miligramo(s), "A" significa adenina o adenosina, "T" significa timina o timidina, "G" significa guanina o guanosina, "C" significa citidina o citosina, "x g" significa veces la gravedad, "nt" significa nucleótido(s), "aa" significa aminoácido(s), "pb" significa par(es) de base(s), "kb" significa kilobase(s), "k" significa kilo, "μ" significa micro, "°C" significa grados centígrados, "C" en el contexto de una ecuación química significa centígrado, "THF" significa tetrahidrofurano, "DME" significa dimetoxietano, "DMF" significa dimetilformamida, "RMN" significa resonancia magnética nuclear, "psi" se refiere a libras por pulgada cuadrada y "CCF" significa cromatografía en capa fina.

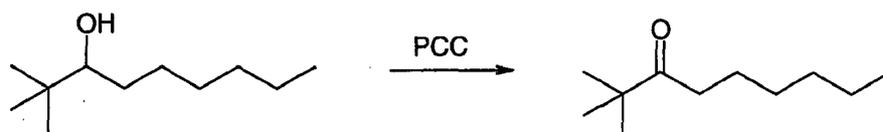
#### EJEMPLO 1: PREPARACIÓN DE COMPUESTOS

Pueden prepararse los compuestos de la presente invención según las siguientes rutas de síntesis.

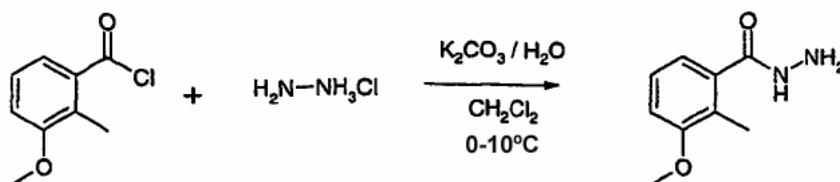
##### 1.1 Preparación de RG-115853 (referencia)



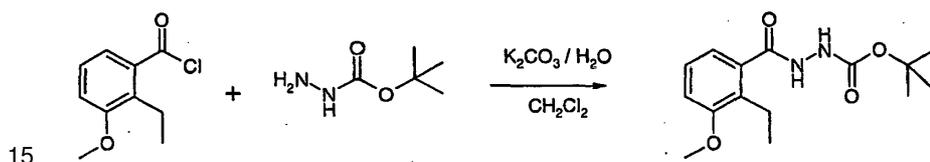
Se disolvieron 20,0 g (0,232 mol) de pivaldehído en 600 ml de THF en un matraz de 2 l de fondo redondo, de 3 bocas, equipado con una barra de agitación magnética y un termómetro. Se purgó el matraz con N<sub>2</sub>. Se enfrió la disolución hasta -65°C en un baño de nieve carbónica/acetona. Se añadieron 100 ml de una disolución 2,3 M de hexil-litio en hexano (0,23 moles) mediante una jeringa de vidrio de 20 ml insertada entre un tapón de caucho y el cuello de vidrio, en porciones pequeñas de 5 ml, manteniendo la temperatura a o por debajo de -60°C. Tras agitar a -60°C durante 1 hora, se permitió que la reacción se calentara hasta aproximadamente -5°C a lo largo de una hora. Se enfrió de nuevo la reacción hasta -60°C y se extinguió lentamente con disolución acuosa de NH<sub>4</sub>Cl, y permitiendo que la temperatura aumentara por encima de -50°C. Se permitió que la mezcla de reacción se calentara a medida que se añadieron 100 ml de agua. Se eliminó el THF en un evaporador rotatorio, manteniendo la temperatura del baño de agua a 25-30°C, para evitar la pérdida de producto volátil. Se extrajo el producto con etil éter; se secó la fase orgánica y se eliminó cuidadosamente el disolvente en un evaporador rotatorio, monitorizando la pérdida de peso. Se usó el producto, 2,2-dimetil-nonan-3-ol, en la etapa de oxidación posterior como una disolución altamente concentrada en éter.



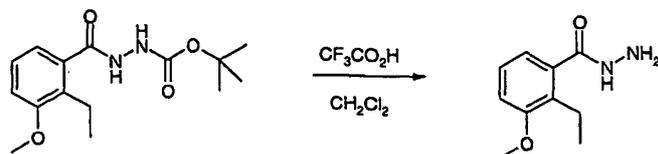
Se disolvió 2,2-dimetil-nonan-3-ol, disponible como una disolución concentrada en éter (aproximadamente 0,23 moles, véase el procedimiento anterior), en aproximadamente 350 ml de CH<sub>2</sub>Cl<sub>2</sub> en un matraz de fondo redondo de 500 ml. Con agitación vigorosa y enfriamiento externo, se añadió lentamente clorocromato de piridinio (PCC, 76,6 g, 0,355 mol). La mezcla de reacción se volvió negra y comenzó a calentarse. Se agitó la reacción durante la noche a temperatura ambiente y se decantó el sobrenadante de una suspensión espesa negra que se había formado. Se extrajo la suspensión espesa con aproximadamente 40 ml de hexano, que se combinaron con la disolución en CH<sub>2</sub>Cl<sub>2</sub>. Se aplicó esta mezcla directamente a una columna de gel de sílice de 100 g y se eluyó en primer lugar con CH<sub>2</sub>Cl<sub>2</sub>/hexano y después con el 10% de acetato de etilo en hexano proporcionando 35,8 g de producto bruto de color verde. Volvió a someterse este material a cromatografía; la elución con hexano y evaporación cuidadosa a 25°C en un evaporador rotatorio proporcionaron 26,1 g de 2,2-dimetil-nonan-3-ona (rendimiento del 67%). <sup>1</sup>H-RMN (500 MHz, CDCl<sub>3</sub>), δ (ppm): 2,47 (t, 2H), 1,28 (a, 8H), 0,89 (t, 3H).



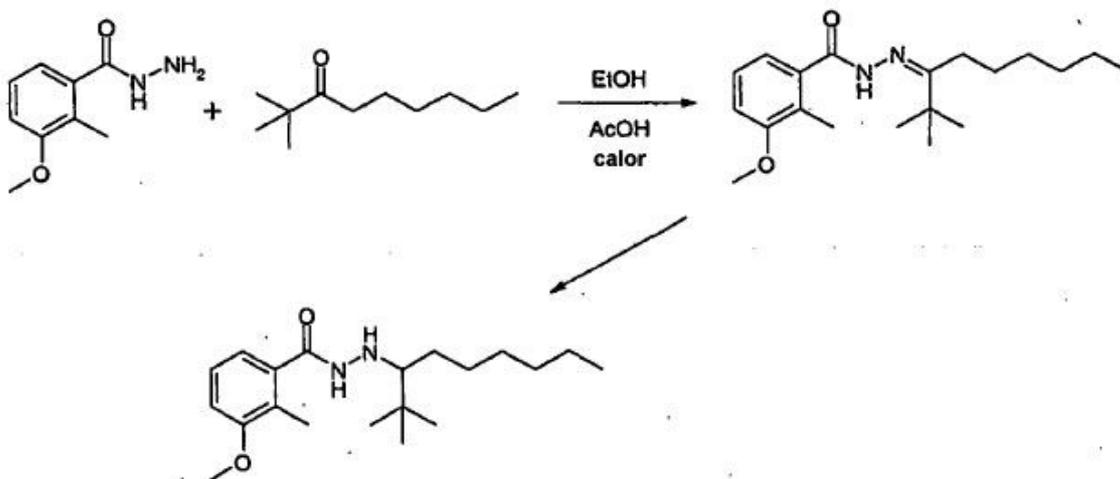
A un matraz de 3 bocas, de 500 ml, equipado con agitación magnética y enfriado en un baño de hielo y agua, se le añadieron 50 ml de  $\text{CH}_2\text{Cl}_2$ , seguido por una disolución de 32 g (231 mmol) de  $\text{K}_2\text{CO}_3$  disueltos en 80 ml de agua, y 24 g (479 mmol) de hidrato de hidrazina. A lo largo de un periodo de 30-60 minutos, se añadió una disolución de 20 g (108,3 mmol) cloruro de 2-etilo-3-metoxibenzoilo disueltos en 100 ml de  $\text{CH}_2\text{Cl}_2$ , mientras se mantenía la temperatura por debajo de  $10^\circ\text{C}$ . Se agitó la mezcla de reacción durante una hora adicional, tiempo durante el cual se formó un precipitado. Se recogió el precipitado y se agitó con una mezcla de  $\text{CH}_2\text{Cl}_2/\text{CHCl}_3$ ; se separó la fase líquida de los sólidos restantes y se eliminó el disolvente a vacío dejando 5,85 g de producto bruto de hidrazida. Mientras tanto, se transfirió la disolución en  $\text{CH}_2\text{Cl}_2$  original a un embudo de decantación, se diluyó con  $\text{CHCl}_3$  y se agitó con agua. Se eliminó la fase orgánica, se lavó de nuevo con agua, se secó y se evaporó el disolvente dejando un residuo sólido. Esto se lavó exhaustivamente con hexano y se filtró proporcionando 6,05 g adicionales de producto bruto de hidrazida. Se recrystalizó la hidrazida en bruto combinada en éter caliente o mezclas de acetato de etilo/hexano proporcionando 10,04 g de hidrazida del ácido 3-metoxi-2-etil-benzoico:  $^1\text{H-RMN}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,2 (t, 1H), 6,95 (s. a., 1H), 6,9 (m, 2H), 4,15 (s. a., 2H), 3,84 (s, 3H), 2,27 (s, 3H).



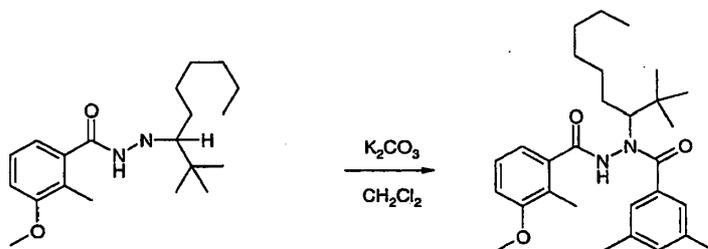
Se agitó carbazato de terc-butilo (80,0 g, 605 mmol) en 800 ml de cloruro de metileno y se enfrió hasta  $0^\circ\text{C}$ . A esto se le añadió la disolución de carbonato de potasio (937 g, 847 mmol). Se disolvió cloruro de 2-etil-3-metoxibenzoilo (132 g, 666 mmol) en 400 ml de cloruro de metileno y se añadió a la mezcla de reacción gota a gota a lo largo de 15 minutos. Se agitó la mezcla a  $0^\circ\text{C}$  durante 15 min, después a temperatura ambiente durante 18 horas. Se diluyó la reacción con más cloruro de metileno y agua. Se separó la fase acuosa. Se lavó la fase orgánica restante con  $\text{HCl}$  1 N, agua, cloruro de sodio saturado, se secó sobre sulfato de magnesio y se evaporó. Se trituró el residuo con hexano dando un sólido blanco, éster terc-butílico del ácido  $\text{N}'$ -(2-etil-3-metoxi-benzoil)-hidrazinacarboxílico (167 g, 567 mmol) con un rendimiento del 94%.  $^1\text{H-RMN}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,4 (s. a., 1H), 7,22 (t, 1H), 7,03 (d. a., 1H), 6,95 (d, 1H), 6,65 (a, 1H), 3,84 (s, 3H), 2,8 (q, 2H), 1:51 (s, 9H), 1,2 (t, 3H).



Se preparó un matraz de fondo redondo con un agitador superior (necesario debido a la gran cantidad de sólido que precipita durante la reacción) y una entrada de nitrógeno. En este matraz se agitó éster terc-butílico del ácido  $\text{N}'$ -(2-etil-3-metoxi-benzoil)-hidrazinacarboxílico (166 g, 564 mmol) en cloruro de metileno (2260 ml). A esta mezcla se le añadió ácido trifluoroacético (217 ml, 2820 mmol). Se agitó la mezcla de reacción a temperatura ambiente durante 18 horas. Se añadieron éter (1000 ml) y hexano (1000 ml) y se agitó la mezcla durante 1 hora. Se separó el precipitado por filtración y se lavó con el 50% de éter/hexano dando un sólido blanco (90,2 g) como la sal de trifluoroacetato del producto. Se combinaron las aguas madres y los lavados, se evaporaron y se trituraron con éter dando una cantidad adicional de un sólido blanco (15,5 g), de nuevo como la sal de trifluoroacetato hidrazida del ácido de 2-etil-3-metoxibenzoico. El rendimiento combinado fue del 60% (105,7 g).  $^1\text{H-RMN}$  (300 MHz,  $\text{DMSO-d}_6$ )  $\delta$  (ppm): 7,3 (t, 1H), 7,17 (d, 1H), 6,95 (d, 1H), 3,85 (s, 3H), 2,65 (q, 2H), 1,1 (t, 3H).  $^{19}\text{F-RMN}$  (282,4 MHz,  $\text{DMSO-d}_6$ )  $\delta$  (ppm): -74,2 (s). Análisis de la base libre de hidrazida del ácido 2-etil-3-metoxi-benzoico:  $^1\text{H-RMN}$  (300 MHz,  $\text{DMSO-d}_6$ )  $\delta$  (ppm): 9,4 (s. a., 1H), 7,2 (t, 1H), 7,05 (d, 1H), 6,85 (d, 1H), 4,45 (a, 2H), 3,85 (s, 3H), 2,6 (q, 2H), 1,1 (t, 3H).



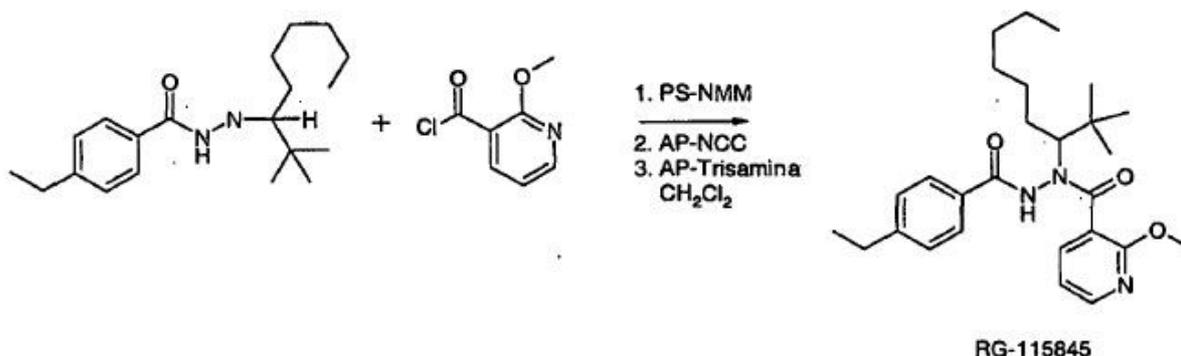
Se disolvieron 3,40 g (20 mmol) de 2,2-dimetilnonan-3-ona en 40 ml de alcohol etílico al 100%. Después se añadieron 3,60 g (20 mmol) de 2-metilo-3-metoxibenzoihidrazida y 20 gotas de ácido acético glacial. Se sometió la mezcla de reacción a reflujo durante 10 horas (requerido para completar la reacción) y se monitorizó mediante CCF. A una disolución de la (1-terc-butil-heptiliden)-hidrazida del ácido 3-metoxi-2-metil-benzoico intermedia, se le añadieron 3,5 ml de ácido acético glacial y 1,89 g (30 mmol) de cianoborohidruro de sodio. Se agitó la reacción a temperatura ambiente durante 24 horas y después se sometió a reflujo durante una hora. Se enfrió la reacción y se añadieron 50 ml de agua y NaOH acuoso al 10% hasta que la reacción fue básica (pH = aproximadamente 14). Se eliminó la mayor parte del alcohol en un evaporador rotatorio y se extrajo el producto con  $\text{CHCl}_3$ . Se secó el extracto acuoso y se concentró hasta obtener un peso constante, proporcionando 6 g de un material viscoso. La CCF (acetato de etilo:hexano 1:1) indicó que el material consistía en cantidades iguales de producto de hidrazida alquilada,  $R_f = 0,6$  (acetato de etilo:hexano 1:1) e hidrazida no alquilada de partida,  $R_f = 0,10$  (acetato de etilo:hexano 1:1). Se obtuvo producto puro mediante cromatografía en columna sobre sílice, eluyéndose  $N'$ -(1-terc-butil-heptil)-hidrazida del ácido 3-metoxi-2-metil-benzoico con el 20-30% de acetato de etilo en hexano.  $^1\text{H-RMN}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,18 (t, 1H), 7,05 (s. a., 1H [NH]), 6,91 (d, 1H), 4,9 (s. a., NH), 3,84 (s, 3H), 2,5 (m, 1H), 2,28 (s, 3H), 1,2-1,8 (múltiples picos amorfos, 10H), 0,97 (s, 9H), 0,9 (t, 3H).



RG-115853

Se disolvieron 164 mg (0,49 mmol) de  $N'$ -(1-terc-butil-heptil)-hidrazida del ácido 3-metoxi-2-metil-benzoico y 82 mg de cloruro de 3,5-dimetilbenzoilo en 10 ml de  $\text{CH}_2\text{Cl}_2$ . Se añadieron 7 ml de  $\text{K}_2\text{CO}_3$  al 25% y se agitó la mezcla de reacción a temperatura ambiente durante la noche. Se monitorizó la reacción mediante CCF. Se separaron las fases, añadiendo  $\text{CH}_2\text{Cl}_2$  y/o agua adicionales según fuera necesario para ayudar en la manipulación. Se secó la fase de  $\text{CH}_2\text{Cl}_2$  y se eliminó el disolvente a vacío proporcionando 240 mg de producto bruto. Se purificó este material mediante cromatografía en columna de gel de sílice, eluyendo con un gradiente en etapas del 5-20% de acetato de etilo en hexano. El producto deseado eluyó en la fracción del 15% de acetato de etilo, CCF  $R_f = 0,62$  (acetato de etilo:hexano 1:1), proporcionando 195 mg de  $N$ -(1-terc-butil-heptil)- $N'$ -(3-metoxi-2-metil-benzoil)-hidrazida del ácido 3,5-dimetil-benzoico.  $^1\text{H-RMN}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,2 (s, NH), 7,05 (s, 2H), 7,03 (m, 1H), 7,01 (s, 1H), 6,83 (d, 1H), 6,25 (d, 1H), 4,67 (d, 1H), 3,81 (s, 3H), 2,28 (s, 6H), 1,79 (s, 3H), 1,55 (m. a., 2H), 1,3 (a, 8H), 1,34 + 1,07 (s+s, 9H), 0,089 (s. a., 3H).

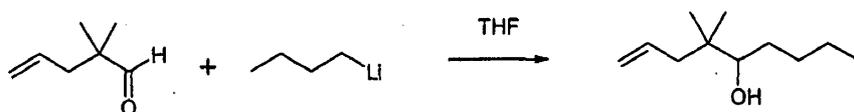
### 1.2 Preparación de RG-115845 (referencia)



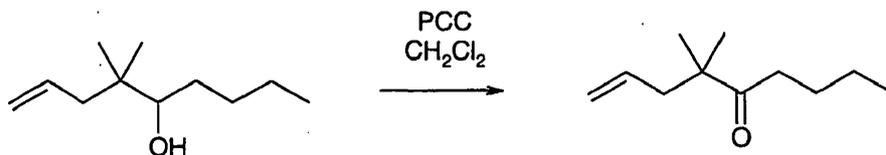
En un vial de 25 ml, se disolvieron 123,1 mg (0,388 mmol) de N'-(1-terc-butil-heptil)-hidrazida del ácido 4-etil-benzoico en 5 ml de CH<sub>2</sub>Cl<sub>2</sub>. Se añadieron 3 equivalentes de PS-NMM (poliestireno-SO<sub>2</sub>NH(CH<sub>2</sub>)<sub>3</sub>-morfolina, Argonaut Tech.), seguido por 3 ml de CH<sub>2</sub>Cl<sub>2</sub> para crear una suspensión que podía agitarse. Se añadió un equivalente de cloruro de ácido y se agitó la mezcla de reacción durante la noche. Al día siguiente, se añadieron 2 equivalentes de resina de AP-NCO (eliminador de isociano, Argonaut Tech.) y 2 equivalentes de AP-trisamina (poliestireno-CH<sub>2</sub>NHCH<sub>2</sub>CH<sub>2</sub>NH(CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>)<sub>2</sub>, Argonaut Tech.) con 3 ml de CH<sub>2</sub>Cl<sub>2</sub>. Se agitó la mezcla durante cuatro horas y se filtraron las resinas. Se analizó la reacción mediante CCF, se eliminó el disolvente y se secó el residuo a vacío. Tras la cromatografía en columna de gel de sílice usando un gradiente del 0-100% de éter en hexano, se aislaron 107,5 mg de N-(1-terc-butil-heptil)-N'-(4-etil-benzoil)-hidrazida del ácido 2-metoxi-nicotínico.

Se prepararon RG-115843, RG-115844 y RG-115854 de una manera análoga a partir de cloruros de ácido e hidrazidas de partida correspondientes.

### 1.3 Preparación de RG-115878 (referencia)

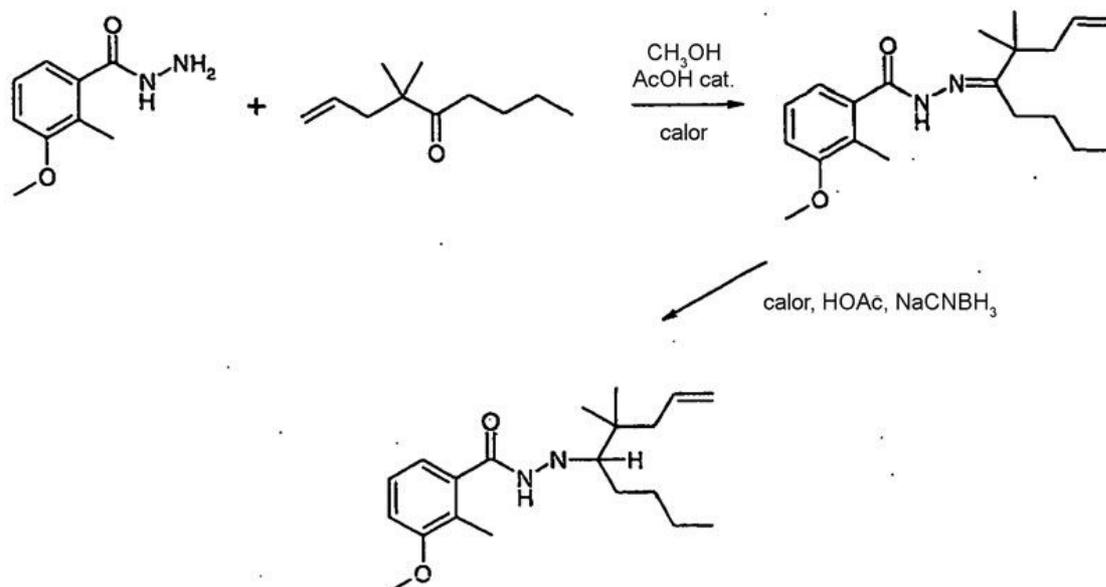


Se añadieron 15 g (133 mmol) de 2,2-dimetil-4-pentenal a 600 ml de THF en un matraz de 2 bocas, de 2 l, se purgó con N<sub>2</sub> y se selló con tapones de caucho. Se enfrió la mezcla de reacción hasta -60°C en un baño de nieve carbónica/acetona. Se añadió una disolución de butil-litio (2,5 M en hexano, 64,4 ml, 160,7 mmol), 4-5 ml cada vez, desde una jeringa de vidrio de 20 ml. Se mantuvo la temperatura de la reacción a -65°C durante la adición y posteriormente se agitó durante una hora adicional. Se permitió que la mezcla de reacción se calentara hasta -5°C a lo largo de 1 hora, se enfrió de nuevo hasta -60°C y se extinguió lentamente con disolución de cloruro de amonio (10,5 g/200 ml de agua, 200 mmol). Se evaporó el THF usando un evaporador rotatorio, manteniendo la temperatura del baño de agua fijada a 30°C. En primer lugar se extrajo la disolución acuosa resultante con CH<sub>2</sub>Cl<sub>2</sub> (3 x 200 ml) y después con éter (1 x 200 ml). Se combinaron los extractos orgánicos, se secaron y se concentraron en un evaporador rotatorio (baño de agua a 30°C) proporcionando 22,64 g de 4,4-dimetil-non-1-en-5-ol, CCF Rf = 0,67 (acetato de etilo:hexanos 1:1, visualización mediante I<sub>2</sub>). <sup>1</sup>H-RMN (CDCl<sub>3</sub>, 500 MHz) δ (ppm): 5,86 (m, 1H), 5,06 (m, 1H), 5,03 (s, 1H), 3,26 (d, 1H), 2,15 (m, 1H), 1,95 (m, 1H), 1,5 (m, 2H), 1,3 (m, 4H), 0,95 (t, 3H), 0,90 (s, 3H), 0,89 (s, 3H).



Se añadieron 17,50 g (102,9 mmol) de 4,4-dimetil-non-1-en-5-ol y 300 ml de CH<sub>2</sub>Cl<sub>2</sub> a un matraz de 500 ml. Se enfrió la reacción en agua con hielo. Mientras se agitaba vigorosamente, se añadieron 33,29 g (154 mmol) de clorocromato de piridinio en porciones (baño de hielo). Se agitó la reacción a temperatura ambiente durante 24 horas, tiempo durante el cual se volvió negra y se formó una suspensión espesa en el fondo del matraz. Se decantó CH<sub>2</sub>Cl<sub>2</sub> de color negro-marrón y se lavó la suspensión espesa dos veces con CH<sub>2</sub>Cl<sub>2</sub>. Se purificó el producto de reacción mediante cromatografía en columna sobre sílice. Se hizo pasar la mezcla de CH<sub>2</sub>Cl<sub>2</sub> marrón a través de una columna de sílice seca y se eluyó producto limpio como disolución en CH<sub>2</sub>Cl<sub>2</sub>, que tras la evaporación del disolvente proporcionó 12,69 g de producto 4,4-dimetil-non-1-en-5-ona. La elución con el 5% de acetato de etilo/hexano proporcionó 1,40 g adicionales (rendimiento del 81%). CCF del producto puro dio un Rf = 0,71 (acetato

de etilo:hexano 1:1, visualizado mediante  $I_2$ ).  $^1\text{H-RMN}$  ( $\text{CDCl}_3$ , 500 MHz)  $\delta$  (ppm): 5,7 (m, 1H), 5,05 (m, 2H), 2,45 (t, 2H), 2,27 (d, 2H), 1,5 (m, 2H), 1,3 (m, 2H), 1,12, (s, 3H), 0,90 (t, 3H).



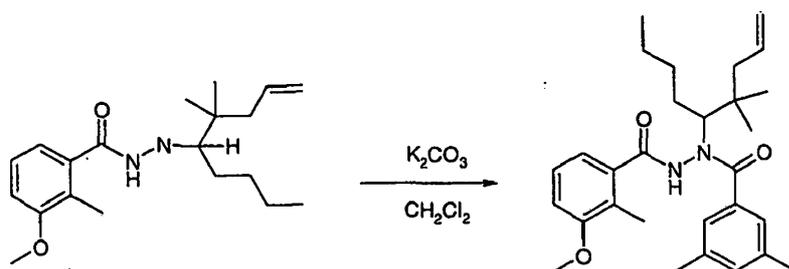
5

Se cargó un matraz de fondo redondo de 50 ml con 10,62 g (59 mmol) de 2-metil-3-metoxibenzoil hidrazida y 10,0 g (59,5 mmol) de 4,4-dimetilnonan-1-en-5-ona y 150 ml de metanol. Se añadieron veinte gotas de ácido acético glacial como catalizador y se sometió la mezcla de reacción a reflujo durante 9 horas y se agitó a temperatura ambiente durante 48 horas. No se aisló la hidrazona intermedia, pero la CCF indicó que se obtuvo el 30% de producto ( $R_f = 0,58$ , hex:acetato de etilo 1:1). A la mezcla de reacción se le añadieron 4 ml de ácido acético y 4,72 g (75 mmol) de  $\text{NaCNBH}_3$  y se sometió la reacción a reflujo durante una hora. Se transfirió la mezcla de reacción a un vaso de precipitados de 600 ml y se añadieron 100 ml de agua, seguido por  $\text{NaOH}$  al 10% hasta que fue básica. Se eliminó la mayor parte del alcohol por evaporación y se extrajo la mezcla restante con acetato de etilo proporcionando 12,2 g de residuo. La cromatografía sobre gel de sílice, eluyendo con el 15-25% de acetato de etilo/hexano, proporcionó el producto bruto, hidrazida (2,86 g, rendimiento del 15%), dos puntos mediante CCF,  $R_f = 0,47$  y  $0,55$  (acetato de etilo/hexano 1:1). La nueva cromatografía proporcionó producto puro mediante elución con un gradiente de acetato de etilo/hexano; la fracción del 10% de acetato de etilo/hexano dio N-(1-butil-2,2-dimetil-pent-4-enil)-hidrazida del ácido 3-metoxi-2-metil-benzoico purificada,  $R_f = 0,56$ , (acetato de etilo/hexano 1:1).  $^1\text{H-RMN}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,18 (t, 1H), 7,03 (s, 1H, NH), 6,90 (d, 1H), 5,87 (m, 1H), 5,04 (m, 2H), 4,9 (s. a., 1H NH), 3,84 (s, 3H), 2,55 (m, 1H), 2,28 (s, 3H), 2,16 (m, 1H), 2,06 (m, 1H), 1,7 (m, 1H), 1,6 (m, 1H), 1,45 (m, 1H), 1,3 (m, 3H), 0,96 (s, 3H), 0,92 (s, 3H), 0,92 (t, 3H).

10

15

20



RG-115878

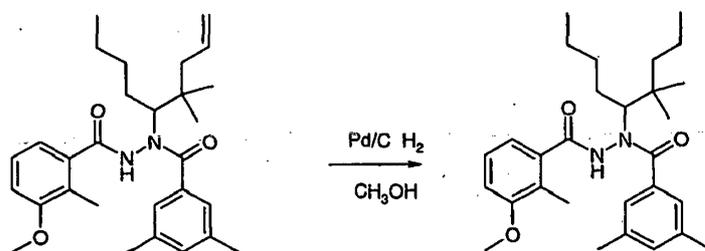
25

Se disolvieron 1,8 g (5,42 mmol) de N'-(1-butil-2,2-dimetil-pent-4-enil)-hidrazida del ácido 3-metoxi-2-metil-benzoico y 1,34 g (8 mmol) de cloruro de 3,5-dimetilbenzoílo en 50 ml de  $\text{CH}_2\text{Cl}_2$ . Se añadieron 20 ml de  $\text{K}_2\text{CO}_3$  al 20% (15 mmol) y se agitó la mezcla de reacción a temperatura ambiente durante la noche. Se separaron las fases, añadiendo  $\text{CH}_2\text{Cl}_2$  y/o agua adicionales según fuera necesario para ayudar en la manipulación. Se secó la fase de  $\text{CH}_2\text{Cl}_2$  y se eliminó el disolvente a vacío proporcionando 1,65 g de un sólido vítreo. Se purificó este material mediante cromatografía en columna de gel de sílice, eluyendo con un gradiente en etapas del 5-25% de acetato de etilo en hexano. El producto deseado eluyó en la fracción del 12% de acetato de etilo, CCF  $R_f=0,59$  (acetato de etilo:hexano 1:1), rendimiento = 1,0 g. La nueva cromatografía de nuevo con un gradiente en etapas de acetato de etilo/hexano proporcionó una muestra más pura de la diacilhidrazida prevista, N-(1-butil-2,2-dimetil-pent-4-enil)-N'-(3-metoxi-2-metil-benzoil)-hidrazida del ácido 3,5-dimetil-benzoico, como un sólido blanco.  $^1\text{H-RMN}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$

30

(ppm): 7,2 (s. a., NH), 7,11 (m, 1H), 7,1 (s, 2H), 7,02 (s, 1H), 6,87 (d, 1H), 6,3 (d, 1H), 5,92 (m, 1H), 5,1 (m, 2H), 4,77 (m, 1H), 3,78 (s, 3H), 2,35 (d, 1H), 2,28 (s, 6H), 2,15 (d, 1H), 1,77 (s, 3H), 1,2-1,6 (m. a., 6H), 1,05 (s, 3H), 1,03 (s, 3H), 0,94 (t, 3H).

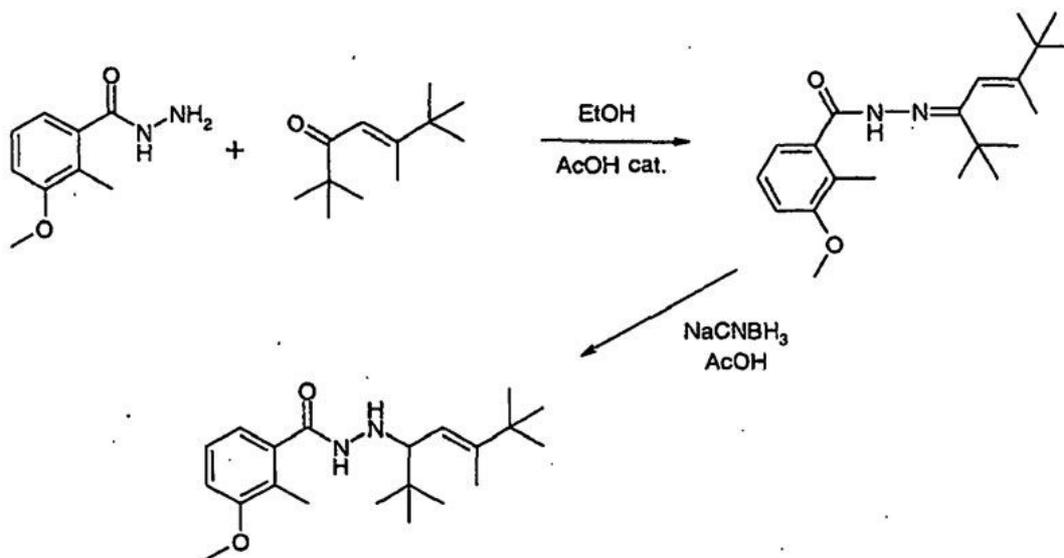
5 1.4 Preparación de RG-115877 (referencia)



RG-115877

10 Se hidrogenó RG-115878 (240 mg) en un agitador Parr en metanol bajo una atmósfera de H<sub>2</sub> usando catalizador de paladio sobre carbón. Se aisló N-(1-butil-2,2-dimetil-pentil)-N'-(3-metoxi-2-metil-benzoil)-hidrazida del ácido 3,5-dimetil-benzoico como un sólido blanco: <sup>1</sup>H-RMN (500 MHz, CDCl<sub>3</sub>) δ (ppm): 7,07 (m, 3H [NH]), 7,00 (m, 2H), 6,83 (d, 1H), 6,29 (d, 1H), 4,73 (d, 1H), 3,79 (s, 3H), 2,29 (s, 6H), 1,78 (s, 3H), 1,58 (a, 4H), 1,38 (a, 6H), 1,03 (m, 6H), 0,96 (m, 6H).

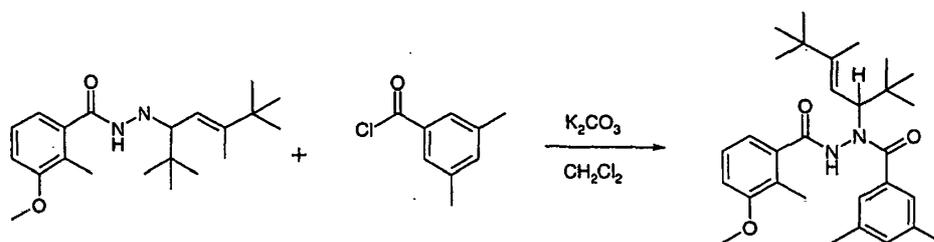
15 Preparación de RG-115855 (referencia)



20 Se pesaron 1,84 g (100 mmol) de 2,2,5,6,6-pentametilhepten-3-ona (Lancaster Synthesis) en un matraz de fondo redondo de 200 ml. Se añadieron 1,80 (10 mmol) de 2-metil-3-metoxibenzoilhidrazina, 50 ml de etanol y 20 gotas de ácido acético glacial. Se sometió la reacción a reflujo con agitación durante 24 horas. No se aisló el producto de hidrazona, sino que se sometió directamente a reducción.

25 A la mezcla de reacción de N'-(1-terc-butil-3,4,4-trimetil-pent-2-enil)-hidrazida del ácido 3-metoxi-2-metil-benzoico se le añadieron 3 ml de ácido acético glacial y 950 mg (14,5 mmol) de cianoborohidruro de sodio al 95%, y se agitó la reacción durante la noche a temperatura ambiente. Se evaporó la mayor parte del etanol y se añadieron 50 ml de agua y se basificó la mezcla a pH 14 con NaOH al 10%. Se evaporó el resto del etanol y se extrajo el producto con cloroformo. La CCF del residuo indicó la presencia del producto de hidrazida previsto (R<sub>f</sub>=0,55, acetato de etilo:hexano 1:1), hidrazida de partida (R<sub>f</sub>=0,08) y varios productos minoritarios. Se obtuvo la hidrazida pura tras cromatografía en gel de sílice en gradiente (acetato de etilo/hexano); el producto eluyó con el 20% de acetato de etilo en hexano. La concentración a vacío proporcionó 515 mg de N'-(1-terc-butil-3,4,4-trimetil-pent-2-enil)-hidrazida del ácido 3-metoxi-2-metil-benzoico cristalina blanca (rendimiento del 15%). <sup>1</sup>H-RMN (500 MHz, CDCl<sub>3</sub>) δ (ppm): 7,2 (t, 1H), 6,93 (s, 1H [NH]), 6,88 (d, 1H), 5,24 (d, 1H), 5,2 (s. a., 1H), 3,83 (s, 3H), 3,63 (d, 1H), 2,26 (s, 3H), 1,69 (s, 3H), 1,07 (s, 9H), 0,99 (s, 9H).

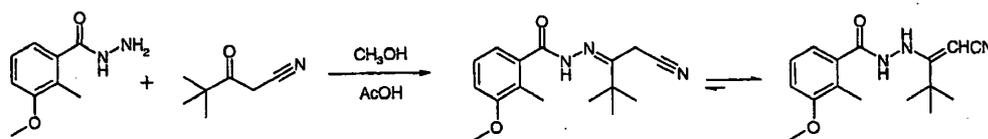
35



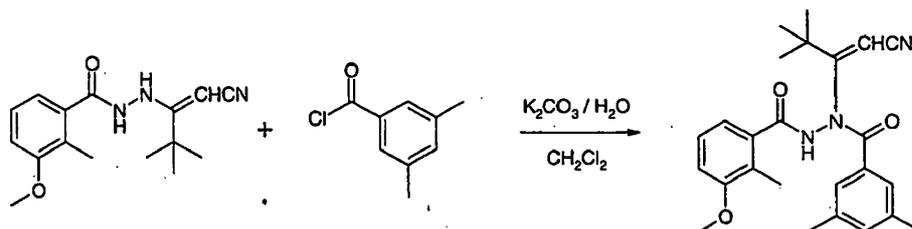
RG-115855

Se añadieron 219 mg (0,63 mmol) de N'-(1-terc-butil-3,4,4-trimetil-pent-2-enil)-hidrazida del ácido 3-metoxi-2-metil-benzoico, 230 mg (1,36 mmol) de cloruro de 3, 5-dimetilbenzoilo, 7 ml de una disolución acuosa al 25% de K<sub>2</sub>CO<sub>3</sub> y 7 ml de CH<sub>2</sub>Cl<sub>2</sub> a un vial de 20 ml. Se agitó la reacción a temperatura ambiente durante 24 horas. Se transfirió la mezcla a un embudo de decantación con CH<sub>2</sub>Cl<sub>2</sub> y se añadieron 45 ml de K<sub>2</sub>CO<sub>3</sub> acuoso al 17%. Se separó la fase de CH<sub>2</sub>Cl<sub>2</sub>, se extrajo la fase acuosa con 2 x porciones de 50 ml de CH<sub>2</sub>Cl<sub>2</sub>. Se combinaron las fases orgánicas, se secaron y se concentraron hasta sequedad a vacío proporcionando 0,53 g de residuo, R<sub>f</sub> = 0,67 (acetato de etilo:hexano 1:1). Se sometió el residuo a cromatografía sobre gel de sílice usando un gradiente de acetato de etilo/hexano y el producto deseado se eluyó con el 10-11% de acetato de etilo en hexano proporcionando 0,21 g de N'-(1-terc-butil-3,4,4-trimetil-pent-2-enil)-N'-(3,5-dimetil-benzoil)-hidrazida del ácido 3-metoxi-2-metil-benzoico pura. <sup>1</sup>H-RMN (500 MHz, CDCl<sub>3</sub>) δ (ppm): 7,2 (1H, [NH]), 7,12 (m, 1H), 7,1 (s. a., 2H), 6,95 (s, 1H), 6,85 (d, 1H), 6,6 (d, 1H), 5,55 (d, 1H), 5,45 (d, 1H), 3,8 (s, 3H), 2,3 (s, 6H), 1,95 (s, 3H), 1,6 (s, 3H), 1,07 (s, 9H), 1,02 (s, 9H).

#### 1.6 Preparación de RG-115860 (referencia)

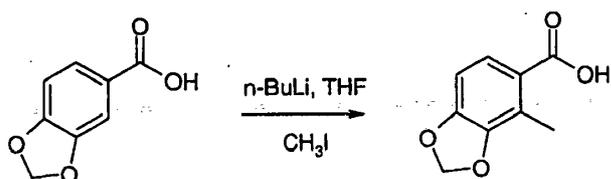


Se añadieron terc-butil-cianometilcetona (1,75 g, 14 mmol, Lancater Synthesis) y 2-metil-3-metoxibenzoilhidrazida (2,52 g, 14 mmol) a 20 ml de metanol acidificado con 20 gotas de ácido acético glacial. Se agitó la mezcla de reacción a temperatura ambiente durante la noche. Se eliminó el disolvente a vacío, volvió a disolverse el residuo en CH<sub>2</sub>Cl<sub>2</sub> y se extrajo la mezcla resultante con K<sub>2</sub>CO<sub>3</sub> acuoso. Se secó la fase orgánica y se eliminó el disolvente a vacío. La cromatografía en columna sobre gel de sílice, eluyendo con un gradiente del 10-35% de acetato de etilo/hexano, proporcionó 0,69 g de N'-(1-terc-butil-2-ciano-vinil)-hidrazida del ácido 3-metoxi-2-metil-benzoico purificada, así como una cantidad comparable de material algo menos puro, CCF: R<sub>f</sub>=0,56 (acetato de etilo:hexano 1:1). <sup>1</sup>H-RMN (500 MHz, CDCl<sub>3</sub>) δ (ppm): 7,18 (t, 1H), 7,07 (d, 1H), 6,91 (d, 1H), 5,56 (s. a., 2H [NH]), 5,27 (s, 1H), 3,82 (s, 3H), 2,13 (s, 3H), 1,145 (s, 9H).



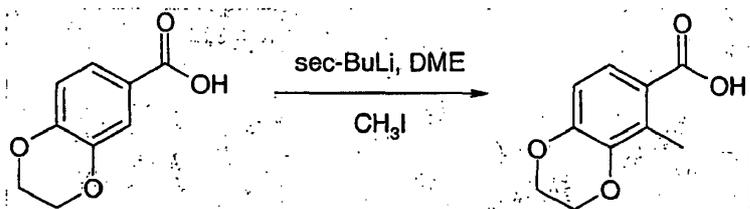
Se agitaron 340 mg (1,18 mmol) de N'-(1-terc-butil-2-ciano-vinil)-hidrazida del ácido 3-metoxi-2-metil-benzoico, 290 mg (1,7 mmol) de cloruro de 3,5-dimetilbenzoilo, 3 ml de una disolución acuosa al 25% de K<sub>2</sub>CO<sub>3</sub> y 5 ml de CH<sub>2</sub>Cl<sub>2</sub> a temperatura ambiente durante 2 días. Se diluyó la mezcla con agua y CH<sub>2</sub>Cl<sub>2</sub>. Se separó la fase orgánica y se extrajo la fase acuosa con CH<sub>2</sub>Cl<sub>2</sub> adicional. Se combinaron las fases orgánicas, se secaron y se concentraron hasta sequedad a vacío proporcionando 0,72 g de semisólido, R<sub>f</sub> = 0,69 (acetato de etilo:hexano 1:1, impureza principal a R<sub>f</sub> = 0,63). Se sometió el residuo a cromatografía sobre gel de sílice usando un gradiente del 2-20% de acetato de etilo/hexano y el producto deseado se eluyó con el 4% de acetato de etilo en hexano proporcionando 106 mg de N'-(1-terc-butil-2-ciano-vinil)-N'-(3,5-dimetilbenzoil) hidrazida del ácido 3-metoxi-2-metil-benzoico purificada. <sup>1</sup>H-RMN (500 MHz, CDCl<sub>3</sub>) δ (ppm): 7,57 (s, 2H), 7,25 (t, 1H), 7,2 (s, 1H), 7,1 (d, 1H [C=CH]), 7,02 (s, 1H), 6,99 (d, 1H), 3,87 (s, 3H), 2,4 (s, 6H), 2,16 (s, 3H), 1,24 (s, 9H).

#### 1.7 Preparación de RG-115790

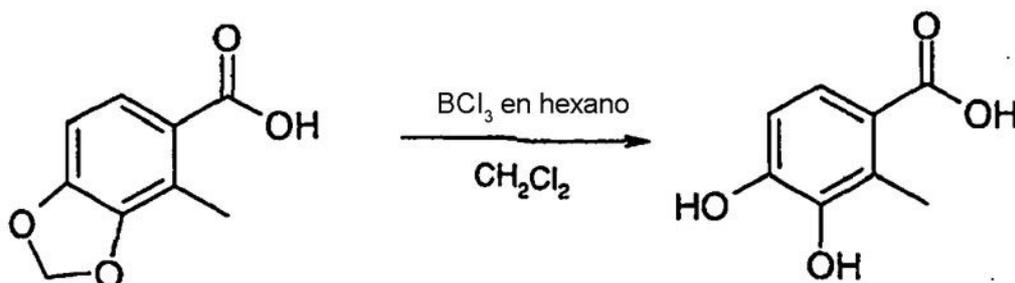


5 A un matraz de fondo redondo equipado con un agitador superior y una entrada de nitrógeno se le añadieron ácido piperonílico (50,0 g, 301 mmol) y tetrahidrofurano (753 ml). Esto se enfrió hasta  $-75^{\circ}\text{C}$  en un baño de nieve carbónica-acetona. Se añadió N-butil-litio (1,6 M en hexanos) (414 ml, 662 mmol) gota a gota manteniendo la temperatura de la reacción a o por debajo de  $-65^{\circ}\text{C}$ . Cuando se completó la adición se retiró el baño de enfriamiento y se permitió que la reacción se calentara hasta  $-20^{\circ}\text{C}$ . Se devolvió la reacción al baño y se enfrió hasta al menos  $-60^{\circ}\text{C}$ , momento en el cual se añadió yodometano (37,5 ml, 602 mmol) gota a gota. Se retiró el enfriamiento y se permitió que la reacción se calentara hasta  $10^{\circ}\text{C}$ , momento en el cual se colocó la reacción en un baño de hielo y se agitó a  $0^{\circ}\text{C}$  durante 30 minutos. Se extinguió la reacción mediante adición de HCl 1 N y se eliminó el tetrahidrofurano mediante evaporación. Se suspendió el residuo en HCl 1 N y se filtró. Se lavó el filtrado con agua y se secó a  $50-60^{\circ}\text{C}$  a vacío dando un sólido de color amarillo pálido, ácido 4-metil-benzo[1,3]dioxol-5-carboxílico (52,8 g, 293 mmol) con un rendimiento del 97%.  $^1\text{H-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): 2,44 (s, 3H), 6,10 (s, 2H), 6,77 (d, 1H), 7,64 (d, 1H). CCF Rf = 0,54 (acetato de etilo/hexano 1:1, ácido piperonílico = 0,41).

15 Notas: Se necesita agitación superior, a diferencia de magnética, ya que la sal de carboxilato forma un precipitado pesado. La adición del primer equivalente de butil-litio es bastante exotérmica (desprotonación del ácido carboxílico). La velocidad de adición de butil-litio puede aumentarse sustancialmente para el segundo equivalente. Temperaturas de reacción de  $-60^{\circ}\text{C}$  o más antes de completarse la adición de butil-litio dan como resultado la formación de butilcetona en cantidades crecientes. La adición de yodometano es exotérmica, con una temperatura inicial de aproximadamente  $-10$  a  $0^{\circ}\text{C}$ , momento en el cual se produce un aumento de temperatura de  $10-15^{\circ}\text{C}$ .

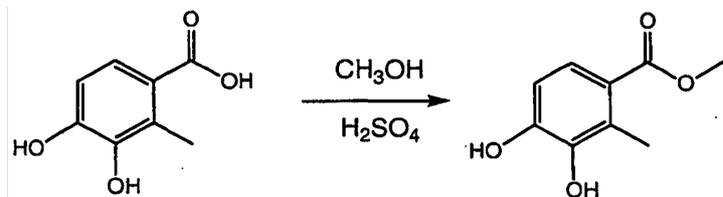


25 A un matraz de fondo redondo equipado con agitación magnética, un embudo de adición y una entrada de nitrógeno, se le añadieron ácido benzo(1,4)dioxan-6-carboxílico (18,00 g, 99,91 mmol) y 1,2-dimetoxietano (667 ml). Se enfrió esta mezcla hasta  $-75^{\circ}\text{C}$  en un baño de nieve carbónica-acetona. A esto se le añadió sec-butil-litio 1,3 M en ciclohexano (230,6 ml, 299,7 mmol) a lo largo de 1 hora, manteniendo la temperatura de reacción por debajo de  $-60^{\circ}\text{C}$ . Se retiró la reacción del baño de enfriamiento, se permitió que se calentara hasta  $-20^{\circ}\text{C}$  y posteriormente se agitó a  $-20^{\circ}\text{C}$  durante 45 min. Se enfrió la reacción hasta  $-50^{\circ}\text{C}$  y se añadió yodometano (15,6 ml, 249,8 mmol). Se retiró de nuevo la reacción del baño de enfriamiento, se permitió que se calentara hasta  $-20^{\circ}\text{C}$  y se agitó a esta temperatura durante 45 min. Se retiró todo el enfriamiento y se agitó la reacción a temperatura ambiente durante 16 horas. Se extinguió la reacción mediante adición de unos pocos ml de HCl 1 N (ac.) y se eliminó el disolvente mediante evaporación. Se volvió el residuo sustancialmente ácido mediante la adición de HCl 1 N acuoso. Se filtró el precipitado resultante y se lavó con agua dando un sólido de color marrón claro, ácido 5-metil-benzo(1,4)dioxan-6-carboxílico, (6,60 g, 33,9 mmol) con un rendimiento del 34%.  $^1\text{H-RMN}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,62 (d, 1H), 6,8 (d, 1H), 4,30 (s. a., 4H), 2,52 (s, 3H).

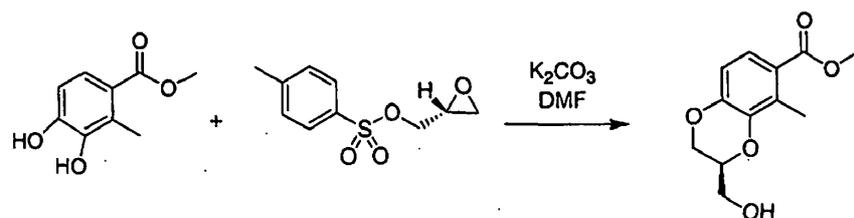


40 A un matraz de fondo redondo equipado con agitación magnética, una trampa de gas NaOH al 25% y una entrada de nitrógeno, se le añadieron ácido 2-metil-piperonílico (26,0 g, 144 mmol) y cloruro de metileno (482 ml). A esto se le añadió la disolución de tricloruro de boro. Se agitó la mezcla a temperatura ambiente durante 3 horas. Se extinguió la reacción mediante la adición cuidadosa de 500 ml de agua. Atención: esta etapa provoca espumación y

genera grandes cantidades de HCl. Se agitó la reacción durante 30 min. Se separaron las fases, se comprobó la CCF de la fase orgánica para confirmar la ausencia de producto y posteriormente se descartó la fase orgánica. Se extrajo la fase acuosa 3 veces con acetato de etilo. Se lavaron los extractos combinados una vez con agua, una vez con disolución saturada de NaCl, se secaron sobre sulfato de sodio, se filtraron y se evaporaron dando un sólido de color tostado, ácido 3,4-dihidroxi-2-metilbenzoico, (20,0 g, 119 mmol) con un rendimiento del 82%. Nota: También se ha usado  $\text{BCl}_3$  en heptano de manera igual de satisfactoria.  $^1\text{H-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): 2,51 (s, 3H), 6,76 (d, 1H), 7,45 (s+d solapantes, 1H+1H), 9,06 (s, 1H). CCF Rf= 0,20 (acetato de etilo/hexano 1:1).

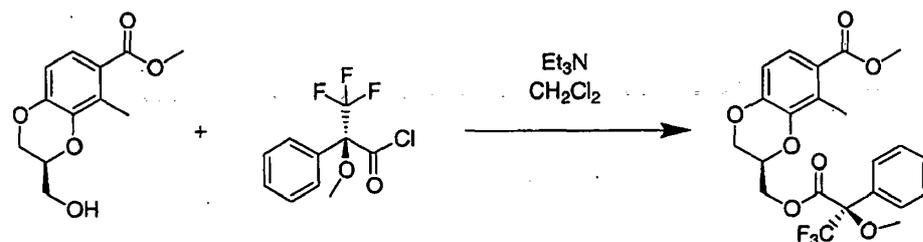


A un matraz de fondo redondo equipado con agitación magnética y un condensador de reflujo se le añadieron ácido 3,4-dihidroxi-2-metilbenzoico (23,3 g, 139 mmol), metanol (139 ml) y ácido sulfúrico concentrado (0,77 ml, 13,9 mmol). Se calentó esta mezcla a reflujo durante 18 horas. Se enfrió la reacción hasta temperatura ambiente y se evaporó el metanol. Se llevó el residuo a éter y se lavó una vez con disolución saturada de bicarbonato de sodio, una vez con disolución saturada de cloruro de sodio, se secó sobre sulfato de sodio, se filtró y se evaporó dando un sólido marrón, 3,4-dihidroxi-2-metilbenzoato de metilo, (20,9 g, 114 mmol) con un rendimiento del 83%.  $^1\text{H-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): 2,47 (s, 3H), 3,79 (s, 3H), 6,78 (d, 1H), 7,46 (d, 1H); CCF: Ref=0,54 (acetato de etilo/hexano 1:1).



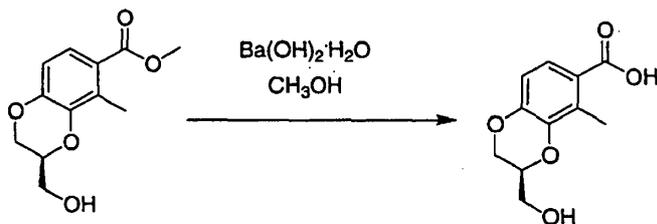
A un matraz de fondo redondo equipado con agitación magnética y una entrada de nitrógeno se le añadieron 3,4-dihidroxi-2-metilbenzoato de metilo (24,0 g, 132 mmol), tosilato de (2R)-(-)-glicidilo (30,1 g, 132 mmol), carbonato de potasio (21,8 g, 158 mmol) y DMF (264 ml). Se calentó esta mezcla hasta  $60^\circ\text{C}$  durante 5 horas. Se enfrió la mezcla hasta temperatura ambiente, se diluyó con éter y se lavó una vez con agua. Se sometió el lavado acuoso a retroextracción una vez con éter y se combinaron las disoluciones en éter. Entonces se lavaron estas disoluciones orgánicas 3 veces con disolución saturada de cloruro de sodio, se secaron sobre sulfato de magnesio, se filtraron y se evaporaron dando un aceite marrón. Se sometió este aceite a cromatografía sobre gel de sílice eluyendo con el 25% de éter en hexanos dando un aceite amarillo, éster metílico del ácido 3-hidroxi-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico, (24,6 g, 103 mmol) con un rendimiento del 78%.  $^1\text{H-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): 2,42 (s, 2H), 3,80 (s, 3H), 3,83 (m, 1H), 4,12 (dd, 1H), 4,23 (m, 2H), 4,42 (dd, 1H), 6,75 (d, 1H), 7,42 (d, 1H); CCF Rf = 0,40 (acetato de etilo / hexano 1:1).

Notas: (a) puede usarse tosilato de (2S)-(+)-glicidilo en condiciones idénticas dando la estereoquímica opuesta, (b) la formación de un éster de Mosher de este compuesto indica la presencia de un único regio y estereoisómero mediante  $^{19}\text{F-RMN}$ , (c) la determinación de la estructura cristalina por rayos X de la amida formada a partir de (R)-(+)-1-(1-naftil)etilamina confirmó la regio y estereoquímica indicadas para el producto 3-hidroxi-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico. Delgado, A.; Leclerc, G.; Lobato, C.; Mauleon, D. *Tetrahedron Lett.* 1988, 29(30), 3671.

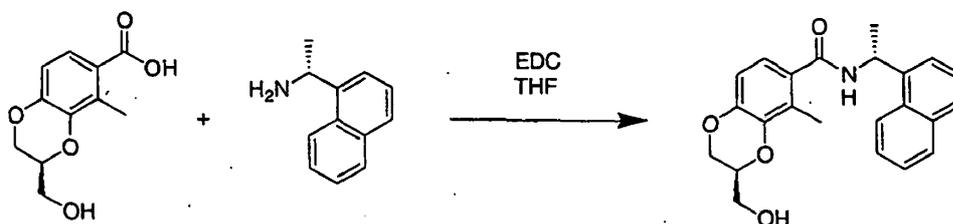


A un matraz de fondo redondo equipado con agitación magnética y una entrada de nitrógeno se le añadieron éster metílico del ácido 3-hidroxi-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico (0,10 g, 0,44 mmol) y cloruro de metileno seco (1,5 ml). Se enfrió la mezcla en un baño de hielo y se añadió trietilamina (0,12 ml, 0,87 mmol) seguido por cloruro de (R)-(-)- $\alpha$ -metoxi- $\alpha$ -trifluorometilfenilacetilo (0,09 ml, 0,48 mmol). Se agitó la mezcla durante 2 horas,

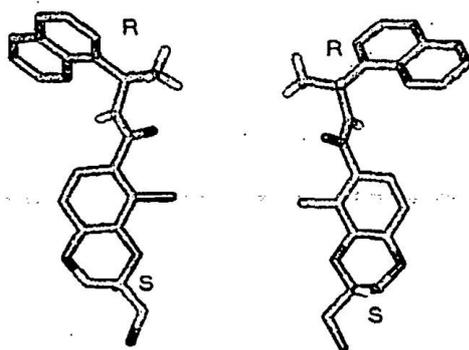
momento en el cual la CCF mostró que la reacción era incompleta. Se añadió más cloruro de (R)-(-)-  $\alpha$ -metoxi- $\alpha$ -trifluorometilfenilacetilo (0,09 ml, 0,48 mmol) y se permitió que la reacción se agitara durante 1 hora a temperatura ambiente. Se diluyó la mezcla con cloruro de metileno, se lavó una vez con HCl 1 N, una vez con bicarbonato de sodio saturado, se secó sobre sulfato de magnesio, se filtró y se evaporó. Se sometió el residuo a cromatografía sobre gel de sílice eluyendo con cloruro de metileno dando un aceite transparente, éster metílico del ácido 5-metil-3-(3,3,3-trifluoro-2-metoxi-2-fenil-propioniloximetil)-2,3-dihidro-benzo[1,4]dioxin-6-carboxílico, para el que no se determinó el rendimiento.  $^1\text{H-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): 2,38 (s, 3H), 3,58 (s, 3H), 3,81 (s, 3H), 4,10 (m, 1H), 4,48 (dd, 1H), 4,65 (m, 2H), 4,86 (m, 1H) 6,78 (d, 1H) 7,47 (m, 5H), 7,60 (d, 1H);  $^{19}\text{F-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): -72,89 (s); tanto  $^1\text{H}$  como  $^{19}\text{F-RMN}$  indican la presencia de tan sólo un estereo y regioisómero; CCF: Rf = 0,62 (acetato de etilo/hexano 1:1).



A un matraz de fondo redondo equipado con agitación magnética se le añadieron éster metílico del ácido 3-hidroxi-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico (6,40 g, 26,9 mmol), hidróxido de bario monohidratado (25,44 g, 134,3 mmol) y metanol (108 ml). Se agitó esta mezcla a temperatura ambiente durante 24 horas. Se evaporó el metanol. Se llevó la suspensión resultante a agua y se lavó una vez con éter. Se acidificó la fase acuosa vertiéndola en HCL concentrado helado. Esto se extrajo dos veces con acetato de etilo. Se combinaron los extractos de acetato de etilo, se secaron sobre sulfato de magnesio, se filtraron y se evaporaron dando un sólido amarillo, ácido 3-hidroxi-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico, (4,82 g, 21,5 mmol) con un rendimiento del 80%.  $^1\text{H-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): 2,45 (s, 3H), 3,85 (m, 2H), 4,12 (m, 1H), 4,24 (m, 1H), 4,41 (m, 1H), 6,76 (d, 1H), 7,51 (d, 1H); CCF: Rf = 0,32 (acetato de etilo/hexano 1:1).



A un matraz de fondo redondo equipado con agitación magnética y una entrada de nitrógeno se le añadieron ácido 3-hidroxi-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico (2,00 g, 8,92 mmol), THF (45 ml), (R)-(+)-1-(1-naftil)etilamina (2,23 ml, 8,92 mmol) y finalmente clorhidrato de 1-(3-dimetilaminopropil)-3-etilcarbodiimida (EDC) (1,53 g, 9,81 mmol). Se agitó esta mezcla a temperatura ambiente durante 36 horas. Se añadieron unas pocas gotas de agua y se agitó la mezcla durante 5 minutos. Se diluyó la reacción con éter y se lavó una vez con HCl 1 N, una vez con bicarbonato de sodio saturado (ac.), de nuevo una vez con HCl 1 N, se secó sobre sulfato de magnesio, se filtró y se evaporó proporcionando (1-naftalen-1-il-etil)-amida del ácido 3-hidroxi-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico. Se obtuvieron cristales de pureza aceptable para cristalografía de rayos X sublimando el material dos veces a vacío a  $60^\circ\text{C}$ . Este análisis de rayos X estableció la regioquímica y la estereoquímica absoluta de los compuestos tal como se muestra, dado que se conoce la configuración del estereocentro de naftiletilamina y no experimenta racemización en las condiciones de la síntesis.  $^1\text{H-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): 1,71 (d, 3H), 2,22 (s, 3H), 3,80 (m, 2H), 4,06 (m, 1H), 4,23 (m, 2H), 4,32 (m, 1H), 6,09 (m, 1H), 6,65 (d, 1H), 6,87 (d, 1H), 7,55 (m, 3H), 7,69 (m, 2H), 7,83 (d, 1H), 7,94 (d, 1H), 8,35 (d, 1H); CCF Rf = 0,12 (acetato de etilo/hexano 1:1).



RG-119097 (referencia)

Datos de cristalografía para:	MSC01712 (RG-119097)
Fecha de entrada:	20-09-2001
Referencia:	No publicado
Síntesis:	RHeoGene
Cristalografía por:	J. C. Huffman (huffman@indiana.edu)
Nombre del compuesto:	desconocido
Fórmula:	C <sub>23</sub> H <sub>23</sub> N <sub>4</sub> O
Fórmula empírica:	C <sub>23</sub> H <sub>23</sub> N <sub>4</sub> O
Color del cristal:	Incoloro
Sistema cristalino:	Monoclínico
Grupo espacial:	P21
Dimensiones de celda:	(a -161°C; 3421 picos)
a=	11,9081(18)
b =	4,8854(7)
c =	16,3819(25)
alfa =	90,00( 0)
beta =	97,758(4)
gamma =	90,00(0)
Z (moléculas/celda):	2
Volumen:	944,31
Densidad calculada:	1,327
Peso molecular:	377,44
Coefficiente de absorción lineal:	0,906
Los errores residuales finales son:	
R(F)=	0,062
Rw(F)=	0,050

Coordenadas fraccionarias XYZ:

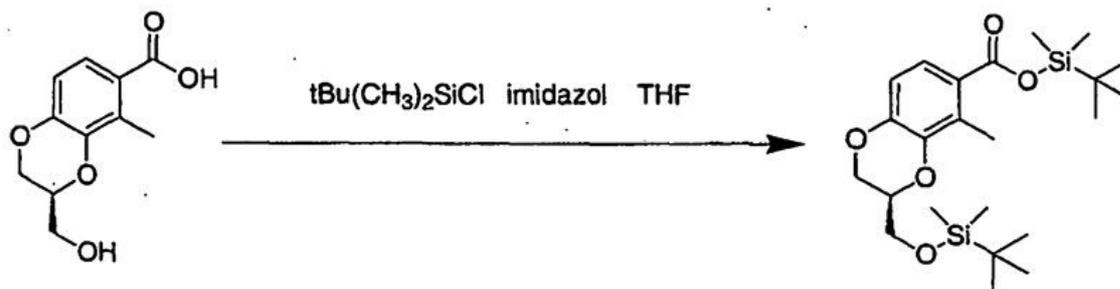
5

C(1)	12796(3)	9729*	835(3)	22
C(2)	11822(3)	8046(11)	622(3)	24
C(3)	11171(4)	8270(12)	-122(3)	31

ES 2 478 622 T3

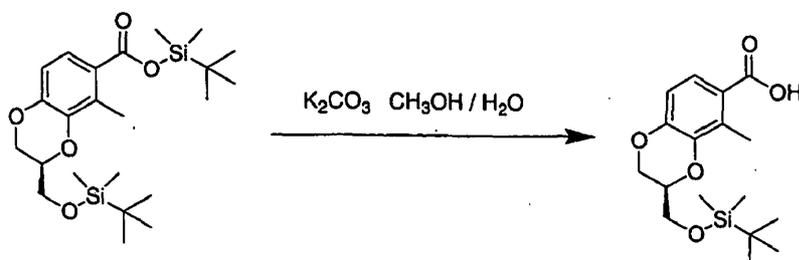
C(4)	11442(4)	10186(11)	-712(3)	33
C(5)	12366(4)	11786(12)	-544(3)	35
C(6)	13084(3)	11594(12)	238(3)	28
C(7)	14062(4)	13263(11)	419(3)	33
C(8)	14737(4)	13001(11)	1149(3)	32
C(9)	14460(3)	11138(11)	1749(3)	27
C(10)	13510(3)	9560(10)	1616(2)	20
C(11)	13203(3)	7583(11)	2258(2)	21
C(12)	14097(3)	7159(13)	2994(3)	30
N(13)	12130(2)	8376(10)	2530(2)	18
C(14)	11390(3)	6516(11)	2737(2)	21
O(15)	11632(2)	4097(9)	2774(2)	34
C(16)	10259(3)	7575(10)	2900(2)	18
C(17)	9704(3)	9513(11)	2367(2)	22
C(18)	8643(3)	10491(11)	2469(2)	23
C(19)	8117(3)	9439(10)	3115(2)	20
O(20)	7068(2)	10502(9)	3208(2)	26
C(21)	6810(3)	10089(11)	4020(3)	25
C(22)	6986(3)	7197(11)	4283(2)	24
C(23)	6706(3)	6660(12)	5139(2)	22
O(24)	5518(2)	6959(9)	5166(2)	26
O(25)	8150(2)	6406(9)	4280(2)	20
C(26)	8647(3)	7491(10)	3634(2)	18
C(27)	9726(3)	6479(11)	3539(2)	18
C(28)	10270(4)	4441(12)	4154(3)	22
H(36)	1306(3)	557(7)	201(2)	15(7)
H(45)	641(3)	577(8)	388(2)	33(10)
H(29)	1168(3)	636(9)	96(3)	35(10)
H(30)	1045(3)	727(8)	-22(2)	34(10)
H(31)	1095(3)	1040(8)	-121(3)	27(9)
H(32)	1264(3)	1319(9)	-89(3)	39(10)
H(33)	1417(3)	1460(9)	1(3)	37(10)
H(34)	1547(3)	1404(8)	132(2)	27(9)
H(35)	1499(4)	1049(12)	220(3)	71(16)
H(37)	1488(3)	652(8)	277(3)	32(9)
H(38)	1424(3)	896(9)	326(3)	37(11)
H(39)	1388(3)	543(8)	334(2)	28(9)
H(40)	1187(3)	995(7)	242(2)	5(7)
H(41)	1010(3)	1024(9)	200(3)	34(10)
H(42)	824(3)	1205(8)	209(2)	23(8)
H(43)	738(3)	1156(9)	439(2)	33(9)
H(44)	601(2)	1061(6)	404(2)	5(6)

H(46)	699(4)	820(11)	549(3)	52(12)
H(47)	702(3)	480(8)	538(2)	20(8)
H(48)	509(4)	508(10)	512(3)	41(11)
H(49)	970(5)	414(13)	455(5)	100(21)
H(50)	1010(3)	274(10)	398(3)	32(10)
H(51)	1099(3)	470(8)	420(2)	26(9)



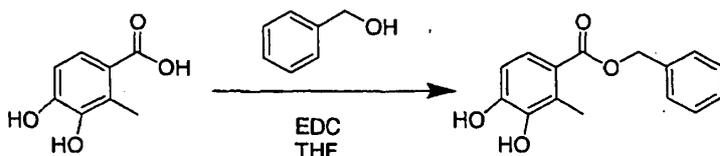
- 5 A un matraz de fondo redondo equipado con agitación magnética y una entrada de nitrógeno se le añadieron ácido 3-hidroximetil-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico (1,40 g, 6,24 mmol), THF (32 ml) y cloruro de t-butildimetilsililo (5,65 g, 18,73 mmol). Se agitó esta mezcla a temperatura ambiente durante 30 min. Se añadió imidazol (1,87 g, 27,47 mmol) y se agitó la mezcla de reacción durante otros 30 min. Se diluyó la mezcla con hexano y se lavó una vez con HCl 1 N, una vez con salmuera, se secó sobre sulfato de magnesio y se evaporó dando un aceite amarillo. Esto se purificó mediante cromatografía ultrarrápida (gel de sílice, el 5% de éter/hexano) dando un aceite amarillo, éster terc-butildimetilsilílico del ácido 3-(terc-butildimetilsililoximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico, (1,26 g, 2,78 mmol) con un rendimiento del 45%. <sup>1</sup>H-RMN (300 MHz, CD<sub>3</sub>COCD<sub>3</sub>) δ (ppm): 0,10 (d, 6H), 0,33 (s, 6H), 0,90 (s, 9H), 0,99 (s, 9H), 2,45 (s, 3H), 3,95 (m, 2H), 4,13 (m, 1H), 4,26 (m, 1H), 4,40 (dd, 1H), 6,75 (d, 1H), 7,51 (d, 1H); CCF: Rf = 0,87 (acetato de etilo/hexano 1:1).

15



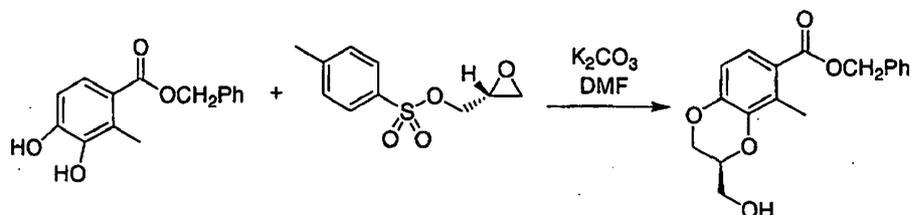
- 20 A un matraz de fondo redondo equipado con agitación magnética y una entrada de nitrógeno se le añadieron el material de partida (1,40 g, 3,09 mmol), metanol (16 ml) y carbonato de potasio al 25% (acuoso) (5 ml). Se agitó esta mezcla a temperatura ambiente durante 4 horas. Se evaporó el metanol y se añadió acetato de etilo al residuo. Se acidificó la mezcla con HCl 1 N y se separaron las fases. Se extrajo la fase acuosa una vez más con acetato de etilo. Se lavaron las fases de acetato de etilo combinadas una vez con salmuera, se secaron sobre sulfato de magnesio y se evaporaron dando un sólido amarillo, éster terc-butildimetilsilílico del ácido 3-(terc-butildimetilsililoximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico, (0,93 g, 2,75 mmol) con un rendimiento del 89%. <sup>1</sup>H-RMN (300 MHz, CD<sub>3</sub>COCD<sub>3</sub>) δ (ppm): 0,10 (s, 6H), 0,90 (s, 9H), 1,94 (s, 3H), 3,95 (m, 2H), 4,13 (m, 1H), 4,26 (m, 1H), 4,40 (dd, 1H), 6,75 (d, 1H), 7,51 (d, 1H); CCF Rf = 0,62 (acetato de etilo/hexano 1:1).

25

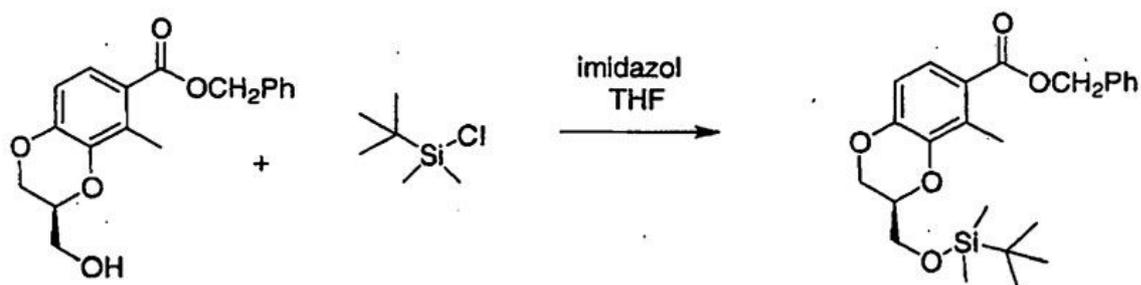


- 30 A un matraz de fondo redondo equipado con agitación magnética y una entrada de nitrógeno se le añadieron ácido 3,4-dihidroxi-2-metilbenzoico (23,0 g, 137 mmol), alcohol bencílico (21,2 ml, 205 mmol) y THF (274 ml). A esta mezcla con agitación se le añadió clorhidrato de 1-(3-dimetilaminopropil)-3-etilcarbodiimida (EDC) (30,2 g 157 mmol). Se agitó esta mezcla a temperatura ambiente durante 36 horas. Se añadieron unos pocos mililitros de agua y se agitó la mezcla durante 10 min tras lo cual se evaporó el THF. Se repartió el residuo entre acetato de etilo

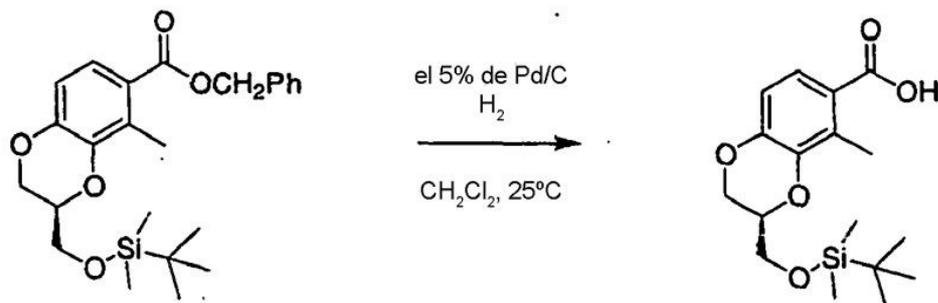
y HCl 1 N. Se separaron las fases y se extrajo la fase acuosa dos veces con acetato de etilo. Se lavaron los extractos orgánicos combinados una vez con disolución saturada de bicarbonato de sodio, una vez con HCl 1 N y una vez con salmuera. Entonces se secó la disolución orgánica sobre sulfato de sodio, se filtró y se evaporó. Se trituró el aceite resultante con el 50% de cloruro de metileno en hexanos a partir de lo cual se filtró un sólido de color tostado con un rendimiento del 43%, 3,4-dihidroxi-2-metilbenzoato de bencilo (15,2 g, 59 mmol). <sup>1</sup>H-RMN (300 MHz, CD<sub>3</sub>COCD<sub>3</sub>) δ (ppm): 2,49 (s, 3H), 5,29 (s, 2H), 6,78 (d, 1H), 7,43 (m, 6H), CCF: Rf = 0,49 (acetato de etilo/hexano 1:1).



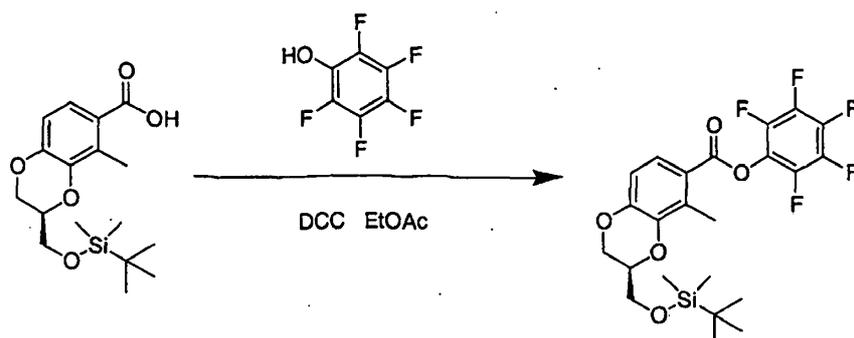
En un matraz de fondo redondo de 500 ml, se añadieron 25,21 g (97,71 mmol) de 3,4-dihidroxi-2-metilbenzoato de bencilo, 22,28 g (97,71 mmol) de tosilato de (2R)-(-)-glicidilo, 16,56 g (120 mmol) de K<sub>2</sub>CO<sub>3</sub> y 200 ml de DMF. Se agitó la mezcla de reacción a 65°C en un baño de aceite durante 5 horas. Se permitió que la reacción se enfriara, se diluyó con agua (750 ml) y se extrajo con éter (3 x 300 ml). Se sometió la fase de éter a retroextracción con 100 ml de agua y se secaron las fases orgánicas combinadas y se evaporaron proporcionando 31,1 g de aceite. La CCF indicó que se completó la reacción: Rf = 0,36 (acetato de etilo:hexano 1:1). La cromatografía en columna usando un gradiente del 10-50% de acetato de etilo en hexano proporcionó 23,0 g de éster bencílico del ácido 3-hidroximetil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carboxílico puro (rendimiento del 80%). <sup>1</sup>H-RMN (300 MHz, CD<sub>3</sub>OD) δ (ppm): 7,4 (m, 6H), 6,75 (d, 1H), 5,28 (s, 2H), 4,4 (d, 1H), 4,2 (m, 1H), 4,1 (dd, 1H), 3,8 (m, 2H), 2,46 (s, 3H).



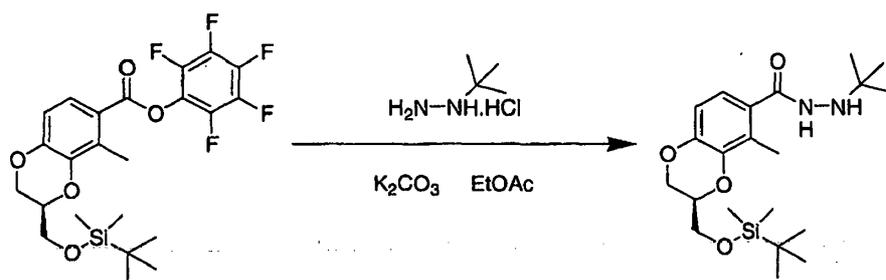
Se disolvieron 23 g (78,2 mmol) de éster bencílico del ácido 3-hidroximetil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carboxílico en 300 ml de THF seco. Mientras se agitaba, se añadieron 18,08 g de cloruro de t-butildimetilsililo y 10,2 g de imidazol. Se agitó la mezcla a temperatura ambiente durante la noche, tiempo durante el cual se desarrolló un precipitado blanco, imidazol-HCl. La monitorización mediante CCF (acetato de etilo:hexano 1:1) demostró evolución de la reacción. Se separó el precipitado blanco por filtración y se concentró la disolución en THF restante en un evaporador rotatorio. Volvió a disolverse el residuo en 300 ml de CH<sub>2</sub>Cl<sub>2</sub> y se extrajo con 300 ml de agua. Se secó el extracto de CH<sub>2</sub>Cl<sub>2</sub> y se eliminó el disolvente a vacío proporcionando 26,55 del producto de silil éter, CCF: Rf = 0,7 (acetato de etilo:hexano 1:1); rendimiento bruto = 83%. Volvió a disolverse el producto en 150 ml de hexano y se sometió a cromatografía mediante cromatografía en columna de gel de sílice, eluyendo con hexano, seguido por el 2% de acetato de etilo en hexano. Se recuperaron 21,28 g de éster bencílico del ácido 3-(terc-butil-dimetil-silaniloximetil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carboxílico puro (rendimiento del 67%). <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,45 (d, 1H), 7,3 (m, 6H), 6,65 (d, 1H), 4,27 (d, 1H), 4,12 (m, 1H), 4,05 (dd, 1H), 4,85 (dd, 1H), 4,75 (dd, 1H), 2,37 (s, 3H), 0,81 (s, 9H), 0,01 (s, 3H).



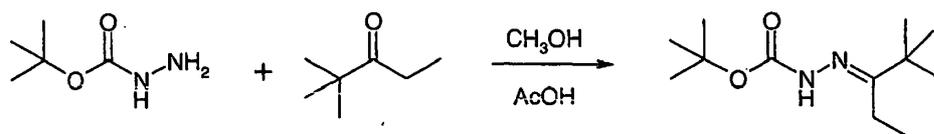
Se disolvieron 9,63 g de éster bencílico del ácido 3-(terc-butil-dimetil-silaniloximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico en aproximadamente 125 ml de  $\text{CH}_2\text{Cl}_2$  seco. Se transfirió la disolución a un frasco de hidrogenación Parr y se añadieron 5 g del 5% de Pd sobre carbono. Se cargó el frasco con hidrógeno y se agitó en un aparato de hidrogenación Parr durante 3 horas; la captación de hidrógeno cesó tras 2 horas, tal como se indicó mediante monitorización de la presión. Se añadieron 100 ml de  $\text{CHCl}_3$  y se separó el Pd/C por filtración tras añadir  $\text{MgSO}_4$  como adyuvante para aumentar el tamaño de partícula. Se lavó la fase de  $\text{CH}_2\text{Cl}_2$ - $\text{CHCl}_3$  con 200 ml de HCl 0,1 N para eliminar el subproducto, alcohol bencílico. Se secó la fase orgánica y se evaporó dando 6,47 g de ácido 3-(terc-butil-dimetil-silaniloximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico, tras la eliminación de disolvente traza en un horno de vacío. CCF Rf=0,55 (acetato de etilo:hexano 1:1). Nota: (a) la hidrogenación catalizada por Pd es un método muy superior al de trietilsilano/diacetato de paladio para esta escisión de éster bencílico, (b) lavar el producto con ácido acuoso débil es una buena manera de eliminar el alcohol bencílico.  $^1\text{H-RMN}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,55 (d, 1H), 6,7 (d, 1H), 4,3 (d, 1H), 4,17 (m, 1H), 4,02 (dd, 1H), 3,82 (dd; 1H), 3,75 (dd, 1H), (2,45 (s, 3H), 0,8 (s, 9H), 0,01 (s, 3H).



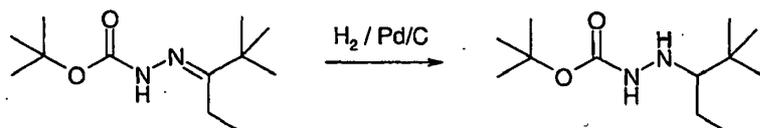
A un matraz de fondo redondo equipado con agitación magnética y una entrada de nitrógeno se le añadieron ácido 3-(terc-butildimetilsililoximetil)-5-metil-2,3-dihidrobenczo[1,4]dioxin-6-carboxílico (0,72 g, 2,13 mmol) y acetato de etilo (11 ml). A esto se le añadió pentafluorofenol (0,82 g, 2,23 mmol) seguido por diciclohexilcarbodiimida 1 M en cloruro de metileno (2,34 ml, 2,34 mmol). Se agitó esta mezcla a temperatura ambiente durante cuatro horas. Se añadieron unos pocos mililitros de agua y se continuó la agitación durante 10 min. Se filtró la reacción, se diluyó con acetato de etilo, se lavó una vez con HCl 1 N, una vez con disolución acuosa saturada de bicarbonato de sodio, una vez con cloruro de sodio saturado, se secó sobre sulfato de magnesio, se filtró y se evaporó dando un aceite. Se sometió este aceite a cromatografía ultrarrápida sobre gel de sílice eluyendo con hexanos dando un aceite incoloro transparente, éster pentafluorofenílico del ácido 3-(terc-butildimetilsililoximetil)-5-metil-2,3-dihidrobenczo[1,4]dioxin-6-carboxílico, (0,90 g, 1,78 mmol) con un rendimiento del 84%.  $^1\text{H-RMN}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 0,10 (s, 6H), 0,90 (s, 9H), 2,50 (s, 3H), 3,90 (m, 2H), 4,14 (m, 1H), 4,24 (m, 1H), 4,40 (m, 1H), 6,83, (d, 1H), 7,86, (d, 1H); CCF Rf = 0,88 (acetato de etilo/hexano 1:1).



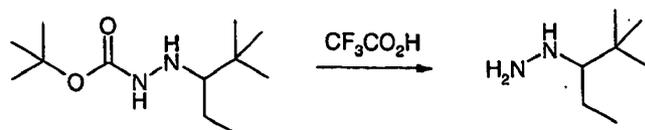
A un matraz de fondo redondo equipado con agitación magnética se le añadieron acetato de etilo (5 ml) y disolución acuosa al 25% en peso de carbonato de potasio (2,96 g, 5,35 mmol). A esto se le añadió clorhidrato de t-butilhidrazina (0,33 g, 2,68 mmol) seguido por éster pentafluorofenílico del ácido 3-(terc-butildimefilsililoximetil)-5-metil-2,3-dihidrobenczo[1,4]dioxin-6-carboxílico (0,90 g, 1,78 mmol) disuelto en acetato de etilo (4 ml). Se agitó esta mezcla a temperatura ambiente durante 18 horas. Se separaron las fases y se lavó la fase orgánica una vez con agua, una vez con NaOH al 10% (ac.), una vez con HCl 1 N y una vez con cloruro de sodio acuoso saturado. Se secó la disolución sobre sulfato de magnesio, se filtró y se evaporó dando un sólido blanco. Se trituró este material con hexano dando un sólido blanco, N-terc-butil-hidrazida del ácido 3-(terc-butil-dimetil-silaniloximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico, (0,50 g, 1,22 mmol) con un rendimiento del 69%.  $^1\text{H-RMN}$  (300 MHz,  $\text{CD}_3\text{COCD}_3$ )  $\delta$  (ppm): 0,12 (s, 6H), 0,92 (s, 9H), 1,53 (s, 9H), 2,32 (s, 3H), 3,97 (m, 2H), 4,14 (m, 1H), 4,28 (m, 1H), 4,43 (m, 1H), 6,79 (d, 1H), 7,22 (d, 1H); CCF Rf = 0,43 (acetato de etilo : hexano 1:1).



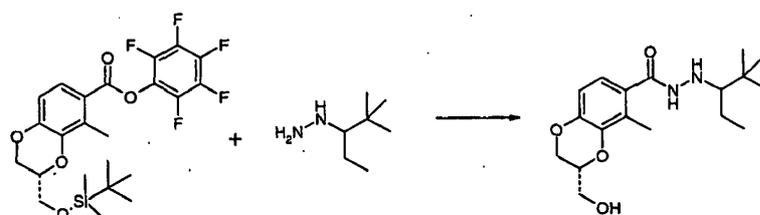
- 5 Se mezcló terc-butil etil cetona (5 g) con carbazato de terc-butilo (6,4 g, 1,1 eq.) en 30 ml de metanol con 3 gotas de ácido acético a temperatura ambiente. Se formó un sólido en el plazo de una hora; se añadió más metanol y se agitó la mezcla a temperatura ambiente durante tres días para producir éster terc-butílico del ácido N'-(1-etil-2,2-dimetilpropiliden)-hidrazinacarboxílico.  $^1\text{H-RMN}$  ( $\text{CDCl}_3$ , 500 MHz)  $\delta$  (ppm): 7,5 (a, 1H), 2,23 (q, 2H), 1,51 (s, 9H), 1,15 (s, 9H), 1,1 (t, 3H).



- 10 Se mezclaron aproximadamente 10,3 g de éster terc-butílico del ácido N'-(1-etil-2,2-dimetilpropiliden)-hidrazinacarboxílico con 4,14 g (3,9 mmol) del 10% de Pd/C en  $\text{CH}_2\text{Cl}_2$  bajo una atmósfera de hidrógeno en un agitador Parr durante cuatro horas. Se monitorizó la reacción mediante CCF (indicador de  $\text{I}_2$ ) y se sometió el producto bruto dos veces a cromatografía usando el 10% de éter en hexano proporcionando éster terc-butílico del ácido N'-(1-etil-2,2-dimetilpropil)-hidrazinacarboxílico.  $^1\text{H-RMN}$  ( $\text{CDCl}_3$ , 500 MHz)  $\delta$  (ppm): 6,05 (a, 1H), 3,9 (a, 1H), 2,45 (m, 1H), 1,6 (m, 1H), 1,5 (s, 9H), 1,25 (m, 1H), 1,1 (t, 3H), 0,92 (s, 9H).

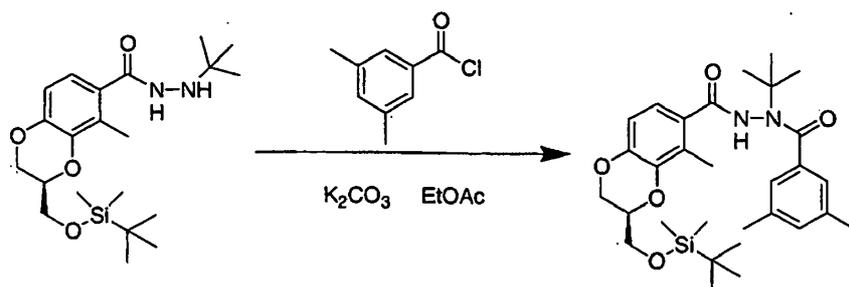


- 20 Se mezcló éster terc-butílico del ácido N'-(1-etil-2,2-dimetilpropil)-hidrazinacarboxílico (1,31 g, 5,68 mmol) disuelto en acetato de etilo con 15 ml de ácido trifluoroacético en un baño de hielo y se agitó a temperatura ambiente durante la noche. Volvió a enfriarse la mezcla de reacción y se añadieron 40 ml de agua fría, dando como resultado la formación de una nueva fase aceitosa que contenía (1-etil-2,2-dimetilpropil)-hidrazina. Entonces se añadió NaOH al 10% (aproximadamente 60 ml) hasta que la mezcla permaneció básica. Se usó esta mezcla en acetato de etilo/acuosa "tal cual" en el acoplamiento con un éster pentafluorfenílico.



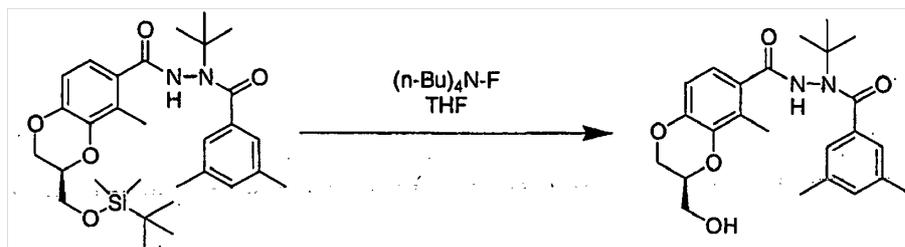
- 30 A la preparación bifásica de (1-etil-2,2-dimetilpropil)-hidrazina descrita anteriormente se le añadió una disolución de aproximadamente un gramo de éster pentafluorofenílico del ácido 3(S)-(terc-butil-dimetil-silaniloximetil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carboxílico disuelto en 10 ml de acetato de etilo. Se agitó la mezcla a temperatura ambiente durante 2 horas, tiempo tras el cual se retiró la fase acuosa y se sustituyó por  $\text{K}_2\text{CO}_3$  1 M. Se agitó la mezcla durante la noche. Se separó la fase acuosa y se extrajo la fase orgánica una vez con  $\text{K}_2\text{CO}_3$  1 M y después una vez con agua. Se secó la fase orgánica. Se sometió el producto bruto a cromatografía sobre gel de sílice usando un gradiente en etapas del 10% de éter en hexano seguido por éter puro, proporcionando N-(1-etil-2,2-dimetilpropil)-hidrazida del ácido 3(S)-hidroximetil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carboxílico.  $^1\text{H-RMN}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,15 (a, 1H), 6,85 (d, 1H), 6,73 (d, 1H), 4,85 (a, 1H), 4,3 (d, 1H), 4,25 (m, 1H), 4,1 (dd, 1H), 3,87 (m, 2H), 2,41 (m, 1H), 2,29 (s, 3H), 1,8 (a, 1H); 1,65 (m, 1H), 1,3 (m, 1H), 1,45 (t, 3H), 0,98 (s, 9H); CCF: Rf = 0,23 (acetato de etilo : hexano 1:1).

40

**RG-115790**

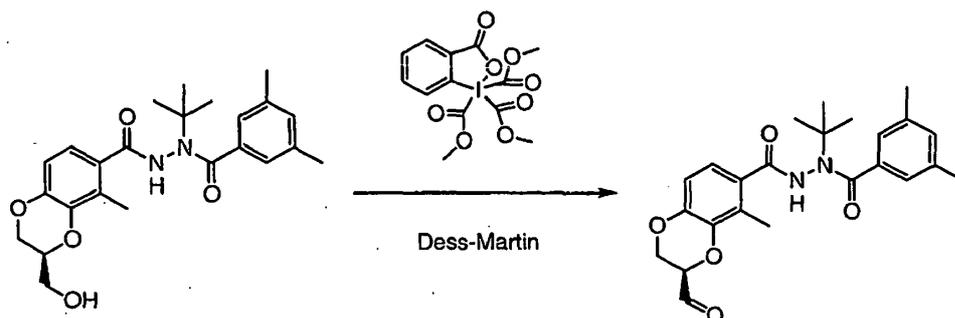
5 A un matraz de fondo redondo equipado con agitación magnética se le añadieron acetato de etilo (7 ml), disolución acuosa al 25% en peso de carbonato de potasio (2,03 g, 3,67 mmol) y N'-terc-butildimetilsililoximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carboxílico (0,50 g, 1,22 mmol). A esto se le añadió cloruro de 3,5-dimetilbenzoílo (0,31 g, 1,84 mmol). Se continuó la agitación durante 18 horas. Se diluyó la reacción con acetato de etilo y se separaron las fases. Se lavó la fase orgánica una vez con agua, una vez con cloruro de sodio acuoso saturado, se secó sobre sulfato de magnesio, se filtró y se evaporó. Se sometió el residuo a cromatografía ultrarrápida sobre gel de sílice eluyendo con el 20% de éter/hexano, después el 50% de éter/hexano dando un sólido blanco, N'-terc-butil-N'-[3-(terc-butildimetilsililoximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carbonil]-hidrazida del ácido 3,5-dimetilbenzoico, (0,50 g, 0,92 mmol) con un rendimiento del 76%. <sup>1</sup>H-RMN (300 MHz, CD<sub>3</sub>COCD<sub>3</sub>) δ (ppm): 0,10 (s, 6H), 0,90 (s, 9H), 1,59 (s, 9H), 1,89 (s, 3H), 2,25 (s, 6H), 3,92 (m, 2H), 4,05 (m, 1H), 4,22 (m, 1H), 4,34 (m, 1H), 6,33 (d, 1H), 6,56 (d, 1H), 7,02 (s, 1H), 7,13 (s, 2H), 9,49 (s, 1H); CCF: Rf = 0,30 (1:1 éter/hexano).

15

1.8 Preparación de RG-115789**RG-115789**

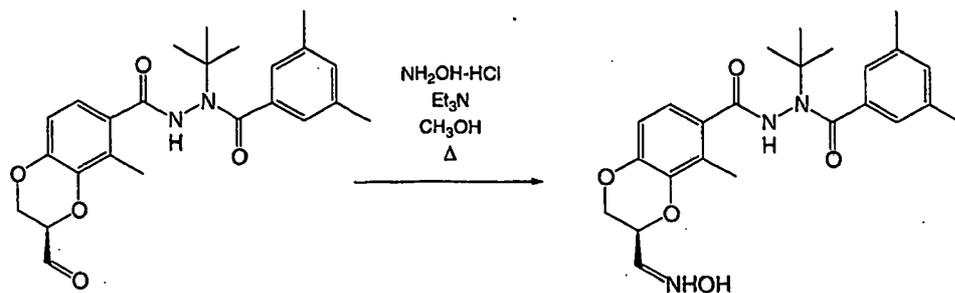
20 A un matraz de fondo redondo equipado con agitación magnética se le añadieron N'-terc-butil-N'-[3-(terc-butildimetilsililoximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carbonil]-hidrazida del ácido 3,5-dimetilbenzoico (0,26 g, 0,48 mmol), THF (5,0 ml) y disolución 1 M de fluoruro de tetrabutilamonio (TBAF) en THF (1,39 ml, 1,39 mmol). Se agitó la mezcla a temperatura ambiente durante tres horas. Se evaporó el THF y se llevó el residuo a acetato de etilo. Se lavó esta disolución una vez con agua, una vez con disolución acuosa saturada de cloruro de sodio, se secó sobre sulfato de magnesio, se filtró y se evaporó. Se trituró el residuo con hexano dando un sólido blanco, N'-terc-butil-N'-[3-(3-hidroximetil)-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carbonil]-hidrazida del ácido 3,5-dimetilbenzoico, (0,205 g, 0,48 mmol) con un rendimiento del 100%. <sup>1</sup>H-RMN (300 MHz, CD<sub>3</sub>COCD<sub>3</sub>) δ (ppm): 1,59 (s, 9H), 1,89 (d, 3H), 2,25 (s, 6H), 3,79 (m, 2H), 4,04 (m, 1H), 4,18 (m, 1H), 4,33 (m, 1H), 6,33 (dd, 1H), 6,56 (dd, 1H), 7,03 (s, 1H), 7,12 (s, 2H), 9,51 (s, 1H); CCF Rf = 0,03 (éter/hexano).

30

1.9 Preparación de RG-115812**RG-115812**

En un matraz de fondo redondo de 50 ml, se disolvieron 1,00 g (2,35 mm) de N-terc-butil-N'-(3-hidroximetil-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetilbenzoico en 10 ml de CH<sub>2</sub>Cl<sub>2</sub> y se añadieron 8,0 g de reactivo de Dess-Martin disponible comercialmente. Se agitó la reacción a temperatura ambiente durante 24 horas. Se transfirió la mezcla de reacción a un embudo de decantación con CH<sub>2</sub>Cl<sub>2</sub> y se extrajo con NaHCO<sub>3</sub> acuoso diluido, después con tiosulfato de sodio diluido (Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>) para extinguir el agente oxidante. Se secó la fase orgánica con MgSO<sub>4</sub> y se evaporó hasta sequedad, proporcionando 1,32 g de producto. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 9,70 (s, 1H), 7,50(s, 1H), 7,05 (s, 2H), 6,95 (s, 1H), 6,6 (m, 1H) 6,1-6,2 (q, 1H), 4,6 (d, 2H), 4-4,3 (m, 3H), 2,27 (s, 6H), 2,02 - 2,06 (d, 3H), 1,60 (s, 9H); CCF Rf = 0,13 (acetato de etilo:hexano 1:1).

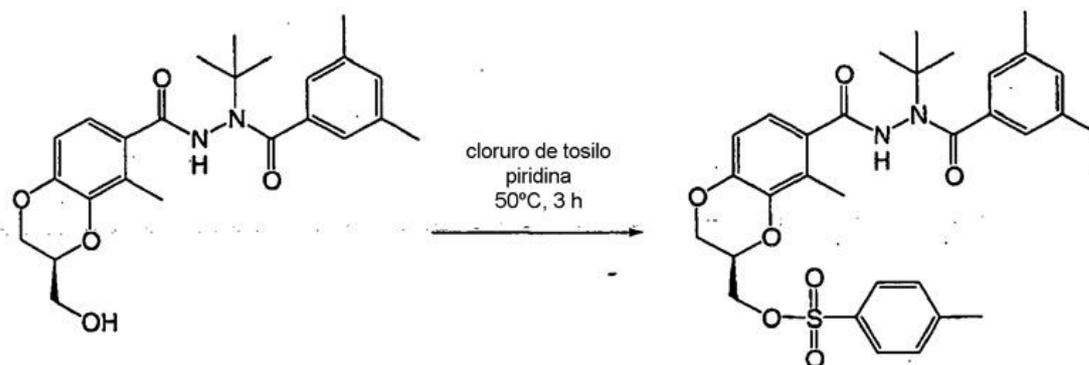
#### 1.10 Preparación de RG-115814



**RG-115814**

A un matraz de fondo redondo de 100 ml, que contenía 100 mg (0,24 mm) de N-terc-butil-N'-(3-formil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, se le añadieron 0,50 g de trietilamina, 70 mg (1 mm) de clorhidrato de hidroxilamina y 25 ml de metanol. Se sometió la mezcla de reacción a reflujo durante 1 hora. Se eliminó el metanol en un evaporador rotatorio y se disolvió el residuo con cloroformo y HCl diluido y se transfirió a un embudo de decantación. Se secó el extracto de cloroformo y se evaporó hasta sequedad. Se sometió el residuo a cromatografía sobre sílice y se eluyó el producto con el 40% de acetato de etilo en hexano. La evaporación de disolvente proporcionó 85 mg de N-terc-butil-N'-[3-(hidroxiimino-metil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil]-hidrazida del ácido 3,5-dimetil-benzoico. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,7 (s, 1H), 7,4 (d, 1H), 7,05 (s, 2H), 6,95 (s, 1H), 6,55 (d, 1H), 6,15 (m, 1H), 4,8 (m, 1H), 4,0-4,4 (m, 4H), 2,248 (s, 6H), 1,915 (s, 3H), 1,574 (s, 9H); CCF: Rf = 0,40 (acetato de etilo : hexano 1:1).

#### 1.11 Preparación de RG-115813

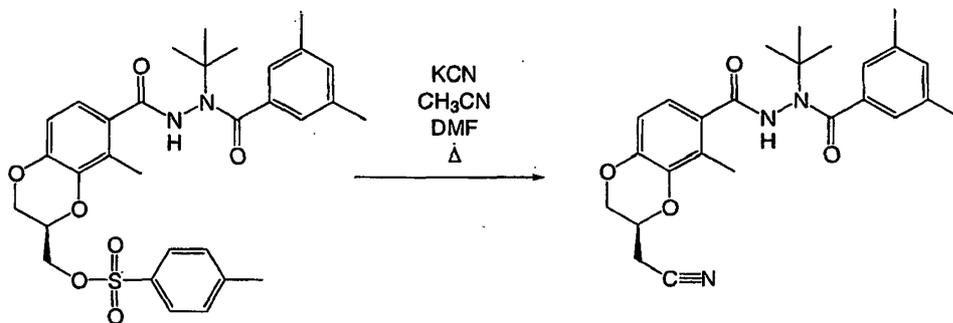


**RG-115813**

En un matraz de fondo redondo de 100 ml se añadieron 1,00 g (2,35 mm) de N-terc-butil-N'-(3-hidroximetil-5-metil-2,3-dihidrobenzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetilbenzoico, 538 mg (2,82 mm) de cloruro de tosilo y 10 ml de piridina. Se calentó la mezcla de reacción y se agitó en un baño de agua a 50-60°C durante 3 horas, después se agitó a temperatura ambiente durante 24 horas. Se disolvió la mezcla de reacción en CH<sub>2</sub>Cl<sub>2</sub> y en primer lugar se extrajo con K<sub>2</sub>CO<sub>3</sub> diluido, después con HCl diluido (para eliminar la piridina). Se secó el extracto de CH<sub>2</sub>Cl<sub>2</sub> y se concentró proporcionando aproximadamente 1,4 g de material. La CCF (acetato de etilo : hexano 1:1) dio un Rf de 0,35 para el compuesto principal. Se purificó el producto mediante cromatografía en columna sobre sílice, eluyendo con el 45% de acetato de etilo en hexano dando aproximadamente 1,1 g de éster 7-[N-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetílico del ácido tolueno-4-sulfónico puro. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,8 ppm (d, 1H), 7,55 (s. a., 1H), 7,36 (d, 1H), 7,05 (s, 2H), 6,95 (s, 1H),

6,5 (m, 1H) 6,1 (q, 1H), 4,0 - 4,4 (m, 5H) 2,45 (s, 3H), 2,25 (s, 6H), 1,85 (d, 3H), 1,58 (s, 9H).

### 1.12 Preparación de RG-115816

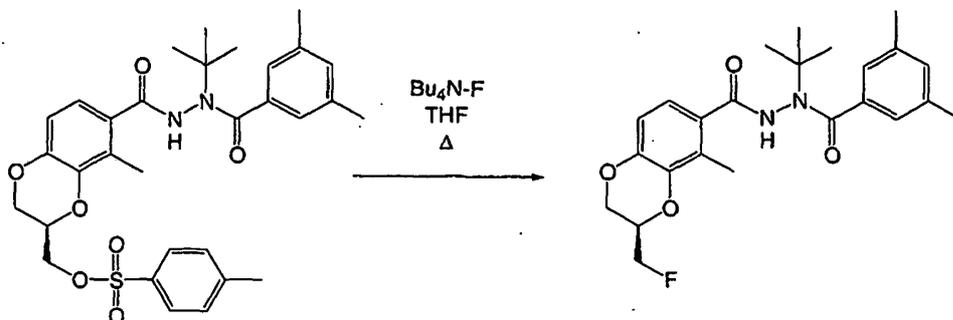


5

### RG-115816

Se sometieron 300 mg (0,52 mm) de éster 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonyl]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetílico del ácido tolueno-4-sulfónico, 100 mg de KCN, 10 mg de KI, 12 ml de CH<sub>3</sub>CN y 4 ml de DMF a reflujo durante 5 horas. Se concentró el producto de reacción en un evaporador rotatorio. Se transfirió el producto de reacción con etil éter y agua a un embudo de decantación y se extrajo dos veces con éter. Se extrajo el extracto de éter con agua, se secó y se concentró dando 0,17 g de un sólido blanco. La CCF (acetato de etilo : hexano 1:1) indicó que el producto de nitrilo tenía un R<sub>f</sub> de 0,27. Se purificó N-terc-butil-N'-(3-cianometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico mediante cromatografía en columna de gel de sílice, eluyendo con el 45% de acetato de etilo en hexano. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,70 (s, 1H), 7,05 (s, 2H), 6,95 (s, 1H) 6,60 (d, 1H), 6,1-6,2 (m, 1H), 4,0-4,4 (m, 3H), 2,78 (d, 2H), 2,25 (s, 6H), 1,93 (d, 3H), 1,57 (s, 9H).

### 1.13 Preparación de RG-115815

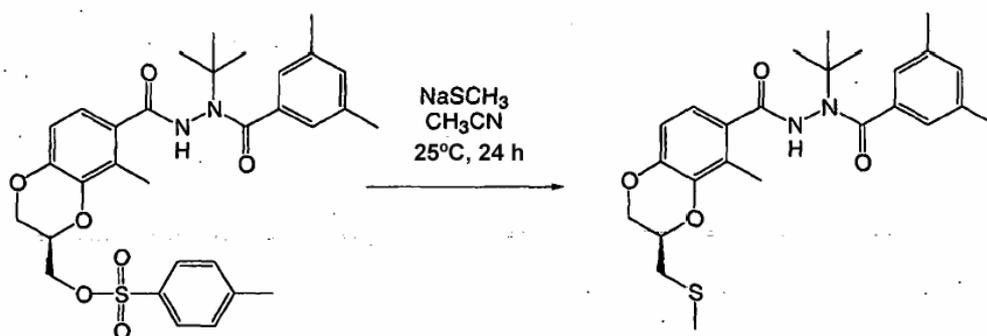


20

### RG-115815

Se sometieron 250 mg (0,43 mm) de éster 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonyl]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetílico del ácido tolueno-4-sulfónico, 2 ml de disolución 1 M de fluoruro de tetrabutilamonio (TBAF) en THF y 15 ml de THF a reflujo durante 4 horas. Se concentró la mezcla de reacción y volvió a disolverse con CH<sub>2</sub>Cl<sub>2</sub>. Se lavó el extracto de CH<sub>2</sub>Cl<sub>2</sub> con bicarbonato de sodio diluido, se secó y se concentró proporcionando 0,37 g de producto. La CCF en acetato de etilo : hexano 1:1 dio un R<sub>f</sub> de 0,48 para el componente principal. La purificación mediante cromatografía en columna sobre gel de sílice, eluyendo con el 26% de acetato de etilo en hexano, proporcionó N-terc-butil-N'-(3-fluorometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico (aproximadamente 0,25 g). <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,80 ppm (s, 1H), 7,02 (s, 2H), 6,95 (s, 1H), 6,5 (d, 1H), 6,1 (t, 1H), 4,70 (d, 1H), 4,5 (d, 1H), 4-4,3 (m, 3H), 2,24 (s, 6H), 1,93 (d, 3H), 1,56 (s, 9H).

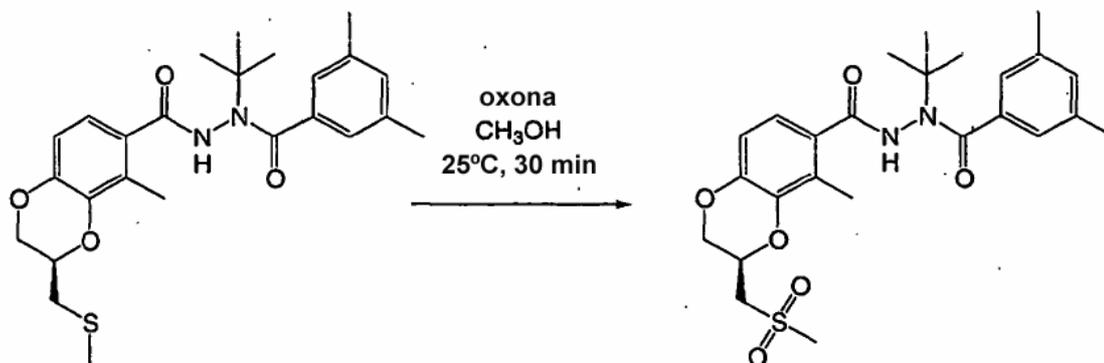
### 1.14 Preparación de RG-115817



RG-115817

Se agitaron 400 mg (0,69 mm) de éster 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetílico del ácido tolueno-4-sulfónico, 147 mg (2,1 mm) de  $\text{CH}_3\text{SNa}$  y 15 ml de  $\text{CH}_3\text{CN}$  a temperatura ambiente durante 24 horas. Se concentró la mezcla de reacción hasta sequedad y volvió a disolverse con  $\text{CH}_2\text{Cl}_2$ . Se lavó el extracto de  $\text{CH}_2\text{Cl}_2$  con  $\text{NaHCO}_3$  acuoso diluido, se secó y se concentró dando 0,26 g de producto. La CCF indicó una mezcla, con un compuesto principal a  $R_f = 0,52$  (acetato de etilo : hexano 1:1). La purificación mediante cromatografía en columna sobre sílice dio el producto deseado, N-terc-butil-N'-(5-metil-3-metilsulfanilmetil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, eluido con el 24% de acetato de etilo en hexano (0,18 g).  $^1\text{H-RMN}$ : (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,40 (s, 1H), 7,05 (s, 2H), 6,96 (s, 1H), 6,6(d, 1H), 6,1(q, 1H), 4,0-4,3(m, 3H), 2,7-2,9 (m, 2H), 2,27 (s, 6H), 2,21(s, 3H), 1,97 (d, 3H) , 1,59 (s, 9H).

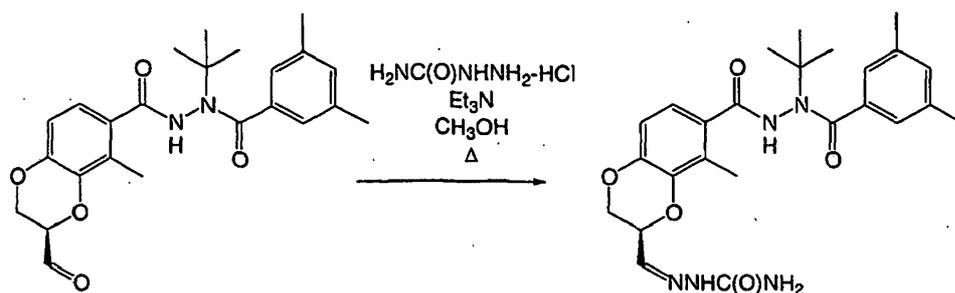
#### 1.15 Preparación de RG-115818



RG-115818

Se disolvieron 125 mg (0,27 mm) de N-terc-butil-N'-(5-metil-3-metilsulfanilmetil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico en 6 ml de  $\text{CH}_3\text{OH}$ . Con agitación, se añadieron 250 mg (0,4 mm) de oxona en 6 ml de agua, dejando agitar posteriormente durante 30 minutos. Se eliminó el metanol en un evaporador rotatorio, volvió a disolverse el residuo en  $\text{CHCl}_3$  y agua. Se extrajo la fase acuosa con cloroformo, que entonces se secó y se concentró proporcionando 130 mg de un sólido blanco. La CCF (acetato de etilo : hexano 1:1) mostró un  $R_f$  de 0,12 para el producto principal, N-terc-butil-N'-(3-metanosulfonil-metil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, que se purificó mediante cromatografía en columna sobre gel de sílice, eluyendo con el 80% de acetato de etilo en hexano.  $^1\text{H-RMN}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,55 (d, 1H), 7,02 (s, 2H), 6,98 (s, 1H), 6,6 (d, 1H), 6,0-6,1 (m, 1H), 4,0-4,5 (m, 3H), 2,8 (d, 2H), 2,25 (s, 6H), 1,92-1,94 (d, 3H), 1,57 (s, 9H).

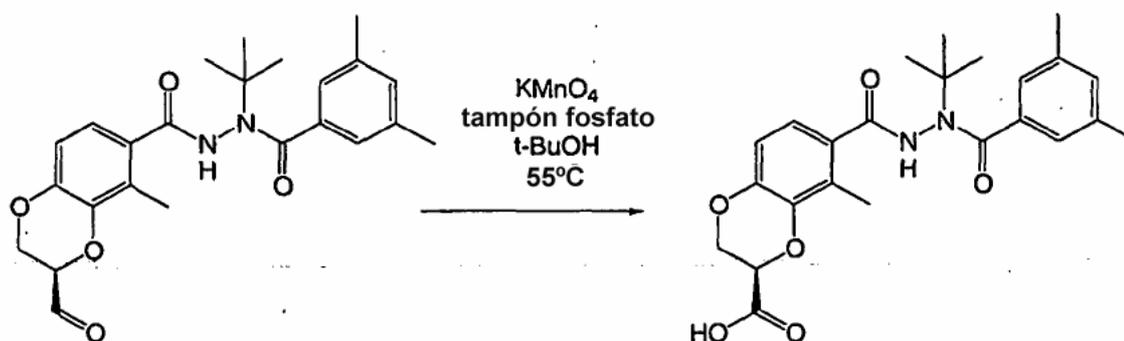
#### 1.16 Preparación de RG-115807



RG-115807

Se añadieron 200 mg (0,47 mm) de N-terc-butil-N'-(3-formil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, 0,5 g de trietilamina, 210 mg (1,9 mm) de clorhidrato de semicarbazida, 10 ml de metanol y 2 gotas de ácido acético glacial a un matraz de fondo redondo de 100 ml y se sometieron a reflujo durante 4 horas. Se concentró la mezcla de reacción en el evaporador rotatorio y volvió a disolverse en  $\text{CHCl}_3$ . Se extrajo la disolución en  $\text{CHCl}_3$  resultante dos veces con  $\text{NaHCO}_3$  diluido, se secó y se evaporó proporcionando 0,16 g de producto bruto. La CCF (acetato de etilo:hexano 1:1) mostró que el componente principal estaba en el origen. Se purificó el producto, N-terc-butil-N'-(3-fomil-semicarbazida-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, mediante cromatografía en columna sobre gel de sílice, eluyendo con el 10% de  $\text{CH}_3\text{OH}$  en acetato de etilo.  $^1\text{H-RMN}$ : (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm):  $\text{CD}_3\text{OD} - \text{CDCl}_3$ : 7,2 (d, 1H), 7,07 (s, 2H), 6,97 (s, 1H), 6,5-6,6 (2d, 1H), 6,2-6,3 (q, 1H), 4,8 (m, 1H) 4,1-4,3 (m, 2H), 2,28 (s, 6H), 1,88 (m, 3H), 1,586 (s, 9H).

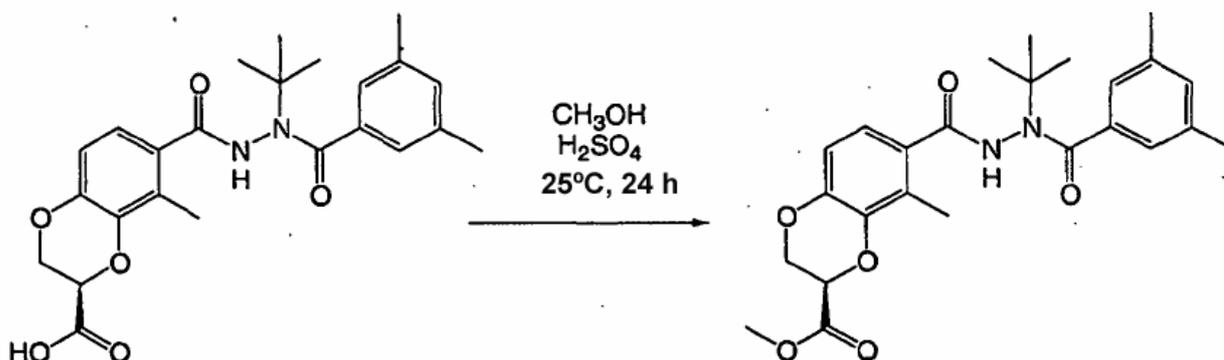
#### 1.17 Preparación de RG-115805



RG-115805

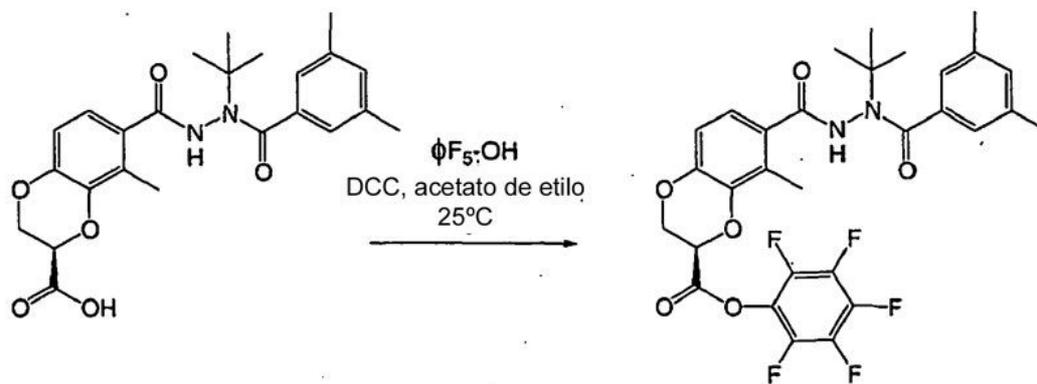
Se añadieron 0,85 g (2 mm) de N-terc-butil-N'-(3-formil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, 10 ml de alcohol t-butílico caliente y 20 ml de tampón fosfato de Aldrich, pH 7,2, n.º 31925-2, a un matraz de fondo redondo de 100 ml, que se colocó en un baño de agua a 55°C. Mientras se agitaba la mezcla de reacción, se añadieron lentamente 380 mg (2,4 mm) de permanganato de potasio y se agitó la reacción a 55°C durante 7 horas y a temperatura ambiente durante 24 horas. Se añadió  $\text{NaOH}$  acuoso al 10% hasta que el pH fue de 12 y se separó por filtración el precipitado negro ( $\text{MnO}_2$ ). Se extrajo la disolución acuosa con acetato de etilo, eliminando así el color marrón, se transfirió a un embudo de decantación, se acidificó con  $\text{HCl}$  1 N (tras lo cual precipitó producto) y después se extrajo dos veces con  $\text{CHCl}_3$ . Se secó el extracto de  $\text{CHCl}_3$ , se evaporó hasta sequedad y se secó en un horno de vacío proporcionando 0,76 g de producto, ácido 7-[N-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico.  $^1\text{H-RMN}$  (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (ppm): 7,06 (s, 2H), 6,97 (s, 1H), 6,6 (q, 1H), 6,3 y 6,0 (d+d, 1H), 4,91 (m, 1H), 4,5-4-3 (m, 2H) 2,265 (s, 6H), 2,13 y 1,90 (s+s, 3H) 1,588 (s, 9H); CCF: Rf = 0-0,17, franja (acetato de etilo : hexano 1:1).

#### 1.18 Preparación de RG-115806

**RG-115806**

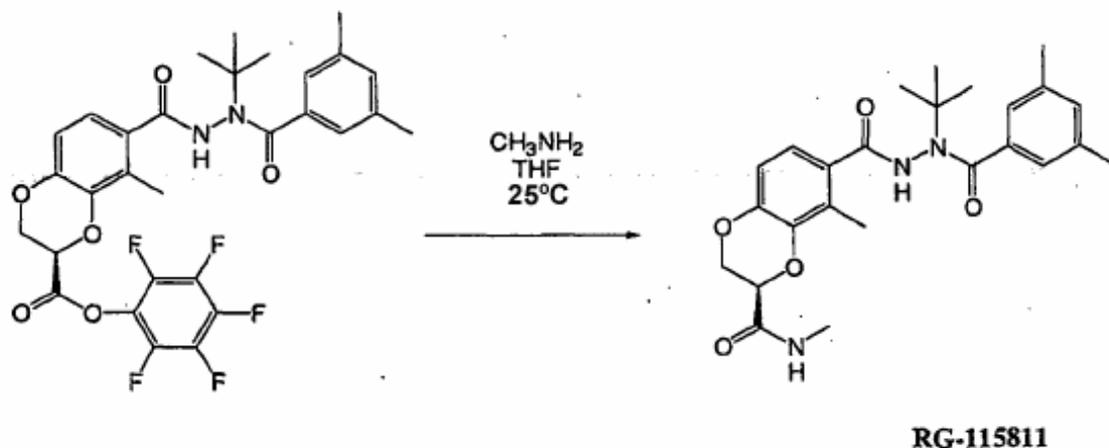
Se agitaron 190 mg (4,3 mm) de ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonyl]-8-metil-2,3-dihidrobenzo[1,4]dioxin-2-carboxílico, 15 ml de CH<sub>3</sub>OH y 1 gota de ácido sulfúrico concentrado a temperatura ambiente durante 24 horas. Se eliminó el CH<sub>3</sub>OH en un evaporador rotatorio y se disolvió el residuo en CHCl<sub>3</sub> y se extrajo con disolución diluida de NaHCO<sub>3</sub>. Se secó el extracto de CHCl<sub>3</sub> y se concentró dando 0,10 g de sólido blanco. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,52 (s, 1H), 7,07 (s, 2H), 6,97 (s, 1H), 6,5-6,6 (q, 1H), 6,3 y 6,0 (d+d, 1H), 4,8 (m, 1H), 4,4-4,2 (m, 2H), 3,772 (3H), 2,26 (s, 6H), 2,035 y 1,919 (s+s, 3H), 1,584 (s, 9H). CCF: R<sub>f</sub> = 0,35 (acetato de etilo:hexano 1:1).

#### 1.19 Preparación de RG-115810

**RG-115810**

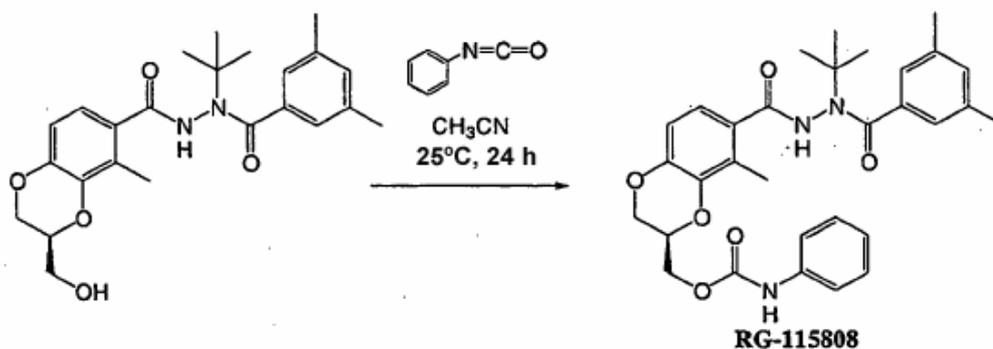
En un vial, se agitaron 97 mg (0,32 mm) de ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonyl]-8-metil-2,3-dihidrobenzo[1,4]dioxin-2-carboxílico, 60 mg (0,35 mm) de pentafluorofenol, 60 mg (0,32 mm) de DCC y 3 ml de acetato de etilo a temperatura ambiente durante 24 horas. Se transfirió la mezcla de reacción a un matraz de fondo redondo y se evaporó hasta sequedad. Volvió a disolverse el residuo en CH<sub>2</sub>Cl<sub>2</sub> y se sometió a cromatografía sobre gel de sílice. La elución con el 40% de acetato de etilo en hexano proporcionó el éster de pentafluorofenol, éster pentafluorofenílico del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonyl]-8-metil-2,3-dihidrobenzo[1,4]dioxin-2-carboxílico, en una cantidad de 80 mg. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,60 (d, 1H), 7,03 (s, 2H), 6,95 (d, 1H), 6,6 (q, 1H), 6,3 y 6,0 (d+d, 1H), 4,6 (m, 1H), 4,4 (m, 1H), 2,251 y 2,230 (s+s, 6H), 2,052 y 1,903 (s+s, 3H), 1,58 y 1,574 (s+s, 9H) (muchas señales divididas); CCF: R<sub>f</sub> = 0,52 (1:1 hexano : acetato de etilo).

#### 1.20 Preparación de RG-115811



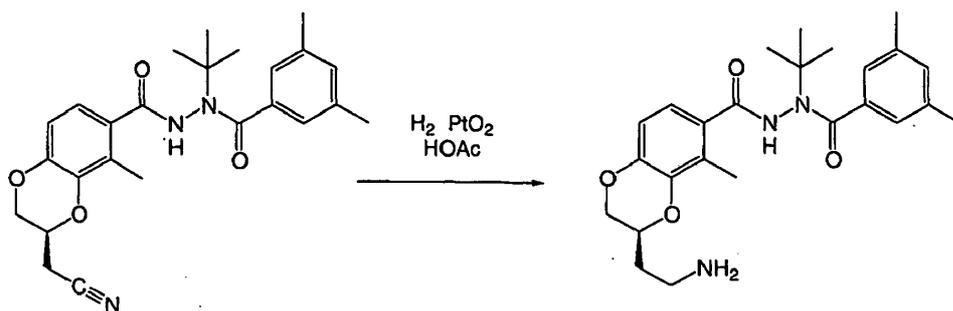
En un vial, se agitaron 80 mg (0,2 mm) de éster pentafluorofenílico del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico y 2 ml de una disolución 2 M de CH<sub>3</sub>NH<sub>2</sub> en THF a temperatura ambiente durante 24 horas. Se transfirió la mezcla de reacción a una columna de cromatografía en sílice. Se eluyeron impurezas con el 45% de acetato de etilo en hexano y se eluyó el producto, metilamida del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico, con el 90% de acetato de etilo en hexano. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 8,1 (d, 1H), 7,05 (s, 2H), 6,97 (d, 1H), 6,40 (d, 1H), 6,00 (d, 1H), 4,45-4,25 (m, 2H), 3,9-4,0 (m, 1H), 2,88 y 2,86 (s+s, 3H), 2,24 y 2,22 (s+s, 6H), 2,07 y 1,87 (s+s, 3H), 1,59 y 1,57 (s+s, 9H) (varias señales divididas); CCF: R<sub>f</sub> = 0,06 (acetato de etilo : hexano 1:1).

#### Preparación de RG-115808



Se agitaron 120 mg (0,28 mm) de N-terc-butil-N'-(3-hidroxi-metil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, 70 mg (0,55 mm) de isocianato de fenilo y 2 ml de CH<sub>3</sub>CN en un vial a temperatura ambiente durante 24 horas. Se eliminó el CH<sub>3</sub>CN mediante purga con una corriente de N<sub>2</sub> y se trituró el residuo con pentano. Se retiró el sobrenadante y se eliminó el pentano residual mediante purga con N<sub>2</sub> dando 0,15 g de producto éster 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetilico del ácido fenil-carbámico. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,65 (d, 1H), 7,4-7,2 (m, 5H), 7,1 (d, 1H), 7,06 (s, 2H), 6,98 (s, 1H), 6,86 (s. a., 1H), 6,53 (d, 1H), 6,12 (d, 1H), 4,5-4,2 (m, 3H), 4,0 (m, 2H), 2,23 (s, 6H), 1,94 (s, 3H), 1,58 (s, 9H); CCF: R<sub>f</sub> 0,4-0,5, franja (acetato de etilo : hexano 1:1).

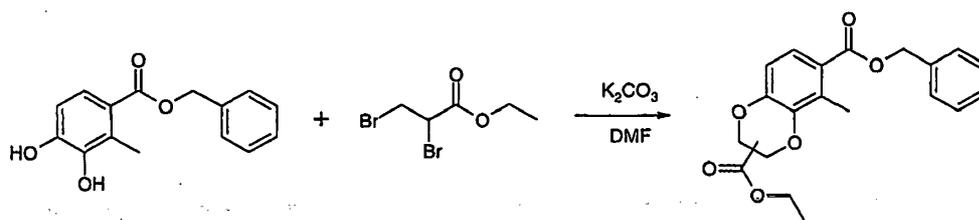
#### 1.22 Preparación de RG-115809



RG-115809

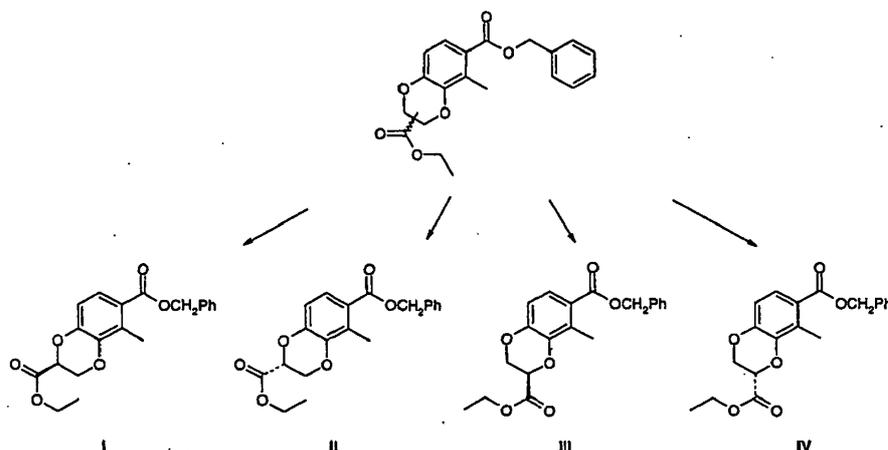
Se disolvieron 150 mg (0,34 mm) de N-terc-butil-N'-(3-cianometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico en 8 ml de ácido acético glacial y se añadieron a un frasco de hidrogenación Parr, junto con aproximadamente 16 mg de PtO<sub>2</sub>. Se realizó la hidrogenación Parr durante 4 horas y se filtró la mezcla de reacción. Se transfirió el contenido a un matraz de fondo redondo con CHCl<sub>3</sub> y acetato de etilo y se concentró la mezcla de reacción hasta sequedad. Volvió a disolverse el residuo con CHCl<sub>3</sub> y KOH 0,1 N y se transfirió a un embudo de decantación. Se extrajo de nuevo la fase acuosa con CHCl<sub>3</sub> y se secó el extracto de CHCl<sub>3</sub> y se concentró proporcionando 0,14 g de producto, N'-[3-(2-amino-etil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil]-N-terc-butil-hidrazida del ácido 3,5-dimetil-benzoico. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,8 (d, 1H), 7,00 (s, 2H), 6,95 (s, 1H), 6,6-6,5 (m, 1H), 6,2-6,1 (m, 1H), 4,4-3,8 (m, 3H), 2,9-2,75 (m, 2H), 2,244 (s, 6H), 1,92 (t, 3H), 1,7 (m, 2H), 1,565 (s, 9H); CCF: Rf = 0,24 (acetato de etilo:hexano 1:1).

### 1.23 Preparación de RG-119098 (referencia)



RG-119098

Se mezclaron 16,7 g (64,7 mmol) de 2-metil-3,4-dihidroxi-benzoato de bencilo con 2,3-dibromopropionato de etilo (20,17 g, 77,6 mmol), carbonato de potasio (10,72 g, 77,6 mmol) y DMF (216 ml) y se calentaron hasta 40-45°C durante 4 horas. Se diluyó la mezcla de reacción con éter, se lavó una vez con agua y tres veces con salmuera. Se secó la fase orgánica sobre sulfato de magnesio y se evaporó. Se purificó una pequeña muestra mediante cromatografía ultrarrápida, eluyendo con el 10% de éter/cloruro de metileno. La destilación de tipo Kugelrohr a alto vacío y el calentamiento hasta 200°C no volatilizan el producto deseado, que permanece en el matraz de destilación como un aceite naranja. No obstante, tal tratamiento proporcionó algo de purificación de éster 2 (o 3)-etílico del éster 7-bencilico del ácido 8-metil-2,3-dihidro-benzo[1,4]dioxin-2 (o 3),7-dicarboxílico. <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) δ (ppm): 7,565 y 7,525 (d+d, 1H), 7,4 (m, 5H), 6,865 y 6,725 (d+d, 1H), 5,3 (s, 2H), 4,9 (m, 1H), 4,4 (m, 2H), 4,3 (m, 2H), 2,54 y 2,46 (s+s, 3H), 1,3 (m, 3H).



30

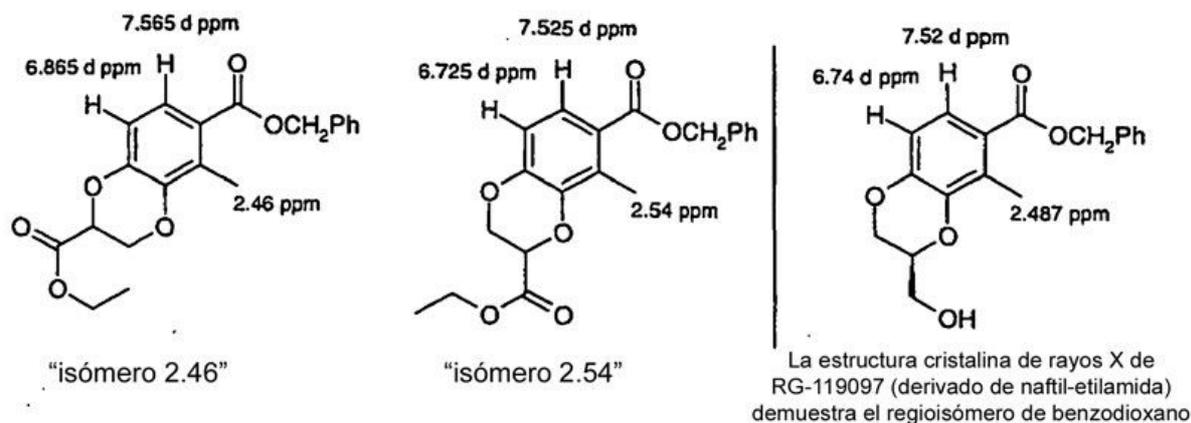
Se resolvió la mezcla aproximadamente 1:1:1 de regio y estereoisómeros de benzodioxano que comprende RG-119098, éster 2 (o 3)-etílico del éster 7-bencílico del ácido 8-metil-2,3-dihidro-benzo[1,4]dioxin-2 (o 3),7-dicarboxílico, unos de otros con una resolución de línea base y se purificó en una escala de múltiples gramos usando una columna de cromatografía quiral Chiralcel® OD-H® HPLC, n.º de serie ODH0CE-CB035. La fase móvil fue hexano:etanol

5 97,5:2,5 a una velocidad de flujo de 1 ml/min a 25°C. Se aisló cada uno de los cuatro isómeros de los tres restantes. El trabajo lo realizó Chiral Technologies, 730 Spring Drive, Exton PA.

En otro experimento, se purificaron aproximadamente 12 g de éster bencílico RG-119098 sobre gel de sílice. La elución con hexano proporcionó aproximadamente 10 g de una mezcla aproximadamente 1:1 de 2 regioisómeros.

10 Se denominaron "isómero 2,46" e "isómero 2,54", basándose en la absorbancia del grupo CH<sub>3</sub> bencílico. La elución continuada con el 10% de acetato de etilo en hexano produjo una muestra que estaba enriquecida en el isómero 2,46 en una razón de aproximadamente 2:1 con respecto al isómero 2,54. La nueva cromatografía de este material en un gradiente del 6-9% de acetato de etilo en hexano demostró un enriquecimiento adicional del isómero 2,46. Por tanto, la fracción eluida con el 8% de acetato de etilo en hexano estaba compuesta por una razón 6:1 y la fracción

15 del 9% estaba compuesta por una razón 8:1 de los isómeros 2,46:2,54.



Los isómeros se asignan provisionalmente como en el diagrama basándose en señales de <sup>1</sup>H-RMN (300 MHz, CDCl<sub>3</sub>) análogas de un 2-hidroximetilbenzodioxano relacionado de regioquímica conocida, basándose en la correlación con la estructura cristalina de RG-119097. La comparación de la <sup>1</sup>H-RMN de isómeros aislados que comprenden RG-119098 con el éster etílico correspondiente correlacionado con RG-119097 (el isómero III indicado en el diagrama) proporcionará una asignación regioquímica inequívoca de cada isómero.

## 25 EJEMPLO 2: ENSAYO BIOLÓGICO DE COMPUESTOS

Los ligandos de la presente invención son útiles en diversas aplicaciones incluyendo terapia génica, expresión de proteínas de interés en células huésped, producción de organismos transgénicos y ensayos basados en células.

### 30 Ensayo 27-63

Casete de expresión génica

GAL4 DBD (1-147)-CfEcR(DEF)/VP16AD-βRXREF-LmUSPEF: Se fusionaron los dominios D, E y F de tipo natural de EcR de gusano de la yema de las píceas, *Choristoneura fumiferana* ("CfEcR-DEF"; SEQ ID NO: 1) a un dominio de unión a ADN de GAL4 ("Gal4DBD1-147"; SEQ ID NO: 2) y se colocaron bajo el control de un promotor de fosfoglicerato cinasa ("PGK"; SEQ ID NO: 3). Se fusionaron las hélices 1 a 8 de los dominios EF de RXRβ de *Homo sapiens* ("HsRXRβ-EF"; nucleótidos 1-465 de SEQ ID NO: 4) y las hélices 9 a 12 de los dominios EF de la proteína ultraespiráculo de *Locusta migratoria* ("LmUSP-EF"; nucleótidos 403-630 de SEQ ID NO: 5) al dominio de transactivación de VP16 ("VP16AD"; SEQ ID NO: 6) y se colocaron bajo el control de un promotor de factor-α de elongación ("EF-1α"; SEQ ID NO: 7). Se fusionaron cinco sitios de unión al elemento de respuesta a GAL4 consenso ("5XGAL4RE"; que comprende 5 copias de un GAL4RE que comprende SEQ ID NO: 8) a un promotor mínimo TATA sintético (SEQ ID NO: 9) y se colocaron aguas arriba del gen indicador de luciferasa (SEQ ID NO: 10).

Se transfectaron transitoriamente células CHO con casetes de transcripción para GAL4 DBD (1-147) CfEcR(DEF) y para VP16AD βRXREF-LmUSPEF controlados por promotores celular activos de manera ubicua (PGK y EF-1α, respectivamente) en un único plásmido. Se seleccionaron células transfectadas de forma estable mediante resistencia a zeocina. Se transfectaron transitoriamente clones de células CHO aislados individualmente con un indicador de GAL4 RE-luciferasa (pFR Luc). Se seleccionó el clon 27-63 usando higromicina.

50 Tratamiento con ligando

## ES 2 478 622 T3

Se tripsinizaron células y se diluyeron hasta una concentración de  $2,5 \times 10^4$  células/ml. Se colocaron 100  $\mu$ l de suspensión celular en cada pocillo de una placa de 96 pocillos y se incubaron a 37°C bajo un 5% de CO<sub>2</sub> durante 24 h. Se prepararon disoluciones madre de ligando en DMSO y se diluyeron 300 veces para todos los tratamientos.

5 Las pruebas de respuesta a la dosis consistieron en 8 concentraciones que oscilaban entre 33  $\mu$ M y 0,01  $\mu$ M.

Ensayo de gen indicador Se midió la expresión de gen indicador de luciferasa 48 h después del tratamiento celular usando el sistema de ensayo de luciferasa Bright-Glo™ de Promega (E2650). Se detectó la luminiscencia a temperatura ambiente usando un luminómetro de placa de microtitulación Dynex MLX. Se calcularon las CE<sub>50</sub> a partir de datos de respuesta a la dosis usando un modelo logístico de tres parámetros.

10

Se muestran los resultados de los ensayos en la tabla 1. Se llevó a cabo cada ensayo en dos pocillos separados y se calculó el promedio de los dos valores. Se determinó FI máxima relativa como la inducción en veces máxima del ligando sometido a prueba (una realización de la descripción) observada a cualquier concentración en relación con la inducción en veces máxima de GS™-E (N-terc-butil-N'-(2-etil-3-metoxi-benzoil)-hidrazida del ácido 3,5-dimetil-benzoico) observada a cualquier concentración. Tabla 1: Resultados del ensayo biológico para compuestos

15

Compuesto	CE <sub>50</sub> ( $\mu$ M)/F1 máxima relativa; ensayo 27-63
RG-115789	1,85/0,81
RG-115790	> 33/0,0
RG-115805	> 33/0,01
RG-115806	~33/0,25
RG-115807	3,34/0,80
RG-115808	1,12/2,37
RG-115809	3,25/1,09
RG-115810	~20/1,23
RG-115811	> 33/0,24
RG-115812	2,68/0,86
RG-115813	> 33/0,01
RG-115814	0,32/1,0
RG-115815	0,3/0,738
RG-115816	4,22/0,85
RG-115817	2,90/0,95
RG-115818	5,26/0,90
RG-115843*	0,85/0,73
RG-115844*	1,23/0,352
RG-115853*	0,10/0,898
RG-115854*	0,20/0,917
RG-115845*	4,47/0,52
RG-115855*	0,0267/1,16
RG-115860*	F1 promedio = 2307 a 33 $\mu$ M
RG-115877*	2,25/1,48
RG-115878*	0,10/1,05
Referencia:	
RG-102317	0,102
Ligando GS™-E	0,288

Ligando GS™-E = N-terc-butil-N'-(2-etil-3-metoxi-benzoil)-hidrazida del ácido 3,5-dimetil-benzoico

\* Compuestos de referencia.

5 Además, un experto habitual en la técnica puede predecir también que los ligandos dados a conocer en el presente documento funcionarán también para modular la expresión génica en diversos tipos de células descritos anteriormente usando sistemas de expresión génica basados en receptores nucleares de grupo H y grupo B.

**Lista de secuencias**

- 10 <110> RheoGene Holdings, Inc.  
 Hormann, Robert E  
 Chortyk, Orestes  
 Smith, Howard  
 Meteyer, Thomas
- 15 Tice, Colin M  
 <120> LIGANDOS DE DIACILHIDRAZINA PARA MODULAR LA EXPRESIÓN DE GENES EXÓGENOS EN SISTEMAS DE MAMÍFERO A TRAVÉS DE UN COMPLEJO DE RECEPTOR DE ECDISONA  
 <130> A01381-PCT  
 <140> No asignado aún
- 20 <141> 10-02-2004  
 <150> Documento US 60/446.233  
 <151> 10-02-2003  
 <150> No asignado aún  
 <151> 09-02-2004
- 25 <160> 10  
 <170> PatentIn versión 3.2  
 <210> 1  
 <211> 1054  
 <212> ADN
- 30 <213> *Choristoneura fumiferana*  
 <400> 1  
 cctgagtgcg tagtaccgca gactcagtgc gccatgaagc ggaaagagaa gaaagcacag 60  
 aaggagaagg acaaactgcc tgtcagcacg acgacggtgg acgaccacat gccgcccatt 120  
 atgcagtgtg aacctccacc tcctgaagca gcaaggattc acgaagtggc cccaaggttt 180  
 ctctccgaca agctgttgga gacaaaccgg cagaaaaaca tccccagtt gacagccaac 240  
 cagcagttcc ttatcgccag gctcatctgg taccaggacg ggtacgagca gccttctgat 300  
 gaagatttga agaggattac gcagacgtgg cagcaagcgg acgatgaaaa cgaagagtct 360  
 gacactccct tccgccagat cacagagatg actatcctca cggccaact tatcgtggag 420  
 ttcgcaaggg gattgccagg gtccgccaag atctcgcagc ctgatcaaat tacgctgctt 480  
 aaggcttgct caagtggagt aatgatgctc cgagtcgctc gacgatacga tgcggcctca 540  
 gacagtgttc tgttcgcgaa caaccaagcg taactcgcg acaactaccg caaggctggc 600  
 atggcctacg tcatcgagga tctactgcac ttctgccggt gcatgtactc tatggcgttg 660  
 gacaacatcc attacgcgct gctcacggct gtcgtcatct tttctgaccg gccagggttg 720  
 gagcagccgc aactggtgga agaaatccag cggactacc tgaatacgct ccgcatctat 780  
 atcctgaacc agctgagcgg gtcggcgcgt tcgtccgtca tatacggcaa gatcctctca 840  
 atcctctctg agctacgcac gctcggcatg caaaactcca acatgtgcat ctccctcaag 900  
 ctcaagaaca gaaagctgcc gcctttctc gaggagatct gggatgtggc ggacatgtcg 960  
 cacaccaac cgccgcctat cctcaggtcc cccacgaatc tctagcccct gcgcgcacgc 1020  
 atcgcctgat ccgctgcccg ccgctgctgct ctga 1054
- 35 <210> 2  
 <211> 441  
 <212> ADN  
 <213> *Saccharomyces cerevisiae*  
 <400> 2

ES 2 478 622 T3

	atgaagctac tgtcttctat cgaacaagca tgcgatattt gccgacttaa aaagctcaag	60
	tgctccaaag aaaaaccgaa gtgCGccaag tgtctgaaga acaactggga gtgtcgctac	120
	tctcccaaaa ccaaaaggtc tccgctgact agggcacatc tgacagaagt ggaatcaagg	180
	ctagaaagac tggAACagct atttctactg atttttctc gagaagacct tgacatgatt	240
	ttgaaaatgg attctttaca ggatataaaa gcattgttaa caggattatt tgtacaagat	300
	aatgtgaata aagatgccgt cacagataga ttggcttcag tggagactga tatgcctcta	360
	acattgagac agcatagaat aagtgcgaca tcatcatcgg aagagagtag taacaaaggT	420
	caaagacagt tgactgtatc g	441
	<210> 3	
	<211> 538	
	<212> ADN	
5	<213> <i>Homo sapiens</i>	
	<400> 3	
	tcgagggccc ctgcaggTca attctaccgg gtaggggagg cgcttttccc aaggcagtct	60
	ggagcatgCg cttagcagc cccgctggca cttggcgcta cacaagtggc ctctggcctc	120
	gcacacattc cacatccacc ggtagcGcca accggctccg ttctttggTg gccccttcgc	180
	gccaccttct actcctcccc tagtcaggaa gttccccccc gcccgcgagc tcgCgtcgtg	240
	caggacgtga caaatggaag tagcacgtct cactagtctc gtgcagatgg acagaccgc	300
	tgagcaatgg aagcgggtag gcctttgggg cagcggccaa tagcagcttt gctccttcgc	360
	tttctgggct cagaggctgg gaaggggtgg gtccgggggc gggctcaggg gcgggctcag	420
	gggcggggcg ggcggaagg tcctcccgag gcccggcatt ctgcacgct tcaaaagcgc	480
	acgtctgccg cgctgttctc ctcttctca tctccgggcc tttcgacctg cagccaat	538
	<210> 4	
	<211> 720	
	<212> ADN	
10	<213> <i>Homo sapiens</i>	
	<400> 4	
	gccccgagg agatgcctgt ggacaggatc ctggaggcag agcttgctgt ggaacagaag	60
	agtgaccagg gcgttgaggg tcctggggga accgggggta gcggcagcag cccaaatgac	120
	cctgtgacta acatctgtca ggCagctgac aaacagctat tcacgcttgt tgagtgggCg	180
	aagaggatcc cacacttttc ctcttgcct ctggatgatc aggtcatatt gctgcgggca	240
	ggctggaatg aactcctcat tgctccttt tcacaccgat ccattgatgt tcgagatggc	300
	atcctccttg ccacaggTct tcacgtgcac cgcaactcag cccattcagc aggagtagga	360
	gccatctttg atcgggtgct gacagagcta gtgtccaaaa tgctgacat gaggatggac	420
	aagacagagc ttggctgcct gagggcaatc attctgttta atccagatgc caagggcctc	480
	tccaacccta gtgaggtgga ggtcctgcgg gagaaagtgt atgcatcact ggagacctac	540
	tgaaacaga agtaccctga gcagcaggga cggtttgcca agctgctgct acgtcttcct	600
	gccctccggt ccattggcct taagtgtcta gagcatctgt tttcttcaa gctcattggt	660
	gacaccccca tcgacacctt cctcatggag atgcttgagg ctcccatca actggcctga	720
15	<210> 5	
	<211> 635	
	<212> ADN	
	<213> <i>Locusta migratoria</i>	
	<400> 5	

ES 2 478 622 T3

tgcatacaga catgcctggt gaacgcatac ttgaagctga aaaacgagtg gagtgcaaag 60  
 cagaaaacca agtggaaatag gagctggtgg agtgggctaa acacatcccg cacttcacat 120  
 ccctacctct ggaggaccag gttctcctcc tcagagcagg ttggaatgaa ctgctaattg 180  
 cagcattttc acatcgatct gtagatgta aagatggcat agtacttgcc actggtctca 240  
 cagtgcacg aaattctgcc catcaagctg gagtcggcac aatatttgac agagttttga 300  
 cagaactggt agcaaagatg agagaaatga aaatggataa aactgaactt ggctgcttgc 360  
 gatctgttat tcttttcaat ccagaggtga ggggtttgaa atccgccag gaagttgaac 420  
 ttctacgtga aaaagtatat gccgctttgg aagaatatac tagaacaaca catcccgatg 480  
 aaccaggaag atttgcaaaa cttttgcttc gtctgccttc ttacgttcc ataggcctta 540  
 agtgtttgga gcaattgttt ttctttgcc ttattggaga tgttccaatt gatacgttcc 600  
 tgatggagat gcttgaatca ccttctgatt cataa 635

<210> 6

<211> 271

<212> ADN

5 <213> virus del herpes simple 7

<400> 6

atgggcccta aaaagaagcg taaagtcgcc cccccgaccg atgtcagcct gggggacgag 60  
 ctccacttag acggcgagga cgtggcgatg gcgcatgccg acgcgctaga cgatttcgat 120  
 ctggacatgt tgggggacgg ggattccccg gggccgggat ttacccccca cgactccgcc 180  
 ccctacggcg ctctggatat ggccgacttc gagtttgagc agatgtttac cgatgccctt 240  
 ggaattgacg agtacggtgg ggaattcccc g 271

<210> 7

<211> 1167

<212> ADN

10 <213> *Homo sapiens*

<400> 7

tgaggctccg gtgcccgtca gtgggcagag cgcacatcgc ccacagtccc cgagaagttg 60

ES 2 478 622 T3

gggggagggg tcggcaattg aaccggtgcc tagagaaggt ggcgcggggg aaactgggaa 120  
 agtgatgtcg tgtactggct ccgcctttt cccgaggggt ggggagaacc gtatataagt 180  
 gcagtagtcg ccgtgaacgt tctttttcgc aacgggtttg ccgccagaac acaggttaagt 240  
 gccgtgtgtg gttccccgcg gcctggcctc tttacgggtt atggcccttg cgtgccttga 300  
 attacttcca cctggctcca gtacgtgatt cttgatcccg agctggagcc aggggcgggc 360  
 cttgcgcttt aggagcccct tcgcctcgtg cttgagttga ggcctggcct gggcgctggg 420  
 gccgcccgtg gcgaatctgg tggcaccttc gcgcctgtct cgctgctttc gataagtctc 480  
 tagcattta aaatttttga tgacctgctg cgacgctttt tttctggcaa gatagtcttg 540  
 taaatgcggg ccaggatctg cacactggta tttcggtttt tgggcccgcg gccggcgacg 600  
 gggcccgtgc gtcccagcgc acatgttcgg cgagggggg cctgcgagcg cggccaccga 660  
 gaatcggacg ggggtagtct caagctggcc ggcctgctct ggtgcctggc ctcgcgccgc 720  
 cgtgtatcgc cccgccctgg gcggaaggc tggcccggtc ggcaccagt gctgagcgg 780  
 aaagatggcc gcttccccgc cctgctccag ggggctcaaa atggaggacg cggcgctcgg 840  
 gagagcgggc gggtgagtca cccacacaaa ggaaaaggc ctttccgtcc tcagccgtcg 900  
 cttcatgtga ctccacggag taccgggcgc cgtccaggca cctcgattag ttctggagct 960  
 tttggagtac gtcgtcttta ggttgggggg aggggtttta tgcgatggag tttcccaca 1020  
 ctgagtgggt ggagactgaa gttagccag cttggcactt gatgtaattc tcgttggaa 1080  
 ttgccctttt tgagtttggg tcttggttca ttctcaagcc tcagacagt gttcaaagtt 1140  
 ttttcttcc atttcagggt tcgtgaa 1167

<210> 8

<211> 19

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Elemento de respuesta GAL4

<400> 8

ggagtactgt cctccgagc 19

10 <210> 9

<211> 6

<212> ADN

<213> Secuencia artificial

<220>

15 <223> promotor sintético

<400> 9

tatata 6

<210> 10

<211> 1653

20 <212> ADN

<213> Secuencia artificial

<220>

<223> gen de luciferasa

<400> 10

ES 2 478 622 T3

atggaagacg ccaaaaacat aaagaaaggc ccggcgccat tctatcctct agaggatgga 60  
 accgctggag agcaactgca taaggctatg aagagatacg ccctggttcc tggacaatt 120  
 gcttttacag atgcacatat cgaggatgac atcacgtacg cggaatactt cgaaatgtcc 180  
 gttcggttgg cagaagctat gaaacgatat gggctgaata caaatcacag aatcgtcgta 240  
 tgcagtgaaa actctcttca attctttatg ccggtggtgg gcgcttatt tatcggagt 300  
 gcagttgcgc ccgcgaacga catttataat gaacgtgaat tgctcaacag tatgaacatt 360  
 tcgcagccta ccgtagtggt tgtttccaaa aaggggttgc aaaaaattht gaacgtgcaa 420  
 aaaaaattac caataatcca gaaaattatt atcatggatt ctaaacgga ttaccagggg 480  
 tttcagtcga tgtacacgtt cgtcacatct catctacctc ccggttttaa tgaatacgat 540  
 tttgtaccag agtcccttga tcgtgacaaa acaattgcac tgataatgaa ttcctctgga 600  
 tctactgggt tacctaaggg tgtggccctt ccgcatagaa ctgcctgcgt cagattctcg 660  
 catgccagag atcctattht tggcaatcaa atcattccgg atactgcgat ttttaagtgt 720  
 gttccattcc atcacggttt tggaaatgtht actacactcg gatatttgat atgtggattt 780  
 cgagtcgtct taatgtatag atttgaagaa gagctgtht tacgatccct tcaggattac 840  
 aaaattcaaa gtgcggtgct agtaccaacc ctatthtcat tcttcgcaa aagcactctg 900  
 attgacaaat acgatttatc taatthtacac gaaattgctt ctgggggagc acctcttctg 960  
 aaagaagtcg gggagcgggt tgcaaaacgc ttccatctt cagggatagc acaaggatat 1020  
 gggctcactg agactacatc agctattctg attacaccg agggggatga taaaccgggc 1080  
 gcggtcggta aagttgttcc atthtttgaa gcgaaggtt tggatctgga taccgggaaa 1140  
 acgctgggag ttaatcagag aggcgaatta tgtgtcagag gacctatgat tatgtccggt 1200  
 tatgtaaaca atccggaagc gaccaacgcc ttgattgaca aggatggatg gctacattct 1260  
 ggagacatag ctactggga cgaagacgaa cacttcttca tagttgaccg cttgaagtct 1320  
 ttaattaaat acaaaggata tcaggtggcc cccgctgaat tggaatcgat attgttacia 1380  
 caccacaaca tcttcgacgc gggcgtggca ggtcttccc acgatgacgc cggatgaact 1440  
 cccgcccggc ttgttgttht ggagcacgga aagacgatga cggaaaaaga gatcgtggat 1500  
 tacgtcgcca gtcaagtaac aaccgcgaaa aagttgcgag gaggagtgt gtttgtggac 1560  
 gaagtaccga aaggtcttac cggaaaactc gacgcaagaa aatcagaga gatcctcata 1620  
 aaggccaaga agggcggaaa gtccaaattg taa 1653

## REIVINDICACIONES

1. Compuesto seleccionado del grupo que consiste en:

- 5 a) N-terc-butil-N'-(3-hidroxiacetil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- b) N-terc-butil-N'-[3-(terc-butil-dimetil-silaniloximetil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil]-hidrazida del  
10 ácido 3,5-dimetil-benzoico,
- c) ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico,
- d) éster metílico del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-  
15 benzo[1,4]dioxin-2-carboxílico,
- e) N-terc-butil-N'-(3-semicarbazidometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- f) éster 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetílico del  
20 ácido fenil-carbámico,
- g) N'-[3-(2-amino-etil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil]-N-terc-butil-hidrazida del ácido 3,5-dimetil-benzoico,
- 25 h) éster pentafluorofenílico del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-carboxílico,
- i) metilamida del ácido 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-  
30 2-carboxílico,
- j) N-terc-butil-N'-(3-formil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- k) éster 7-[N'-terc-butil-N'-(3,5-dimetil-benzoil)-hidrazinocarbonil]-8-metil-2,3-dihidro-benzo[1,4]dioxin-2-ilmetílico del  
35 ácido tolueno-4-sulfónico,
- l) N-terc-butil-N'-[3-(hidroxiimino-metil)-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil]-hidrazida del ácido 3,5-dimetil-benzoico,
- m) N-terc-butil-N'-(3-cianometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-  
40 benzoico,
- n) N-terc-butil-N'-(5-metil-3-metilsulfanilmetil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico,
- 45 o) N-terc-butil-N'-(3-metanosulfonilmetil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-benzoico, y
- p) N-terc-butil-N'-(3-fluorometil-5-metil-2,3-dihidro-benzo[1,4]dioxin-6-carbonil)-hidrazida del ácido 3,5-dimetil-  
50 benzoico.
2. Composición farmacéutica que comprende un compuesto según la reivindicación 1 y un portador farmacéuticamente aceptable.
3. Compuesto según la reivindicación 1 para su uso en la modulación de la expresión de un gen diana en una célula huésped no humana embrionaria, aislada, en el que se pone en contacto dicha célula huésped con dicho compuesto, y en el que dicha célula huésped incluye un primer casete de expresión génica que comprende un primer polinucleótido que codifica para un primer polipéptido que comprende:
- 55 (i) un dominio de transactivación;
- 60 (ii) un dominio de unión a ADN; y
- (iii) un dominio de unión a ligando del receptor nuclear de grupo H; y
- 65 un segundo casete de expresión génica que comprende:

- (i) un elemento de respuesta que puede unirse a dicho dominio de unión a ADN;
  - (ii) un promotor que se activa mediante dicho dominio de transactivación; y
- 5 (iii) dicho gen diana.
4. Compuesto según la reivindicación 1 para su uso en el tratamiento de una enfermedad en un sujeto en el que dicho compuesto modula la expresión de uno o más genes exógenos que codifican para proteínas biológicamente activas.
- 10 5. Compuesto según la reivindicación 4 para su uso según la reivindicación 4, en el que dicha enfermedad es cáncer o un trastorno genético.
- 15 6. Compuesto según la reivindicación 4 para su uso según la reivindicación 4, en el que dichos uno o más genes exógenos codifican para proteínas secretoras, enzimas, proteínas reguladoras, receptores de superficie celular, factores de coagulación sanguíneos, hormonas, citocinas, sustancias inhibitoras, toxina diftérica, toxina colérica, anticuerpos monoclonales o proteasas.
- 20 7. Compuesto según la reivindicación 6 para su uso según la reivindicación 6, en el que dichos uno o más genes exógenos codifican para citocinas.
- 25 8. Compuesto según la reivindicación 7 para su uso según la reivindicación 7, en el que dichas citocinas son linfocinas.
9. Compuesto según la reivindicación 7 para su uso según la reivindicación 7, en el que dichas citocinas son interferones, factor estimulante de colonias de macrófagos y granulocitos, factor estimulante de colonias 1, factor de necrosis tumoral o eritropoyetina.
- 30 10. Compuesto según la reivindicación 6 para su uso según la reivindicación 6, en el que dichas enzimas metabolizan un sustrato de una sustancia tóxica a una sustancia no tóxica o metabolizan un sustrato de una sustancia inactiva a una sustancia activa.
- 35 11. Compuesto según la reivindicación 6 para su uso según la reivindicación 6, en el que dichas hormonas son insulina, hormona paratiroidea, factor liberador de hormona luteinizante, inhibinas seminales alfa y beta u hormona de crecimiento humana.
- 40 12. Compuesto según la reivindicación 6 para su uso según la reivindicación 6, en el que dicha sustancia inhibitora es alfa1-antitripsina.
- 45 13. Compuesto según la reivindicación 1 para su uso en la regulación de la expresión génica endógena o heteróloga en un sujeto transgénico no animal o no humano que comprende poner en contacto dicho compuesto con un complejo de receptor de ecdisona dentro de las células de dicho sujeto, en el que dichas células contienen además una secuencia de unión a ADN para dicho complejo de receptor de ecdisona cuando está en combinación con dicho compuesto y en el que la formación de un complejo de receptor de ecdisona-compuesto-secuencia de unión a ADN induce la expresión de dicho gen.
- 50 14. Compuesto según la reivindicación 13 para su uso según la reivindicación 13, en el que dicho complejo de receptor de ecdisona es un complejo de receptor de ecdisona quimérico y dicho constructo de ADN comprende además un promotor.
- 55 15. Compuesto según la reivindicación 13 para su uso según la reivindicación 13, en el que dicho sujeto es una planta.
16. Compuesto según la reivindicación 1 para su uso en la modulación de la expresión de un gen en una célula huésped no humana embrionaria, aislada, que comprende las etapas de:
- a) introducir en dicha célula huésped un sistema de modulación de la expresión génica que comprende:
    - i) un primer casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que codifica para un primer polipéptido híbrido que comprende:
      - (a) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y
      - 65 (b) un dominio de unión a ligando del receptor de ecdisona;

ii) un segundo casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que codifica para un segundo polipéptido híbrido que comprende:

5 (a) un dominio de transactivación; y

(b) un dominio de unión a ligando del receptor X retinoide quimérico; y

10 iii) un tercer casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que comprende:

(a) un elemento de respuesta reconocido por dicho dominio de unión a ADN de dicho primer polipéptido híbrido;

(b) un promotor que se activa mediante dicho dominio de transactivación de dicho segundo polipéptido híbrido; y

15 (c) un gen cuya expresión va a modularse; y

b) introducir en dicha célula huésped dicho compuesto.

20 17. Compuesto según la reivindicación 1 para su uso en la producción de un polipéptido en una célula huésped no humana embrionaria, aislada, que comprende las etapas de:

a) seleccionar una célula que es sustancialmente insensible a la exposición a dicho compuesto;

25 b) introducir en dicha célula:

1) un constructo de ADN que comprende:

i) un gen exógeno que codifica para el polipéptido; y

30 ii) un elemento de respuesta; en el que dicho gen está bajo el control de dicho elemento de respuesta; y

2) un constructo de ADN que codifica para un complejo de receptor de ecdisona que comprende:

35 i) un dominio de unión a ADN;

ii) un dominio de unión a ligando del receptor de ecdisona; y

iii) un dominio de transactivación; y

40 c) exponer dicha célula a dicho compuesto.

18. Uso de un compuesto según la reivindicación 1 para la fabricación de un medicamento para:

45 1) modular la expresión de un gen diana en una célula huésped no humana embrionaria, en el que la célula huésped incluye un primer casete de expresión génica que comprende un primer polinucleótido que codifica para un primer polipéptido que comprende:

(i) un dominio de transactivación;

50 (ii) un dominio de unión a ADN; y

(iii) un dominio de unión a ligando del receptor nuclear de grupo H; y

un segundo casete de expresión génica que comprende:

55 (i) un elemento de respuesta que puede unirse a dicho dominio de unión a ADN;

(ii) un promotor que se activa mediante dicho dominio de transactivación; y

60 (iii) dicho gen diana;

2) modular la expresión de uno o más genes exógenos en un sujeto;

65 3) regular la expresión génica endógena o heteróloga en un sujeto transgénico no humano que comprende poner en contacto dicho compuesto con un complejo de receptor de ecdisona dentro de las células de dicho sujeto, en el que dichas células contienen además una secuencia de unión a ADN para dicho complejo de receptor de ecdisona

cuando está en combinación con dicho compuesto y en el que la formación de un complejo de complejo de receptor de ecdisona-compuesto-secuencia de unión a ADN induce la expresión de dicho gen;

4) modular la expresión de un gen en una célula huésped que comprende las etapas de:

- 5
- a) introducir en dicha célula huésped un sistema de modulación de la expresión génica que comprende:
- 10
- i) un primer casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que codifica para un primer polipéptido híbrido que comprende:
- 15
- (a) un dominio de unión a ADN que reconoce un elemento de respuesta asociado con un gen cuya expresión va a modularse; y
- (b) un dominio de unión a ligando del receptor de ecdisona;
- 20
- ii) un segundo casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que codifica para un segundo polipéptido híbrido que comprende:
- 25
- (a) un dominio de transactivación; y
- (b) un dominio de unión a ligando del receptor X retinoide quimérico; y
- 30
- iii) un tercer casete de expresión génica que puede expresarse en dicha célula huésped que comprende una secuencia de polinucleótido que comprende:
- 35
- (a) un elemento de respuesta reconocido por dicho dominio de unión a ADN de dicho primer polipéptido híbrido;
- (b) un promotor que se activa mediante dicho dominio de transactivación de dicho segundo polipéptido híbrido; y
- 40
- (c) un gen cuya expresión va a modularse; y
- b) introducir en dicha célula huésped dicho compuesto; o
- 5) producir un polipéptido que comprende las etapas de:
- 45
- a) seleccionar una célula que es sustancialmente insensible a la exposición a dicho compuesto;
- b) introducir en dicha célula:
- 50
- 1) un constructo de ADN que comprende:
- i) un gen exógeno que codifica para el polipéptido; y
- ii) un elemento de respuesta; en el que dicho gen está bajo el control de dicho elemento de respuesta; y
- 45
- 2) un constructo de ADN que codifica para un complejo de receptor de ecdisona que comprende:
- i) un dominio de unión a ADN;
- 50
- ii) un dominio de unión a ligando del receptor de ecdisona; y
- iii) un dominio de transactivación; y
- c) exponer dicha célula a dicho compuesto.