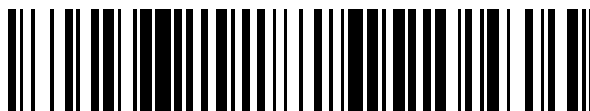


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 409 080**

51 Int. Cl.:

C07K 14/00 (2006.01)

C07K 14/47 (2006.01)

C12N 9/10 (2006.01)

C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **22.01.2002 E 02707545 (6)**

97 Fecha y número de publicación de la concesión europea: **13.03.2013 EP 1353941**

54 Título: **Proteínas de unión con dedos de zinc modificadas**

30 Prioridad:

22.01.2001 US 263445 P

11.05.2001 US 290716 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
24.06.2013

73 Titular/es:

**SANGAMO BIOSCIENCES INC. (100.0%)
POINT RICHMOND TECH CENTER, SUITE A100,
501 CANAL BOULEVARD
RICHMOND, CA 94804, US**

72 Inventor/es:

**REBAR, EDWARD y
JAMIESON, ANDREW**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 409 080 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

Proteínas de unión con dedos de zinc modificadas

Campo técnico

- 5 Los métodos y composiciones descritos en la presente memoria se refieren en general al campo de la regulación de la expresión génica y específicamente a métodos para modular la expresión génica mediante la utilización de polipéptidos derivados de proteínas de unión a nucleótidos con dedos de zinc.

Antecedentes

- 10 La unión de proteínas a secuencias específicas de ADN, ARN, proteínas y otras moléculas participa en cantidad de procesos celulares tales como, por ejemplo, la transcripción, la replicación, la estructura de la cromatina, la recombinación, la reparación del ADN, el procesamiento del ARN y la traducción. La especificidad de unión de las proteínas de unión celulares que participan en las interacciones proteína-ADN, proteína-ARN y proteína-proteína contribuye al desarrollo, la diferenciación y la homeostasis. Las alteraciones en las interacciones de proteínas específicas pueden estar implicadas en varios tipos de patologías tales como, por ejemplo, el cáncer, la enfermedad cardiovascular y la infección.

- 15 Las proteínas con dedos de zinc (ZFP) son proteínas que pueden unirse a una secuencia de ADN de forma específica. Los dedos de zinc se identificaron por primera vez en el factor de transcripción TFIIIA de los ovocitos de la rana africana de uñas, *Xenopus laevis*. Un único dominio de un dedo de zinc de esta clase de ZFP tiene una longitud de aproximadamente 30 aminoácidos, y varios estudios estructurales han demostrado que contiene un giro beta (que contiene los dos restos invariables de cisteína) y una hélice alfa (que contiene los dos restos invariables de histidina), que adopta una conformación particular mediante la coordinación de un átomo de zinc con dos cisteínas y dos histidinas. Esta clase de ZFP se conoce también como ZFP C2H2. También se han propuesto otras clases de ZFP. (Véase, por ejemplo, Jiang *et al.* (1996) *J. Biol. Chem.* **271**:10723-10730 for a discussion of Cys-Cys-His-Cys (C3H) ZPFs.) Hasta la fecha, se han identificado más de 10.000 secuencias de dedos de zinc en varios miles de factores de transcripción conocidos o hipotéticos. Los dominios de los dedos de zinc están implicados no sólo en el reconocimiento del ADN, sino también en la unión al ARN y en la unión proteína-proteína. Las estimaciones actuales son que esta clase de moléculas constituirá aproximadamente el 2% de todos los genes humanos.

- 20 La mayoría de las proteínas con dedos de zinc han conservado restos de cisteína e histidina que coordinan tetraédricamente con el único átomo de zinc de cada dominio del dedo. En particular, la mayoría de las ZFP se caracterizan por componentes de los dedos de la secuencia general: -Cys(X)_{2,4}-Cys(X)₁₂-Su(X)_{3,5}-His (SEQ ID NO: 1), donde X es cualquier aminoácido (ZFP C2H2). Las secuencias de coordinación de zinc de la clase más ampliamente representada contienen dos cisteínas y dos histidinas con espaciados particulares, por ejemplo, los dedos de zinc que se encuentran en la proteína ADRI de levadura, la proteína ZFY asociada a humanos varones, la proteína potenciadora del VIH y la proteína Xfin de *Xenopus* han sido resueltas por métodos de RMN de alta resolución (Kochoyan, *et al.*, *Biochemistry*, **30**:3371-3386, 1991; Omichinski, *et al.*, *Biochemistry*, **29**:9324-9334, 1990; Lee, *et al.*, *Science*, **245**:635-637, 1989). Basándose en la cristalografía de rayos X, se ha resuelto la estructura tridimensional de un complejo ADN-polipéptido de tres dedos derivado de la proteína temprana inmediata de ratón zif268 (también conocida como Krox-24). (Pavletich and Pabo, *Science*, **252**:809-817, 1991). La estructura plegada de cada dedo contiene un giro β antiparalelo, una región punta de dedo y una α -hélice anfipática corta. Los ligandos de coordinación del metal se unen al ion Zn y, en el caso de los dedos de zinc zif268, la α -hélice anfipática corta se une al surco mayor del ADN. Además, los aminoácidos hidrófobos conservados y la coordinación del zinc por los restos cisteína e histidina estabilizan la estructura del dominio del dedo individual.

- 25 El plegado de una ZFP C2H2 en la estructura de dedo adecuada puede ser completamente interrumpido por el cambio de los aminoácidos ligando de C2H2. Miura *et al.* (1998) *Biochim. Biophys. Acta* **1384**:171-179. Además, se puede alterar la especificidad de unión al metal de los péptidos basada en la secuencia consenso de C2H2. Krizek *et al.* (1993) *Inorg. Chem.* **32**:937-940; Merkle *et al.* (1991) *J. Am. Chem. Soc.* **113**:5450-5451. Aunque también se han propuesto modelos detallados para la interacción de los dedos de zinc y el ADN (Berg, 1988; Berg, 1990; Churchill, *et al.*, 1990), las mutaciones en el dedo 2 de la ZFP C2H2 zif268 de tres dedos han demostrado que eliminan totalmente la unión al ADN (Green *et al.* (1998) *Biochem. J.* **333**:85-90).

- 30 No obstante, una mayor comprensión de la naturaleza y el mecanismo de la especificidad de unión a proteínas ha alentado la esperanza de que la especificidad de unión de una proteína podría ser alterada de una manera predecible, o que se podría construir de novo una proteína de unión de especificidad predeterminada. Véase, por ejemplo, Blackburn (2000) *Curr. Opin. Struct. Biol.* **10**:399-400; Segal *et al.* (2000) *Curr. Opin. Chem. Biol.* **4**:34-39.

- 35 Para este fin, se han hecho intentos de modificar proteínas con dedos de zinc C2H2. Véase, por ejemplo, las patentes de EE.UU. nº 6.007.988; 6.013.453; 6.140.081; PCT WO 98/53057; PCT WO 98/53058; PCT WO 98/53059; PCT WO 98/53060; PCT WO 00/23464; PCT WO 00/42219; Choo *et al.* (2000) *Curr. Opin. Struct. Biol.* **10**:411-416; Segal *et al.* (2000) *Curr. Opin. Chem. Biol.* **4**:34-39, y las referencias citadas en estas publicaciones.

Hasta la fecha, sin embargo, en los estudios celulares utilizando ZFP C2H2 de diseño se han utilizado relativamente pocas posiciones en el dedo de zinc como parámetros ajustables para obtener una actividad óptima. En particular, hasta la fecha los estudios han modificado sólo aquellos restos en la interfaz dedo-ADN. Estos incluyen posiciones conocidas por hacer contactos directos con la base, restos de 'apoyo' o 'yuxtaposición' inmediatamente adyacentes a las posiciones de contacto de la base, y posiciones capaces de contactar con el esqueleto de fosfato del ADN. Además, muchos de los efectos observados han sido bastante modestos, y la posibilidad de que la mejora de las actividades de ZFP se podría lograr a través de la sustitución de restos en otras posiciones del dedo o utilizando polipéptidos no-C2H2 queda completamente sin investigar.

Por tanto, existe una necesidad de diseños adicionales o de proteínas de unión con dedos de zinc seleccionadas.

10 Compendio

En la presente memoria se describen proteínas de unión, en particular proteínas de unión con dedos de zinc, con sitios de coordinación metálicos modificados. Se proporcionan también métodos de fabricación y uso de estas proteínas. Según la presente invención, las proteínas de unión contienen tres dedos de zinc coordinados y el tercero de estos dedos es un componente del dedo modificado, no canónico (por ejemplo, no-C2H2). Preferiblemente, el tercer dedo de una ZFP con tres dedos está modificado y no canónico.

La proteína de unión con dedo de zinc aislada, no canónica de la presente invención se une a una secuencia diana. La proteína de unión con dedo de zinc aislada se puede proporcionar como una molécula de ácido nucleico o como un polipéptido. Además, la secuencia diana puede ser un aminoácido, ADN (por ejemplo, una secuencia promotora) o ARN y, adicionalmente, puede estar en una célula procariota (por ejemplo, bacterias) o eucariota (por ejemplo, célula vegetal, célula de levadura, célula fúngica, célula animal humana).

La secuencia de aminoácidos del componente del tercer dedo de zinc es $X_3\text{-Cys-X}_{2-4}\text{-Cys-X}_{12}\text{-His-X}_{1-7}\text{-Z-X}_4$ en la que X es cualquier aminoácido, y Z es Cys.

Las proteínas con dedos de zinc modificadas descritas en la presente memoria pueden incluir cualquier número componentes del coordinados en el que uno o más de los dedos de zinc coordinados son no canónicos. Según la presente invención, la ZFP comprende tres dedos, en donde el componente del tercer dedo es no canónico. En otras realizaciones, cualquiera de las ZFP descrita en este documento comprende un esqueleto de ZFP vegetal modificado.

En otros aspectos, se proporcionan los polipéptidos de fusión que comprenden (a) cualquiera de las proteínas de unión con dedos de zinc descritas en la presente memoria y (b) al menos un dominio funcional. El dominio funcional puede ser, por ejemplo, un dominio represor tales como KRAB, MBD-2B, v-ErbA, MBD3, TR, y miembros de la familia DNMT; un dominio de activación, tales como VP16, la subunidad p65 de NF-kappa B, y VP64; un dominio aislante; una proteína de remodelación de la cromatina, y/o un dominio de unión a metilo.

En otros aspectos, se proporcionan los polinucleótidos que codifican cualquiera de las proteínas con dedos de zinc (o moléculas de fusión) descritos aquí. También se proporcionan los vectores de expresión y las células huésped que comprenden estos polinucleótidos.

En otros aspectos, se proporciona un método de modulación de la expresión de un gen. El método comprende la etapa de poner en contacto una región de ADN con cualquiera de los dedos de zinc que contienen las moléculas de fusión que se describen en la presente memoria.

En ciertas realizaciones, la proteína de unión del dedo de zinc de la molécula de fusión se une a un sitio diana en un gen que codifica un producto seleccionado del grupo que consiste en el factor de crecimiento endotelial vascular, la eritropoyetina, receptor de andrógenos, PPAR- γ 2, p16, p53, pRb, distrofina y e-cadherina, desaturasa delta-9, desaturasa delta-1 2 de otras plantas, desaturasa delta-1 5, acetil-CoA carboxilasa, acil-ACP-tioesterasa, ADP-glucosa pirofosforilasa, almidón sintasa, celulosa sintasa, sacarosa sintasa, genes asociados a la senescencia, quelantes de metales pesados, liasa de hidroperóxidos de ácidos grasos, poligalacturonasa, EPSP sintasa, genes de virus vegetales, genes de hongos patógenos vegetales, y genes de bacterias patógenas vegetales. (Véase, también el documento WO 00/41566). El gen puede en cualquier célula, por ejemplo una célula vegetal o célula animal (por ejemplo, humana).

En otros aspectos adicionales, se proporcionan composiciones que comprenden cualquiera de las moléculas de proteínas con dedos de zinc (o de fusión) descritas en la presente memoria y un excipiente farmacéuticamente aceptable.

Estas y otras realizaciones se les ocurrirán fácilmente a los expertos en la técnica a la luz de la presente descripción.

Breve descripción de los dibujos

Figura 1 es un gráfico que representa los niveles del ARNm del gen LCK (normalizado respecto a los niveles de RNAr 18s) en células transfectadas con construcciones que codifican fusiones del dominio de activación VP16 con una ZFP canónica (PTP2), una ZFP modificada (PTP2 (H->C), y una construcción de control (NVF).

- 5 Figura 2 muestra los niveles de VEGF-A en el medio de cultivo de células que habían sido transfectadas con plásmidos que codifican proteínas de fusión ZFP no canónicas que comprenden un dominio de activación VP16, que fueron dirigidos al gen VEGF.

- 10 Prueba indica las células no transfectadas; vector vacío indica la transfección con una construcción de ADN que carece de secuencias que codifican una proteína de fusión, y C2H2 indica células transfectadas con plásmidos que codifican proteínas de fusión VOP30A C2H2 y VOP32B ZFP-VP16 canónicas. S, E, K, CT, C, GC y GGC indican derivados de VOP30A y VOP 32B no canónicos que contienen un dedo de zinc C2HC, como se describe en la Tabla 1. La barra izquierda de cada par muestra los resultados para VOP30A y sus derivados no canónicos, la barra de la derecha de cada par muestra los resultados para VOP32B y su derivado no canónico. El derivado C de VOP32B y el derivado GC de VOP30A no fueron probados. Los resultados son el promedio de dos determinaciones.

- 15 Figura 3, los paneles A y B, son esquemas que muestran la construcción del vector de expresión YCF3 útil en la expresión de ZFP modificadas.

- 20 Figura 4 muestra los resultados del análisis de mRNA de GMT en ARN aislado a partir de protoplastos de *Arabidopsis thaliana* que había sido transfectados con construcciones que codifican la fusión de un dominio de activación transcripcional con varias ZFP vegetales modificada. Los resultados se expresan como ARNm de GMT normalizado respecto RNAr 18S. Los números AGMT en el eje de abscisas se refieren a los dominios de unión de ZFP de plantas modificadas que se muestran en la Tabla 2.

Se muestran duplicados TaqMan (R) de los análisis para cada muestra de ARN.

Descripción detallada

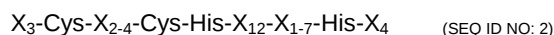
General

- 25 La presente invención proporciona polipéptidos de unión a dedos de zinc no canónicos, aislados (ZFP), en el que uno o más de los componentes con dedos de zinc difiere de la secuencia consenso canónico de Cys-Cys-His-His (por ejemplo, Cys2-His2). El polipéptido puede ser un polipéptido de fusión y, por sí mismo o como parte de una fusión, puede potenciar o suprimir la transcripción de un gen, y puede unirse al ADN, ARN y / o proteína.

- 30 Los polinucleótidos que codifican no canónicos ZFP y proteínas de fusión que comprenden uno o más países no canónicos ZFP también se proporcionan. Adicionalmente, se proporcionan composiciones farmacéuticas que comprenden una cantidad terapéuticamente eficaz de cualquiera de los polipéptidos de unión a nucleótidos con dedos zinc modificados descritos en la presente memoria o fragmentos funcionales de los mismos, o una cantidad terapéuticamente eficaz de una secuencia de nucleótidos que codifica cualquiera de los polipéptidos de unión nucleótido-dedo de zinc modificados o fragmentos funcionales de los mismos, en donde el polipéptido del dedo de zinc o el fragmento funcional del mismo se une a una secuencia celular de nucleótidos para modular la función de la secuencia celular de nucleótidos, en combinación con un vehículo farmacéuticamente aceptable. También se proporcionan métodos de selección para la obtención de un polipéptido de unión nucleótido-dedo de zinc modificado que se une a una secuencia de nucleótidos celular o viral.

- 40 En la actualidad, las ZFP diseñadas y/o seleccionadas utilizan relativamente pocas posiciones en el dedo de zinc como parámetros ajustables para obtener una actividad óptima. En particular, los estudios hasta la fecha han alterado sólo aquellos restos en la interfaz dedo-ADN. Véase, por ejemplo, las patentes de EE.UU. nº 6.007.988; 6.013.453; 6.140.081 y 6.140.466, así como el documento PCT WO 00/42219. Como se señaló anteriormente, los efectos observados han sido bastante modestos, y la posibilidad de que las actividades mejoradas de ZFP podrían ser accesibles a través de la sustitución de restos en otras posiciones en el dedo no ha sido investigada.

- 45 Por consiguiente, en la presente invención, se describen proteínas con dedos de zinc modificadas (por ejemplo, no canónica) que comprenden tres dedos de zinc en las que la secuencia de la tercera ZFP difiere de la secuencia consenso canónica que contiene dos restos cisteína (Cys) y dos restos histidina (His):



- 50 (también conocida como la secuencia consenso "Cys2-His2" o "C2H2"). Como la coordinación del zinc proporciona la energía de plegado principal para los dedos de zinc, el ajuste de los restos de coordinación del zinc parece proporcionar un medio fácil para la modificación de la estabilidad y estructura del dedo, que podría afectar a una variedad de características funcionales importantes de los factores de transcripción-proteína con dedos de zinc. En particular, sería de esperar que todas las características tales como vida media celular, las interacciones con otros

factores celulares, especificidad y afinidad de unión a ADN, y la orientación relativa de los dominios funcionales se viesen influidas por la elección del resto de las posiciones de coordinación de zinc.

Así, en la presente invención, el tercer dedo de coordinación de zinc que forma la proteína de dedo de zinc tiene de la siguiente secuencia:

5 $X_3\text{-Cys-X}_{2,4}\text{-Cys-X}_{12}\text{-His-X}_{1,7}\text{-Z-X}_4$

donde X = cualquier aminoácido

Z = Cys

También se describen en la presente memoria los dedos de coordinación de zinc que tienen cualquiera de las siguientes secuencias:

$X_3\text{-B-X}_{2,4}\text{-Cys-X}_{12}\text{-His-X}_{1,7}\text{-His-X}_4$
 $X_3\text{-Cys-X}_{2,4}\text{-B-X}_{12}\text{-His-X}_{1,7}\text{-His-X}_4$
 $X_3\text{-Cys-X}_{2,4}\text{-Cys-X}_{12}\text{-Z-X}_{1,7}\text{-His-X}_4$
 $X_3\text{-Cys-X}_{2,4}\text{-Cys-X}_{12}\text{-His-X}_{1,7}\text{-Z-X}_4$
 $X_3\text{-B-X}_{2,4}\text{-B-X}_{12}\text{-His-X}_{1,7}\text{-His-X}_4$
 $X_3\text{-B-X}_{2,4}\text{-Cys-X}_{12}\text{-Z-X}_{1,7}\text{-His-X}_4$
 $X_3\text{-B-X}_{2,4}\text{-Cys-X}_{12}\text{-His-X}_{1,7}\text{-Z-X}_4$
 $X_3\text{-Cys-X}_{2,4}\text{-B-X}_{12}\text{-Z-X}_{1,7}\text{-His-X}_4$
 $X_3\text{-Cys-X}_{2,4}\text{-B-X}_{12}\text{-His-X}_{1,7}\text{-Z-X}_4$
 $X_3\text{-Cys-X}_{2,4}\text{-Cys-X}_{12}\text{-Z-X}_{1,7}\text{-Z-X}_4$
 $X_3\text{-Cys-X}_{2,4}\text{-B-X}_{12}\text{-Z-X}_{1,7}\text{-Z-X}_4$
 $X_3\text{-B-X}_{2,4}\text{-Cys-X}_{12}\text{-Z-X}_{1,7}\text{-Z-X}_4$
 $X_3\text{-B-X}_{2,4}\text{-B-X}_{12}\text{-His-X}_{1,7}\text{-Z-X}_4$
 $X_3\text{-B-X}_{2,4}\text{-B-X}_{12}\text{-Z-X}_{1,7}\text{-His-X}_4$
 $X_3\text{-B-X}_{2,4}\text{-B-X}_{12}\text{-Z-X}_{1,7}\text{-Z-X}_4$

10

donde X = cualquier aminoácido

B = cualquier aminoácido excepto cisteína

Z = cualquier aminoácido excepto histidina

15

En la nomenclatura estándar para ZFP, el "primer" dedo es el dedo más N-terminal de la proteína (con respecto a los otros dedos) y se une al extremo del subsitio del triplete -(o cuadruplete) más 3' en el sitio de destino. Los dedos adicionales, que se desplazan hacia el C-terminal de la proteína, se numeran secuencialmente. En la presente invención se proporciona una proteína con dedos de zinc que comprende tres dedos en la que los dos primeros dedos son de la clase C2-H2, pero el segundo resto histidina del tercer dedo (y opcionalmente los restos de aminoácido adyacentes) está sustituido por Cys o con Cys y aminoácidos adicionales, tales como la glicina.

20

También se describe en la presente memoria, una proteína con dedos de zinc de tres dedos en la que el primer o el segundo resto de cisteína del dedo primero están sustituidos por histidina o con histidina y otros aminoácidos tales como glicina. Además, en la presente memoria se describe, un dedo de una proteína con dedos de zinc modificado de modo que, en uno o más de los dedos, se sustituyen uno o más restos cisteína o histidina, con un aminoácido diferente tal como, por ejemplo, serina. En la presente invención se describe que, el segundo dedo de una proteína

25

con dedos de zinc de tres dedos se modifica de modo que uno o ambos restos de cisteína se sustituyen con serina (y/o aminoácidos adicionales). Además, los aminoácidos que contienen carboxilo, tales como, por ejemplo, ácido aspártico y ácido glutámico son sustituidos por cisteína y/o histidina en un dedo de zinc. Además, se proporcionan ZFP que comprenden dos o más dedos en que los también se se modifica más de un dedo.

30

Por lo tanto, las ZFP de la presente invención difieren de los factores de transcripción proteínas con dedos de zinc de diseño descritos previamente en que comprenden al menos un dedo de zinc de coordinación que difiere de la secuencia consenso canónica (Cys-Cys-His-His). Es evidente que diversas combinaciones de dedos de zinc modificados puede ser utilizadas en una única proteína, por ejemplo, todos los componentes de los dedos pueden ser modificados usando los mismos o diferentes dedos de zinc modificados. Alternativamente, no todos los dedos se pueden modificar utilizando los mismos dedos o diferentes modificados. Además, los componentes de los dedos modificados no canónicos descritos en la presente memoria también se pueden utilizar en combinación con los componentes de los dedos C2H2 de ZFP anteriormente descritos.

35

En realizaciones adicionales, los de zinc aislados no canónicos de la presente invención se utilizan en las proteínas de fusión, por ejemplo fusiones de un dominio de unión ADN-ZFP con dominios de represión o activación o con dominios de remodelación de la cromatina. También se proporcionan los polinucleótidos que codifican cualquiera de las proteínas con dedos de zinc, sus componentes y sus fusiones.

- 5 La práctica de los métodos descritos, a menos que se indique lo contrario, emplea técnicas convencionales de biología molecular, bioquímica, genética, química computacional, cultivo celular, ADN recombinante y campos relacionados dentro del este sector de la técnica. Estas técnicas se explican totalmente en la literatura. Véase, por ejemplo, Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL, Segunda edición, Cold Spring Harbor Laboratory Press, 1989, Ausubel y otros, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987, y actualizaciones periódicas, y la serie METHODS IN ENZYMOLOGY, Academic. Press, San Diego.

Definiciones

- 15 Los términos "ácido nucleico", "polinucleótido" y "oligonucleótido" se usan indistintamente y se refieren a un polímero de desoxirribonucleótidos o ribonucleótidos en forma monocatenaria o bicatenaria. Para los fines de la presente descripción, estos términos no deben interpretarse como limitantes con respecto a la longitud de un polímero. Los términos puede abarcar análogos conocidos de nucleótidos naturales, así como los nucleótidos con modificaciones en las bases, azúcares y/o fosfatos. En general, un análogo de un nucleótido particular tiene la misma especificidad de emparejamiento de bases, es decir, un análogo de A emparejará con T.

- 20 Los términos "polipéptido", "péptido" y "proteína" se usan indistintamente para referirse a un polímero de restos de aminoácidos. El término también se aplica a polímeros de aminoácidos en los que uno o más aminoácidos son análogos químicos o derivados modificados de un aminoácido correspondiente de origen natural, por ejemplo selenocisteína (Bock et al.(1991) *Trends Biochem.Sci.***16**:463-467; Nasim et al. (2000) *J. Biol.Chem.* **275**:14846-14852) y similares.

- 25 Una "proteína de unión" es una proteína que es capaz de unirse no covalentemente a otra molécula. Una proteína de unión se puede unir a, por ejemplo, una molécula de ADN (una proteína de unión a ADN), a una molécula de ARN (una proteína de unión a ARN) y/o a una molécula de proteína (una proteína de unión a proteínas). En el caso de una proteína de unión a proteínas, se puede unir a sí misma (para formar homodímeros, homotrímeros, etc) y/o se puede unir a una o más moléculas de una proteína o proteínas diferentes. Una proteína de unión puede tener más de un tipo de actividad de unión. Por ejemplo, las proteínas con dedos de zinc tienen actividad de unión al ADN, al ARN y a proteína. Un "perfil de unión" se refiere a una pluralidad de secuencias diana que son reconocidas y unidas a una proteína de unión particular. Por ejemplo, un perfil de unión se puede determinar poniendo en contacto una proteína de unión con una población de secuencias diana aleatoria para identificar una sub-población de secuencias diana unida a esa proteína de unión particular.

- 35 Una "proteína de unión con dedos de zinc" es una proteína o un segmento dentro de una proteína más grande que se une de una manera específica a una secuencia de ADN, ARN y/o proteína como resultado de la estabilización de la estructura de la proteína a través de la coordinación de un ion de zinc. El término proteína de unión con dedos de zinc se abrevia a menudo como proteína con zinc dedo o ZFP. Un dedo de zinc "canónico" se refiere a un componente de coordinación-zinc (por ejemplo, dedo de zinc) de una proteína con dedos de zinc que tiene la secuencia de aminoácidos general: $X_3\text{-Cys-}X_{2-4}\text{-Cys-His-}X_{12}\text{-}X_{1-7}\text{-Su-}X_4$ donde X es cualquier aminoácido (también conocido como un dedo de zinc C2H2).

- 45 Una proteína con dedos de zinc "modificada" es una proteína no presente en la naturaleza que se ha diseñado y/o seleccionado de modo que comprenda una sustitución de al menos un aminoácido, en comparación con una proteína con dedos de zinc de origen natural. Además, una proteína con dedos de zinc "de diseño" es una proteína que no ocurre en la naturaleza, cuya estructura y composición resulta principalmente de criterios racionales. Los criterios racionales de diseño incluyen la aplicación de reglas de sustitución y algoritmos informáticos para procesar la información en una base de datos de almacenamiento de información de diseños de ZFP existentes y de datos de unión, por ejemplo como se describe en la solicitud de propiedad conjunta PCT WO 00/42219. Una proteína con dedos de zinc "seleccionada" es una proteína que no se encuentra en la naturaleza cuya producción resulta en primer lugar a partir de un proceso empírico tal como expresión de fagos. Véase, por ejemplo, EE.UU 5.789.538; EE.UU 6.007.988; EE.UU 6.013.453, WO 95/19431, WO 96/06166 y WO 98/54311. También se hace referencia a las ZFP de diseño y/o seleccionadas como ZFP "de ingeniería" y se pueden modificar de acuerdo con los métodos y composiciones descritas en este documento (por ejemplo, mediante la conversión a C3H y/o que comprenden un esqueleto de una planta).

- 55 El término "de origen natural" se usa para describir un objeto que puede encontrarse en la naturaleza, a diferencia del producido artificialmente por un humano.

Un "esqueleto" de dedo de zinc es la porción de un dedo de zinc fuera de la región implicada en las interacciones del surco mayor del ADN, es decir, las regiones del dedo de zinc exteriores a los restos -1 a +6 de la hélice alfa. El

esqueleto comprende las láminas beta, la zona de unión entre la segunda lámina beta y la hélice alfa, la porción de hélice alfa distal al primer resto histidina conservado, y la(s) secuencia(s) de enlace entre los dedos.

Las secuencias de ácidos nucleicos o de aminoácidos están "unidas operativamente" (o "unidas operativamente") cuando se disponen en una relación funcional entre sí. Por ejemplo, un promotor o potenciador está unido operativamente a una secuencia codificante si regula, o contribuye a la modulación de la transcripción de la secuencia codificante. Las secuencias de ADN unidas operativamente son típicamente contiguas, y unidas operativamente secuencias de aminoácidos están típicamente contiguas y en el mismo marco de lectura. Sin embargo, puesto que los potenciadores funcionan generalmente cuando están separados del promotor por hasta varias kilobases o más, y las secuencias intrónicas pueden ser de longitudes variables, algunos elementos de polinucleótidos pueden unirse operativamente pero estar no contiguos. De manera similar, determinadas secuencias de aminoácidos que están no contiguas en una secuencia de polipéptidos primaria pueden sin embargo estar unidas operativamente debido a, por ejemplo el plegamiento de una cadena polipeptídica.

Con respecto a los polipéptidos de fusión, el término "unido operativamente" puede referirse al hecho de que cada uno de los componentes realiza la misma función unido al otro componente que realizaría si no estuviera unido. Por ejemplo, con respecto a un polipéptido de fusión en el que el dominio de unión-ADN de ZFP está fusionado a un dominio de activación transcripcional (o un fragmento funcional del mismo), el dominio de unión-ADN de ZFP y el dominio de activación transcripcional (o un fragmento funcional del mismo) están en unión operativa si, en el polipéptido de fusión, la porción del dominio de unión-ADN de ZFP es capaz de unir su sitio diana y/o su sitio de unión, mientras que el dominio de activación transcripcional (o un fragmento funcional del mismo) es capaz de activar la transcripción.

"Unión específica" entre, por ejemplo, una ZFP y un sitio diana específico significa una afinidad de unión de al menos $1 \times 10^6 \text{ M}^{-1}$.

Una "molécula de fusión" es una molécula en la que dos o más moléculas de subunidades están unidas, preferentemente de forma covalente. Las moléculas de subunidades pueden ser moléculas del mismo grupo químico, o pueden ser moléculas de grupos químicos diferentes. Ejemplos del primer grupo de moléculas de fusión incluyen, pero no se limitan a, polipéptidos de fusión (por ejemplo, una fusión entre un dominio de unión-ADN de ZFP y un dominio de activación transcripcional) y ácidos nucleicos de fusión (por ejemplo, un ácido nucleico que codifica el polipéptido de fusión descrito en la presente memoria). Ejemplos del segundo tipo de molécula de fusión incluyen, pero no se limitan a, una fusión entre un ácido nucleico de forma triplex y un polipéptido, y una fusión entre un ligando del surco menor y un ácido nucleico.

Un "gen", para los fines de la presente descripción, incluye una región de ADN que codifica un producto génico (véase más adelante), así como todas las regiones de ADN que regulan la producción del producto génico, ya sean o no tales secuencias reguladoras adyacentes a las secuencias de codificación y/o transcritas. En consecuencia, un gen incluye, pero no se limita necesariamente a, secuencias promotoras, terminadores, secuencias traslacionales reguladoras tales como sitios de unión a ribosomas y sitios internos de entrada al ribosoma, potenciadores, silenciadores, aislantes, elementos de contorno, orígenes de replicación, sitios de fijación a la matriz y regiones de control del locus. Además, un promotor puede ser un promotor celular normal o, por ejemplo, un promotor de un microorganismo infeccioso tal como, por ejemplo, una bacteria o un virus. Por ejemplo, la repetición terminal larga (LTR) de retrovirus es una región promotora que puede ser una diana para un polipéptido de unión con dedo de zinc modificado. Los promotores de miembros del grupo *Lentivirus*, que incluyen patógenos tales como el virus linfotrópico de células T humano (HTLV) 1 y 2, o virus de la inmunodeficiencia humana (VIH) 1 o 2, son ejemplos de regiones promotoras virales que pueden ser dianas para la modulación transcripcional por un polipéptido de unión con dedo de zinc modificado como se describe en la presente memoria.

"Expresión génica" se refiere a la conversión de la información, contenida en un gen, en un producto génico. Un producto génico puede ser el producto directo de la transcripción de un gen (por ejemplo, ARNm, ARNt, ARNr, ARN antisentido, ribozima, ARN estructural o cualquier otro tipo de ARN) o una proteína producida por traducción de un ARNm. Los productos génicos también incluyen los ARN que son modificados, por procesos tales como "capping", poliadenilación, metilación, y edición, y proteínas modificadas, por ejemplo, metilación, acetilación, fosforilación, ubiquitinación, ADP-ribosilación, miristilación, y glicosilación.

"Activación génica" y "aumento de la expresión génica" se refieren a cualquier proceso que tenga como resultado un aumento en la producción de un producto génico. Un producto génico puede ser ARN (incluyendo, pero sin limitarse a, ARNm, ARNr, ARNt, y ARN estructural) o una proteína. En consecuencia, la activación génica incluye aquellos procesos que aumentan la transcripción de un gen y/o la traducción de un ARNm. Ejemplos de procesos de activación génica que aumentan la transcripción incluyen, pero no se limitan a, aquellos que facilitan la formación de un complejo de iniciación de la transcripción, aquellos que aumentan la tasa de iniciación de la transcripción, aquellos que aumentan la tasa de elongación de la transcripción, aquellos que aumentan la capacidad de procesamiento de la transcripción y aquellas que disminuyen la represión transcripcional (por ejemplo, mediante el bloqueo de la unión de un represor transcripcional). La activación génica puede constituir, por ejemplo, la inhibición de la represión, así como la estimulación de la expresión por encima de un nivel existente. Ejemplos de procesos de activación génica que aumentan la traducción incluyen aquellos que aumentan iniciación de la traducción, aquellos

que aumentan la elongación de traslación y aquellos que aumentan la estabilidad del ARNm. En general, la activación génica comprende cualquier incremento detectable en la producción de un producto génico, preferiblemente un aumento del doble en la producción de un producto génico, más preferiblemente aproximadamente de 2 a aproximadamente 5 veces o cualquier valor entero entre los mismos, más preferiblemente de aproximadamente 5 a aproximadamente 10 veces o cualquier valor entero entre los mismos, más preferiblemente entre aproximadamente 10 y aproximadamente 20 veces o cualquier valor entero entre las mismas, aún más preferiblemente entre aproximadamente 20 y aproximadamente 50 veces o cualquier valor entero entre ellas, más preferiblemente entre aproximadamente 50 - y aproximadamente 100 veces o cualquier valor entero entre los mismos, más preferiblemente 100 veces o más.

"Represión génica" y "inhibición de la expresión génica" se refieren a cualquier proceso que tenga como resultado una disminución en la producción de un producto génico. Un producto génico puede ser bien ARN (incluyendo, pero sin limitarse a, ARNm, ARNr, ARNt, y ARN estructural) o proteína. Por consiguiente, la represión génica incluye aquellos procesos que disminuyen la transcripción de un gen y/o la traducción de un ARNm. Ejemplos de procesos de represión génica que disminuyen la transcripción incluyen, pero no se limitan a, aquellos que inhiben la formación de un complejo de iniciación de la transcripción, aquellos que disminuyen la tasa de iniciación de la transcripción, aquellos que disminuyen la tasa de transcripción de elongación, aquellos que disminuyen la capacidad de procesamiento de la transcripción y aquellos que antagonizan la activación de la transcripción (por ejemplo, al bloquear la unión de un activador transcripcional). La represión génica puede constituir, por ejemplo, la prevención de la activación, así como la inhibición de la expresión por debajo de un nivel existente. Ejemplos de procesos de represión génica que disminuyen la traducción incluyen aquellos que disminuyen la iniciación de la traducción, aquellos que disminuyen la elongación de traslación y aquellos que disminuyen la estabilidad del ARNm. La represión transcripcional incluye la inactivación tanto reversible como irreversible de la transcripción génica. En general, la represión génica comprende cualquier disminución detectable en la producción de un producto génico, preferiblemente una disminución en la producción de un producto génico, preferiblemente una disminución de la mitad en la producción de un producto génico, más preferiblemente aproximadamente de 2 a aproximadamente 5 veces o cualquier valor entero entre los mismos, más preferiblemente de aproximadamente 5 a aproximadamente 10 veces o cualquier valor entero entre los mismos, más preferiblemente entre aproximadamente 10 y aproximadamente 20 veces o cualquier valor entero entre las mismas, aún más preferiblemente entre aproximadamente 20 y aproximadamente 50 veces o cualquier valor entero entre ellas, más preferiblemente entre aproximadamente 50 - y aproximadamente 100 veces o cualquier valor entero entre los mismos, más preferiblemente 100 veces o más. Más preferiblemente, la represión génica tiene como resultado la completa inhibición de la expresión génica, de tal manera que no se detecta ningún producto génico.

El término "modular" se refiere a un cambio en la cantidad, grado o extensión de una función. Por ejemplo, polipéptidos de unión nucleótido-dedo de zinc modificados descritos en este documento pueden modular la actividad de una secuencia de un promotor mediante la unión a un motivo dentro del promotor, induciendo de ese modo, el incremento o la supresión de la transcripción de un gen ligado operativamente a la secuencia promotora. Alternativamente, la modulación puede incluir la inhibición de la transcripción de un gen donde el polipéptido de unión nucleótido-dedo de zinc modificado se une al gen estructural y bloquea la lectura de la ARN polimerasa dependiente de ADN a lo largo del gen, lo que inhibe la transcripción del gen. El gen estructural puede ser un gen celular normal o un oncogén, por ejemplo. Alternativamente, la modulación puede incluir la inhibición de la traducción de una transcripción. Por lo tanto, la "modulación" de la expresión génica incluye tanto la activación génica y la represión de genes.

La modulación puede ensayarse por determinación de cualquier parámetro que este indirectamente o directamente afectado por la expresión del gen diana. Tales parámetros incluyen, por ejemplo, cambios en los niveles de proteína o de ARN; cambios en la actividad de la proteína, cambios en los niveles de los productos; cambios en la expresión génica aguas abajo; cambios en la transcripción o actividad de genes indicadores, tales como, por ejemplo, luciferasa, CAT, beta-galactosidasa, o GFP (véase, por ejemplo., Mistili & Spector, (1997) *Nature Biotechnology* 15:961-964); cambios en la transducción de la señal; cambios en la fosforilación y defosforilación; cambios en las interacciones receptor-ligando; cambios en las concentraciones de segundos mensajeros, tales como, por ejemplo, cGMP, cAMP, IP₃, y Ca²⁺; cambios en el crecimiento celular; cambios en la neovascularización, y/o cambios en cualquier efecto funcional de la expresión génica. Las mediciones pueden realizarse *in vitro*, *in vivo* y/o *ex vivo*. Tales efectos funcionales se pueden medir por métodos convencionales, por ejemplo, medición de los niveles de proteína o de ARN, medición de la estabilidad del ARN, y/o identificación aguas abajo o expresión del gen indicador. La lectura puede ser por medio de, por ejemplo, quimioluminiscencia, fluorescencia, reacciones colorimétricas, unión de anticuerpos, marcadores inducibles, ensayos de unión a ligando; cambios en segundos mensajeros intracelulares tales como cGMP e inositol trifosfato (IP₃); cambios en los niveles de calcio intracelular; liberación de citoquinas, y similares.

Las "células eucariotas" incluyen, pero no se limitan a, células de hongos (tales como levaduras), células vegetales, células animales, células de mamíferos y células humanas. De manera similar, "las células procariotas incluyen, pero no se limitan a, bacterias.

Un "dominio regulador" o "dominio funcional" se refiere a una proteína o una secuencia de polipéptidos que tiene actividad moduladora de la transcripción, o que es capaz de interactuar con proteínas y/o con dominios de proteínas

que tienen actividad de moduladora de la transcripción. Típicamente, un dominio funcional está unido covalente o no covalentemente a una ZFP para modular la transcripción de un gen de interés. Alternativamente, una ZFP puede actuar, en ausencia de un dominio funcional, para modular la transcripción. Además, la transcripción de un gen de interés puede ser modulada por una ZFP ligada a múltiples dominios funcionales.

Un "fragmento funcional" de una proteína, polipéptido o ácido nucleico es una proteína, polipéptido o ácido nucleico cuya secuencia no es idéntica a la proteína de longitud completa, al polipéptido o al ácido nucleico, sin embargo, conserva la misma función que la proteína de longitud completa, polipéptido o ácido nucleico. Un fragmento funcional puede poseer más, menos, o el mismo número de restos que la molécula nativa correspondiente, y/o puede contener uno o más sustituciones de aminoácidos o de nucleótidos. Los métodos para determinar la función de un ácido nucleico (por ejemplo, la función de codificación, capacidad de hibridar con otro ácido nucleico) son bien conocidas en la técnica. Del mismo modo, los métodos para determinar la función de la proteína son bien conocidos. Por ejemplo, la función de unión al ADN de un polipéptido se puede determinar, por ejemplo, mediante unión a filtros, cambios de movilidad electroforética, o ensayos de inmunoprecipitación. Véase Ausubel *et al.*, *supra*. La capacidad de una proteína para interactuar con otra proteína se puede determinar, por ejemplo, por co-inmunoprecipitación, ensayos de doble híbrido o complementación, tanto genéticos y bioquímicos. Véase, por ejemplo, Fields *et al.* (1989) *Nature* 340:245-246; EE.UU nº 5.585.245 y el documento PCT WO 98/44350.

Un "sitio diana" o "secuencia diana" es una secuencia que está unida por una proteína de unión tal como, por ejemplo, una ZFP. Las secuencias diana pueden ser secuencias de nucleótidos (bien ADN o ARN) o secuencias de aminoácidos. A modo de ejemplo, una secuencia de ADN diana para una ZFP de tres dedos tiene generalmente una longitud de 9 ó 10 nucleótidos, dependiendo de la presencia y/o naturaleza de las interacciones de cadena cruzadas entre la ZFP y la secuencia diana. Se pueden encontrar secuencias diana en cualquier secuencia de ADN o ARN, incluyendo secuencias reguladoras, exones, intrones, o cualquier secuencia no codificante.

Un "subsitio diana" o "subsitio" es la parte de un sitio diana de ADN que se une mediante un solo dedo de zinc, excluyendo las interacciones de cadenas cruzadas. Así, en ausencia de interacciones de cadenas cruzadas, un subsitio tiene generalmente de tres nucleótidos de longitud. En los casos en los que se produce una interacción de cadenas cruzadas (por ejemplo, un "subsitio D", tal como se describe por ejemplo en la solicitud de propiedad conjunta PCT WO 00/42219, que se incorpora por referencia en su totalidad en la presente memoria) un subsitio tiene cuatro nucleótidos de longitud y se solapa con otro subsitio de 3 o 4 nucleótidos.

El término "cantidad eficaz" incluye la cantidad que da lugar al resultado deseado, por ejemplo, la desactivación de un gen previamente activado, la activación de un gen previamente reprimido, o la inhibición de la transcripción de un gen estructural o la traducción de ARN.

Proteínas con dedos de zinc

Las proteínas con dedos de zinc se forman a partir de los componentes de los dedos de zinc. Por ejemplo, las proteínas con dedos de zinc puede tener desde uno hasta treinta y siete dedos, normalmente tienen 2, 3, 4, 5 o 6 dedos. Las proteínas con dedos de zinc que se unen a ADN se describen, por ejemplo, en Miller *et al.* (1985) *EMBO J.* 4:1609-1614; Rhodes *et al.* (1993) *Scientific American* Feb.:56-65, y Klug (1999) *J. Mol. Biol.* **293**:215-218. Una proteína con dedos de zinc reconoce y se une a un sitio diana (a veces referido como un segmento diana) que representa una subsecuencia relativamente pequeña dentro de un gen diana. Cada dedo componente de una proteína con dedos de zinc se une a un subsitio dentro de la zona diana. El subsitio incluye un triplete de tres bases contiguas en la misma cadena (a veces referida como cadena diana). Las tres bases del subsitio se pueden denominar individualmente la base 5', base intermedia, y base 3' del triplete, respectivamente. El subsitio puede o no puede incluir también una cuarta base en la cadena no diana que es complementaria a la base al lado de 3' de las tres bases contiguas en la cadena diana. La base al lado de 3' de las tres bases contiguas en la cadena diana se refiere a veces como la 3' de la base de 3'. Alternativamente, las cuatro bases de la cadena diana en un subsitio de cuatro bases pueden ser numeradas 4, 3, 2, y 1, respectivamente, a partir de la base 5'.

Al analizar las regiones determinantes de la especificidad de un dedo de zinc, ácido amino + 1 se refiere al primer aminoácido en la porción α -helicoidal del dedo de zinc. Se cree que la porción de un dedo de zinc que es generalmente responsable de su especificidad de unión se encuentra entre -1 y +6. Aminoácido ++2 se refiere al aminoácido en la posición +2 de un segundo dedo de zinc adyacente (en la dirección C-terminal) del dedo de zinc en estudio. En determinadas circunstancias, un dedo de zinc se une a su subsitio del triplete de forma sustancialmente de manera independiente a los otros dedos de la misma proteína con dedos de zinc. En consecuencia, la especificidad de unión de una proteína con dedos de zinc que contiene varios dedos, en una primera aproximación, es la suma de las especificidades de los dedos que la componen. Por ejemplo, si una proteína con dedos de zinc se forma con un primer, segundo y tercer dedo que se unen individualmente a los tripletes XXX, YYY y ZZZ, la especificidad de unión de la proteína con dedos de zinc es 3'-XXX YYY ZZZ-5'.

El orden relativo de los dedos en una proteína con dedos de zinc, desde el N-terminal al C-terminal, determina el orden relativo de los tripletes en la secuencia diana, en el sentido 3' a 5' que será reconocido por los dedos. Por ejemplo, si una proteína con dedos de zinc comprende, a partir de N-terminal hacia C-terminal, los dedos primero, segundo y tercero que se unen individualmente a los tripletes 5'-GAC-3', 5'-GTA-3' y 5'-GGC-3', respectivamente,

entonces la proteína con dedos de zinc se une a la secuencia diana 5'-GGCGTAGAC-3' (SEQ ID NO: 3). Si la proteína con dedos de zinc comprende los dedos en otro orden, por ejemplo, segundo dedo, dedo primero, tercer dedo, entonces la proteína con dedos de zinc se une a un segmento diana que comprende una permutación diferente de tripletes, en este ejemplo, 5'-GGCGACGTA-3' (SEQ ID NO: 4). Véase Berg *et al.* (1996) *Science* 271:1081-1086.

Un dedo componente de una proteína con dedos de zinc contiene normalmente aproximadamente 30 aminoácidos y comprende la siguiente secuencia consenso canónica (de N a C):

Cys-(X)₂₋₄-Cys-X₁₂-His-(X)₃₋₅His (SEQ ID NO: 2)

Por lo tanto, la mayoría de los dedos de zinc tipo C2H2 contienen dos restos cisteína invariantes en el giro beta y dos restos de histidina invariantes, estos cuatro restos están coordinados a través de un átomo de zinc para mantener la estructura de dedo de zinc característica. Véase, por ejemplo, Berg y Shi (1996) *Science* 271:1081-1085. El acuerdo de numeración utilizado anteriormente es estándar en el sector técnico para la región de dedo de zinc que confiere especificidad de unión. Al aminoácido junto al N-terminal del primer resto His invariante se le asigna el número +6, y a los otros aminoácidos, procediendo en un sentido N-terminal, se les asignan sucesivamente números decrecientes. La hélice alfa empieza en el resto +1 y se extiende hasta el resto que sigue a la segunda histidina conservada. La hélice completa tiene por tanto una longitud variable, entre 11 y 13 restos.

Determinados dominios de unión al ADN son capaces de unirse al ADN que se empaqueta en los nucleosomas. Véase, por ejemplo, Cordingley *et al.* (1987) *Cell* **48**:261-270; Pina *et al.* (1990) *Cell* **60**:719-731; y Cirillo *et al.* (1998) *EMBO J.* **17**:244-254. Determinadas proteínas que contienen ZFP tales como, por ejemplo, miembros de la superfamilia de receptores hormonales nucleares, son capaces de unirse a secuencias de ADN empaquetado en la cromatina. Estos incluyen, pero no se limitan, al receptor de glucocorticoides y al receptor de hormona tiroidea. Archer *et al.* (1992) *Science* **255**:1573-1576; Wong *et al.* (1997) *EMBO J.* **16**:7130-7145. Otros dominios de unión al ADN, que incluyen determinados dominios de unión que contienen ZFP, requieren ADN más accesible para la unión. En este último caso, la especificidad de unión del dominio de unión al ADN necesaria se puede determinar mediante la identificación de regiones accesibles en la cromatina celular. Las regiones accesibles se pueden determinar cómo se describe en las Publicaciones Internacionales de propiedad conjunta WO 01/83751 y WO 01/83732, cuyas descripciones se incorporan aquí por referencia. Un dominio de unión a ADN de una ZFP modificada está diseñado y/o seleccionado para unirse a un sitio diana dentro de la región accesible.

A. ZFP no-canónicas

Las composiciones y métodos descritos en la presente memoria incluyen proteínas con dedos de zinc modificadas, preferiblemente no canónicas (por ejemplo, no C2H2), que se unen específicamente a una secuencia diana. Se pueden diseñar y/o seleccionar dominios de unión ADN-ZFP no canónicos para reconocer un sitio diana particular, por ejemplo como se describe en las solicitudes de propiedad conjunta WO 00/42219, WO 00/41566, así como en las patentes de EE.UU. nº 5.789.538; 6.007.408; 6.013.453; 6.140.081 y 6.140.466, y las publicaciones PCT WO 95/19431, WO 98/54311, WO 00/23464 y WO 00/27878. En las realizaciones preferidas, el proceso de diseño o selección de una ZFP no canónica de origen no natural empieza normalmente con una ZFP natural como fuente de restos marco, como se describe en las solicitudes de propiedad conjunta PCT WO 00/42219, WO 98/53057; WO 98/53058, WO 98/53059 y WO 98/53060.

En resumen, los métodos descritos en este documento sirven para modificar los restos Cys e His típicamente invariantes manteniendo al mismo tiempo (o aumentando) la especificidad de unión deseada de una ZFP. El proceso de obtención de una ZFP de origen no natural con una especificidad de unión predeterminada comienza normalmente con una ZFP natural como fuente de restos marco. El proceso de diseño o selección sirve para definir posiciones no conservadas (es decir, las posiciones -1 a +6) de modo que confieran una especificidad de unión deseada. Una ZFP adecuada para uso como marco es el dominio de unión al ADN del factor de transcripción Zif268 de ratón. Otra proteína natural con dedos de zinc adecuada como fuente de restos marco es Sp-1. La secuencia SP-1 utilizada para la construcción de proteínas con dedos de zinc corresponde a los aminoácidos 531 a 624 del factor de transcripción Sp-1. Un esqueleto adicional de ZFP útil es el de la secuencia consenso de Sp-1, descrita por Shi *et al.* (1995) *Chemistry and Biology* 1:83-89. Las secuencias de aminoácidos de estos marcos de ZFP se describen en PCT WO 00/42219 de propiedad conjunta, cuya descripción se incorpora por referencia. En otros aspectos, el esqueleto de ZFP comprenderá una esqueleto ZFP de plantas modificado en el que uno o más de los no canónicos descritos en la presente memoria se insertan de manera que se unan a una secuencia diana. Otras ZFP adecuadas son conocidos por los expertos en la técnica y se describen en la presente memoria. Los documentos citados arriba describen también métodos para evaluar la especificidad de unión de ZFP modificadas.

Las proteínas con dedos de zinc no canónicas por lo tanto, incluyen uno o más componentes con dedos de zinc en donde al menos uno de los aminoácidos de C2H2 se ha sustituido por uno o más aminoácidos. En la presente memoria se describe que más de uno de los aminoácidos canónicos puede ser sustituido. Ejemplos de componentes de dedos de zinc no-canónicos incluyen:

$X_3-B-X_{2.4}-Cys-X_{12}-His-X_{1.7}-His-X_4$
 $X_3-Cys-X_{2.4}-B-X_{12}-His-X_{1.7}-His-X_4$
 $X_3-Cys-X_{2.4}-Cys-X_{12}-Z-X_{1.7}-His-X_4$
 $X_3-Cys-X_{2.4}-Cys-X_{12}-His-X_{1.7}-Z-X_4$
 $X_3-B-X_{2.4}-B-X_{12}-His-X_{1.7}-His-X_4$
 $X_3-B-X_{2.4}-Cys-X_{12}-Z-X_{1.7}-His-X_4$
 $X_3-B-X_{2.4}-Cys-X_{12}-His-X_{1.7}-Z-X_4$
 $X_3-Cys-X_{2.4}-B-X_{12}-Z-X_{1.7}-His-X_4$
 $X_3-Cys-X_{2.4}-B-X_{12}-His-X_{1.7}-Z-X_4$
 $X_3-Cys-X_{2.4}-Cys-X_{12}-Z-X_{1.7}-Z-X_4$
 $X_3-Cys-X_{2.4}-B-X_{12}-Z-X_{1.7}-Z-X_4$
 $X_3-B-X_{2.4}-Cys-X_{12}-Z-X_{1.7}-Z-X_4$
 $X_3-B-X_{2.4}-B-X_{12}-His-X_{1.7}-Z-X_4$
 $X_3-B-X_{2.4}-B-X_{12}-Z-X_{1.7}-His-X_4$
 $X_3-B-X_{2.4}-B-X_{12}-Z-X_{1.7}-Z-X_4$
 $X_3-Y-X_{2.4}-Cys-X_{12}-His-X_{1.7}-His-X_4$
 $X_3-Cys-X_{2.4}-Y-X_{12}-His-X_{1.7}-His-X_4$
 $X_3-Cys-X_{2.4}-Cys-X_{12}-Y-X_{1.7}-His-X_4$
 $X_3-Cys-X_{2.4}-Cys-X_{12}-His-X_{1.7}-Y-X_4$
 $X_3-Y-X_{2.4}-Y-X_{12}-His-X_{1.7}-His-X_4$
 $X_3-Y-X_{2.4}-Cys-X_{12}-Y-X_{1.7}-His-X_4$
 $X_3-Y-X_{2.4}-Cys-X_{12}-His-X_{1.7}-Y-X_4$
 $X_3-Cys-X_{2.4}-Y-X_{12}-Y-X_{1.7}-His-X_4$
 $X_3-Cys-X_{2.4}-Y-X_{12}-His-X_{1.7}-Y-X_4$
 $X_3-Cys-X_{2.4}-Cys-X_{12}-Y-X_{1.7}-Y-X_4$
 $X_3-Cys-X_{2.4}-Y-X_{12}-Y-X_{1.7}-Y-X_4$
 $X_3-Y-X_{2.4}-Cys-X_{12}-Y-X_{1.7}-Y-X_4$
 $X_3-Y-X_{2.4}-Y-X_{12}-His-X_{1.7}-Y-X_4$
 $X_3-Y-X_{2.4}-Y-X_{12}-Y-X_{1.7}-His-X_4$
 $X_3-Y-X_{2.4}-Y-X_{12}-Y-X_{1.7}-Y-X_4$

donde X = cualquier aminoácido

B = cualquier aminoácido excepto cisteína

Z = cualquier aminoácido excepto histidina

5 Y = cualquier aminoácido excepto cisteína o histidina

Una ZFP modificada puede estar compuesta por un número cualquiera de dedos de zinc, aunque una estructura de tres dedos es generalmente preferida. Normalmente, el dedo más C-terminal (por ejemplo, el tercero) de la ZFP está modificado y es no canónico. Los otros dedos de los que se compone la proteína pueden ser dedos de zinc naturales, no canónicos modificados, dedos C2H2 modificados o combinaciones de éstos. Por lo tanto, como se describe a continuación en el Ejemplo 2, en determinadas realizaciones, se proporciona una proteína de unión con 10 dedos de zinc de tres dedos en la que los dos primeros dedos son de la clase C2-H2 y, en el tercer dedo (el más C-terminal), la segunda histidina está sustituida por Cys o por Cys y otros aminoácidos, tales como la glicina. En otras realizaciones, se proporciona una proteína con dedos de zinc de tres dedos en la que, en el primer dedo (el más N-terminal), el primer resto de cisteína está sustituido por histidina o por histidina y otros aminoácidos tales como glicina. Además, en determinadas realizaciones, el segundo dedo (medio) de una ZFP de tres dedos se modifica de 15 modo que una o ambas de las cisteínas se reemplazan por serinas (y/u otros aminoácidos).

También se incluyen en este documento ácidos nucleicos que codifican una ZFP que comprende al menos un dedo de zinc no canónico como se describe en la presente memoria.

B. Asociación

20 Dos o más proteínas con dedos de zinc se pueden asociar para tener una especificidad para un sitio diana que sea, en una primera aproximación, la suma de las proteínas de los dedos de zinc que las componen. Por ejemplo, una

primera proteína con dedos de zinc compuesta por los dedos primero, segundo y tercero, respectivamente, que se unen a XXX, YYY y ZZZ puede asociarse con una segunda proteína con dedos de zinc compuesta por los dedos primero, segundo, y tercero con las especificidades de unión, AAA, BBB y CCC. La especificidad de unión de las proteínas primera y segunda combinadas es por lo tanto, 5'-CCCBBAANZZZYYXXX-3', donde N indica una región intermedia corta (típicamente de 0-5 bases de cualquier tipo). En esta situación, se puede ver que el sitio diana comprende dos segmentos diana separados por un segmento intermedio.

La asociación de las proteínas con dedos de zinc puede realizarse usando cualquiera de los siguientes enlazadores peptídicos:

TGEKP (SEQ ID NO: 5) Liu *et al.* (1997) *Proc.Natl.Acad.Sci.EE.UU* **94**:5525-5530.

(G₄S)_n (SEQ ID NO: 6) Kim *et al.* (1996) *Proc. Natl. Acad.Sci.EE.UU.* **93**:1156-1160.

GGRRGGGS (SEQ ID NO: 7)

LRQRDGERP (SEQ ID NO: 8)

LRQKDGGGSERP (SEQ ID NO: 9)

LRQKD(G₃S)₂ERP (SEQ ID NO: 10)

Alternativamente, se pueden diseñar enlazadores flexibles usando programas de ordenador capaces de modelizar tanto los sitios de unión al ADN y como los propios péptidos, o por métodos de presentación de fagos. En una variación adicional, se puede lograr asociaciones no covalentes mediante la fusión de dos proteínas con dedos de zinc con dominios que promuevan la formación del heterodímero de dos proteínas con dedos de zinc. Por ejemplo, una proteína con dedos de zinc se puede fusionar con *fos* y la otra con *jun* (ver Barbas *et al.*, WO 95/119431).

Alternativamente, las interfaces de dimerización se pueden obtener por selección. Véase, por ejemplo, Wang *et al.*(1999) *Proc. Natl. Acad. Sci. EE.UU.* **96**:9568-9573.

La asociación de dos o más proteínas con dedos de zinc es ventajosa para conferir una especificidad de unión única dentro de un genoma de mamífero. Un genoma diploide típico de mamífero consta de 3×10^9 pb. Suponiendo que los cuatro nucleótidos A, C, G y T están distribuidos al azar, una determinada secuencia de 9 pb está presente 23.000 veces. Así pues, una ZFP con tres dedos que reconoce una diana de 9 pb con absoluta especificidad podría tener potencial para unirse a 23.000 sitios en el genoma. Una secuencia de 18 pb está presente en $3,4 \times 10^{10}$ pb una vez, o una vez aproximadamente en una secuencia aleatoria de ADN cuya complejidad sea diez veces mayor que la de un genoma de mamífero. Por lo tanto, la unión dos ZFP con tres dedos, para reconocer una secuencia diana de 18 pb, proporciona la especificidad requerida para dirigirse a un sitio único en un genoma de mamífero típico.

C. Moléculas de fusión

La selección y/o el diseño de proteínas que contienen dedos de zinc no canónicos permite también el diseño de moléculas de fusión que facilitan la regulación de la expresión génica. Así, en ciertas realizaciones, las composiciones y métodos descritos en la presente memoria implican fusiones entre al menos una de las proteínas con dedos de zinc que se describen en la presente memoria (o fragmentos funcionales de las mismas) y uno o más dominios funcionales (o fragmentos funcionales de los mismos), o un polinucleótido que codifica tal fusión. La presencia de una molécula de fusión en una célula permite que un dominio funcional se aproxime a una secuencia de un gen que este unida por la porción del dedo de zinc de la molécula de fusión. La función reguladora de la transcripción del dominio funcional es entonces capaz de actuar sobre el gen, mediante, por ejemplo, modulación de la expresión génica.

En determinadas realizaciones, las proteínas de fusión que comprenden un dominio de unión a ADN de un dedo de zinc modificado y un dominio funcional se utilizan para la modulación de la expresión del gen endógeno como se describe, por ejemplo en la solicitud de propiedad conjunta, PCT WO 00/41566. La modulación incluye la represión y la activación de la expresión génica; la naturaleza de la modulación generalmente depende del tipo de dominio funcional presente en la proteína de fusión. Cualquier secuencia polipeptídica o dominio capaz de influir en la expresión génica (o su fragmento funcional) que pueda estar fusionada a un dominio de unión a ADN, es adecuada para su uso.

Un ejemplo de dominio funcional para la fusión con un dominio de unión a ADN de una ZFP, que se utiliza para reprimir la expresión génica, es un dominio de represión KRAB de la proteína KOX-1 humana de (véase, por ejemplo, Thiesen *et al.*, *New Biologist* **2**, 363-374 (1990); Margolin *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* **91**, 4509-4513 (1994); Pengue *et al.*, *Nucl. Acids. Res.* **22**:2908-2914 (1994); Witzgall *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* **91**, 4.514-4518 (1994). Otro dominio de represión adecuado es el dominio de unión a metilo de la proteína 2B (MBD-2B) (véase, también Hendrich *et al.* (1999) *Mamm Genome* **10**:906-912 para la descripción de las proteínas MBD). Otro dominio de represión útil es el que está asociado con la proteína v-ErbA. Véase, por ejemplo, Damm, *et al.* (1989) *Nature* **339**:593-597; Evans (1989) *Int. J. Cancer Suppl.* **4**:26-28; Pain *et al.* (1990) *New Biol.* **2**:284-294; Sap *et al.*

(1989) *Nature* **340**:242-244; Zenke *et al.* (1988) *Cell* **52**:107-119; y Zenke *et al.* (1990) *Cell* **61**:1035-1049. Otros ejemplos de dominios de represión incluyen, pero no se limitan a, receptores de hormona tiroidea (TR), SID, MBD1, MBD2, MBD3, MBD4, proteínas similares a MBD, miembros de la familia DNMT (por ejemplo, DNMT1, DNMT3A, DNMT3B), Rb, MeCP1 y MeCP2. Véase, por ejemplo, Zhang *et al.* (2000) *Ann Rev Physiol* **62**:439-466; Bird *et al.* (1999) *Cell* **99**:451-454; Tyler *et al.* (1999) *Cell* **99**:443-446; Knoepfler *et al.* (1999) *Cell* **99**:447-450, y Robertson *et al.* (2000) *Nature Genet.* **25**:338-342. Otros ejemplos de dominios de represión incluyen, pero no se limitan a, ROM2 y AtHD2A. Véase, por ejemplo, Chem. *et al.* (1996) *Plant Cell* **8**:305-321, y Wu *et al.* (2000) *Plant J.* **22**:19-27.

Los dominios adecuados para lograr la activación incluyen el dominio de activación de VP16 de HSV (véase, por ejemplo, Hagmann *et al.*, *J. Virol.* **71**, 5952-5962 (1997)) receptores nucleares de hormonas (véase, por ejemplo, Torchia *et al.*, *Curr. Opin. Cell. Biol.* **10**:373-383 (1998)), la subunidad p65 del factor nuclear kappa B (Bitko and Barik, *J. Virol* **72**:5610-5618 (1998.) y Doyle & Hunt, *Neuroreport* **8**:2937-2942 (1997)); Liu *et al.*, *Cancer Gene Ther.* **5**:3-28 (1998)), o dominios funcionales artificiales quiméricos, tales como VP64 (Seifpal *et al.*, *EMBO J.* **11**, 4961-4968 (1992)).

Otros ejemplos de dominios de activación incluyen, pero no se limitan a, p300, CBP, PCAF, PVALF SRC1, AtHD2A y ERF-2. Véase, por ejemplo, Robyr *et al.* (2000) *Mol. Endocrinol.* **14**:329-347; Collingwood *et al.* (1999) *J. Mol. Endocrinol.* **23**:255-275; Leo *et al.* (2000) *Gene* **245**:1-11; Manteuffel-Cymborowska (1999) *Acta Biochim. Pol.* **46**:77-89; McKenna *et al.* (1999) *J. Steroid Biochem. Mol. Biol.* **69**:3-12; Malik *et al.* (2000) *Trends Biochem. Sci.* **25**:277-283, y Lemon *et al.* (1999) *Curr. Opin. Genet. Dev.* **9**:499-504. Otros ejemplos de dominios de activación incluyen, pero no se limitan a, OsGAI, HALF-1, C1, AP1, ARF-5, -6, -7, y -8, CPRF1, CPRF4, MYC-RP/GP, y TRAB1. Véase, por ejemplo, Ogawa *et al.* (2000) *Gene* **245**:21-29; Okanami *et al.* (1996) *Genes Cells* **1**:87-99; Goff *et al.* (1991) *Genes Dev.* **5**:298-309; Cho *et al.* (1999) *Plant Mol. Biol.* **40**:419-429; Ulmason *et al.* (1999) *Proc. Natl. Acad. Sci. EE.UU.* **96**:5844-5849; Sprenger-Hausells *et al.* (2000) *Plant J.* **22**:1-8; Gong *et al.* (1999) *Plant Mol. Biol.* **41**:33-44, y Hobo *et al.* (1999) *Proc. Natl. Acad. Sci. EE.UU.* **96**:15,348-15,353.

Además se describen, dominios funcionales, por ejemplo, en la solicitud de propiedad conjunta WO 00/41566. Además se describen dominios aislante, proteínas de remodelación de la cromatina, tales como las que contienen dominios ISWI y/o dominios de proteínas de unión a metilo adecuados para uso en las moléculas de fusión, por ejemplo, en el documento WO 01/83793, WO 02/26959, WO 02/26960 y WO 02/44376.

En otras realizaciones, la remodelación dirigida de la cromatina, como se describe en la Publicación Internacional WO 01/83793 de propiedad conjunta se puede utilizar para generar uno o más sitios en la cromatina celular que sean accesibles para unirse a una molécula de fusión ZFP modificada/ dominio funcional.

Las moléculas de fusión se construyen mediante métodos de clonación y conjugación bioquímica que son bien conocidos por los expertos en la técnica. Las moléculas de fusión comprenden un dominio de unión de ZFP modificada y, por ejemplo, un dominio de activación de la transcripción, un dominio de represión de la transcripción, un componente de un complejo de remodelación de la cromatina, un dominio de aislante o un fragmento funcional de cualquiera de estos dominios. En determinadas realizaciones, las moléculas de fusión comprenden una proteína con dedos de zinc no canónicos y por lo menos dos dominios funcionales (por ejemplo, un dominio aislante o un dominio de proteínas de unión a metilo y, adicionalmente, un dominio de activación o de represión de la transcripción). Las moléculas de fusión también comprenden opcionalmente señales de localización nuclear (tales como, por ejemplo, aquellas a partir del antígeno-T del SV40) y etiquetas epítipo (tal como, por ejemplo, FLAG, véase el Ejemplo 2, y hemaglutinina). Las proteínas de fusión (y los ácidos nucleicos que las codifican) están diseñados de tal manera que el marco de lectura de la traducción está conservado entre los componentes de la fusión.

Las moléculas de fusión descritas en la presente memoria comprenden una proteína de unión con dedos de zinc no canónica que se une a un sitio diana. En determinadas realizaciones, el sitio diana está presente en una región accesible de la cromatina celular. Las regiones accesibles se pueden determinar cómo se describe en las Publicaciones Internacionales PCT WO 01/83751 y WO 01/83732 de cotitularidad. Si el sitio diana no está presente en una región accesible de la cromatina celular, se pueden generar una o más regiones accesibles como se describe, por ejemplo, en la Publicación Internacional PCT WO 01/83793 de propiedad conjunta. En otras formas de realización, el componente del dedo zinc no canónico de una molécula de fusión es capaz de unirse a la cromatina celular, independientemente de si su sitio diana está en una región accesible o no. Por ejemplo, tal y como se describe en la presente memoria una ZFP modificada puede ser capaz de unirse a ADN enlazador y/o a un ADN nucleosomal. Ejemplos de este tipo de ADN de dominio de unión "precursor" se encuentran en un determinado receptor de esteroides y en el factor nuclear de hepatocitos 3 (HNF3). Cordingley *et al.* (1987) *Cell* **48**:261-270; Pina *et al.* (1990) *Cell* **60**:719-731; y Cirillo *et al.* (1998) *EMBO J.* **17**:244-254.

Los métodos de regulación génica utilizando un dominio funcional, dirigido a una secuencia específica en virtud de un dominio de unión a ADN fusionado, pueden conseguir la modulación de la expresión génica. Los genes así modulados pueden ser genes endógenos o genes exógenos. La modulación de la expresión génica puede ser en la forma de represión (por ejemplo, reprimiendo la expresión de genes exógenos, por ejemplo, cuando el gen diana reside en un microorganismo infeccioso patológico, o por represión de un gen endógeno del sujeto, tal como un oncogén o un receptor viral, que contribuya a un estado de enfermedad). Como se describe en la presente memoria,

la represión de un gen diana específico se puede lograr mediante el uso de una molécula de fusión que comprende una proteína con dedos de zinc no canónica y un dominio funcional.

Alternativamente, la modulación puede ser en forma de activación, si la activación de un gen (por ejemplo, un gen supresor de tumor o un transgén) puede mejorar un estado de enfermedad. En este caso, la cromatina celular se pone en contacto con cualquiera de las moléculas de fusión descritas en el presente documento, donde la parte del dedo de zinc modificada de la molécula de fusión es específica para el gen diana. El dominio funcional (por ejemplo, dominio aislante, dominio de activación, etc.) permite la expresión incrementada y/o mantenida del gen diana.

Para cualquiera de estas aplicaciones, la(s) molécula(s) de fusión se puede formular con un vehículo farmacéuticamente aceptable, conocido por los expertos en la técnica. Véase, por ejemplo, *Remington' Pharmaceutical Sciences*, 7^o ed., 1985; y en la solicitud de propiedad conjunta WO 00/42219.

Introducción de polinucleótidos y polipéptidos

Las composiciones descritas en la presente memoria se puede proporcionar a la célula diana *in vitro* o *in vivo*. Además, las composiciones pueden proporcionarse como polipéptidos, polinucleótidos o combinación de los mismos.

A. Introducción de polinucleótidos

En determinadas realizaciones, las composiciones se proporcionan como uno o más polinucleótidos. Además, como se señaló anteriormente, una composición que contiene proteína con dedos de zinc no canónica se puede diseñar como una fusión entre un polipéptido dedo de zinc y un dominio funcional que este codificado por un ácido nucleico de fusión. En ambos casos de fusión y de no fusión, el ácido nucleico puede ser clonado en vectores intermedios para transformación en células procariotas o eucariotas para replicación y/o expresión. Los vectores intermedios para el almacenamiento o manipulación del ácido nucleico o para la producción de proteína pueden ser vectores procariotas, (por ejemplo, plásmidos), vectores lanzadera, vectores de insectos, o vectores virales, por ejemplo. Un ácido nucleico que codifica una proteína con dedos de zinc no canónica también puede clonarse en un vector de expresión, para administración a una célula bacteriana, a una célula fúngica, a una célula de protozoos, a una célula vegetal, o a una célula animal, preferiblemente a una célula de mamífero, más preferiblemente a una célula humana.

Para obtener la expresión de un ácido nucleico clonado, generalmente se subclona en un vector de expresión que contiene un promotor para la transcripción directa. Los promotores bacterianos y eucariotas adecuados son bien conocidos en la técnica y se describen, por ejemplo, en Sambrook et al, *arriba*; Ausubel et al, *arriba*, y Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990). Los sistemas de expresión bacterianos están disponibles en, por ejemplo, *E. coli*, *Bacillus sp.*, y *Salmonella*. Palva et al. (1983) *Gene* **22**:229-235. Los kits para tales sistemas de expresión están disponibles comercialmente. Los sistemas de expresión eucariotas para células de mamíferos, levaduras, y células de insectos son bien conocidos en la técnica y están también disponibles comercialmente, por ejemplo, de Invitrogen, Carlsbad, CA, y Clontech, Palo Alto, CA.

El promotor utilizado para dirigir la expresión del ácido nucleico de elección depende de la aplicación particular. Por ejemplo, normalmente se utiliza un promotor constitutivo fuerte para la expresión y la purificación. En contraste, cuando una proteína se va a utilizar *in vivo*, se utiliza ya sea un promotor constitutivo o uno inducible, dependiendo del uso particular de la proteína. Además, se puede utilizar un promotor débil, como TK de HSV o un promotor que tenga una actividad similar. El promotor normalmente puede incluir también elementos que sean sensibles a la transactivación, por ejemplo, elementos de respuesta a hipoxia, elementos de respuesta Gal4, elementos de respuesta al represor *lac*, y sistemas de control de pequeñas moléculas tales como sistemas regulados por tet y el sistema RU-486. Véase, por ejemplo, Gossen et al. (1992) *Proc. Natl. Acad. Sci. EE.UU* **89**:5547-5551; Oligino et al (1998) *Gene Ther.* **5**:491-496; Wang et al. (1997) *Gene Ther.* **4**:432-441; Neering et al. (1996) *Blood* **88**:1147-1155; y Rendahl et al. (1998) *Nat. Biotechnol.* **16**:757-761.

Además de un promotor, un vector de expresión contiene normalmente una unidad de transcripción o casete de expresión que contenga los elementos adicionales requeridos para la expresión del ácido nucleico en las células huésped, ya sea procariota o eucariota. Un casete de expresión típico contiene por lo tanto un promotor unido operativamente, por ejemplo, a la secuencia de ácido nucleico, y las señales requeridas, por ejemplo, para la poliadenilación eficiente del transcrito, para la terminación de la transcripción, para la unión a ribosomas y/o para la terminación de la traducción. Otros elementos del casete pueden incluir, por ejemplo, potenciadores, y señales intrónicas de empalme heterólogas.

El vector de expresión específico usado para transportar la información genética dentro de la célula se selecciona en relación a la utilización prevista del polipéptido de ZFP resultante, por ejemplo, la expresión en plantas, animales, bacterias, hongos, protozoos, etc. Los vectores de expresión bacterianos estándar incluyen plásmidos tales como pBR322, plásmidos basados en pBR322, pSKF, pET23D, y sistemas de expresión de fusión disponibles comercialmente tales como GST y LacZ. Las etiquetas del epítipo también se pueden añadir a las proteínas recombinantes para proporcionar los métodos de aislamiento convenientes, para el seguimiento de la expresión, y para el seguimiento de la localización celular y subcelular, por ejemplo, c-myc o FLAG.

Los vectores de expresión que contienen elementos reguladores de virus eucariotas se utilizan a menudo en vectores de expresión eucariotas, por ejemplo, vectores de SV40, vectores del virus del papiloma, y vectores derivados del virus de Epstein-Barr. Otros ejemplos de vectores eucariotas incluyen pMSG, pAV009/A+, pMTO10/A+, pMAMneo-5, baculovirus de pDSVE, y cualquier otro vector que permita la expresión de proteínas bajo la dirección del promotor temprano de SV40, del promotor SV40 tardío, del promotor de metalotioneína, del promotor del virus del tumor mamario murino, del promotor de virus del sarcoma de Rous, del promotor de polihedrina, o de otros que se muestren eficaces para la expresión en células eucariotas.

Algunos sistemas de expresión tienen marcadores para la selección de líneas celulares transfectadas de manera estable tales como timidina quinasa, fosfotransferasa de higromicina B, y dihidrofolato reductasa. Los sistemas de expresión de alto rendimiento también son adecuados, tales como los vectores de baculovirus en células de insectos, con una secuencia de ácido nucleico que codifica una ZFP como se describe en la presente memoria bajo el control transcripcional del promotor de polihedrina o cualquier otro promotor de baculovirus fuerte.

Los elementos que se incluyen normalmente en los vectores de expresión también incluyen un replicón que funciona en *E. coli* (o en el huésped procarionota, distinto de *E. coli*), un marcador selectivo, por ejemplo, un gen que codifica para resistencia a antibióticos, que permita la selección de bacterias que albergan plásmidos recombinantes, y sitios de restricción únicos en regiones no esenciales del vector que permitan la inserción de secuencias recombinantes.

Los métodos de transfección estándar se puede utilizar para producir líneas celulares de bacterias, de mamíferos, de insectos, o otras que expresen grandes cantidades de proteínas con dedos de zinc no canónicas, que se pueden purificar, si se desea, usando técnicas estándar. Véase, por ejemplo, Colley *et al.* (1989) *J. Biol. Chem.* **264**:17619-17622, y *Guide to Protein Purification, in Methods in Enzymology*, vol.182 (Deutscher, ed.)1990.

La transformación de células eucariotas y procariotas se realiza según técnicas estándar. Véase, por ejemplo, Morrison (1977) *J. Bacteriol.* **132**:349-351; Clark-Curtiss *et al.* (1983) en *Methods in Enzymology* **101**:347-362 (Wu *et al.*, Eds).

Se puede utilizar cualquier procedimiento para la introducción de secuencias de nucleótidos extraños en las células huésped. Estos incluyen, pero no se limitan a, el uso de transfección con fosfato cálcico, transfección mediada por DEAE-dextrano, polibreno, fusión de protoplastos, electroporación, introducción por medio de lípidos (por ejemplo, liposomas), microinyección, bombardeo de partículas, introducción de ADN desnudo, vectores plásmidos, vectores virales (tanto episomal como integrador) y cualquiera de los otros métodos bien conocidos para introducir ADN genómico clonado, ADNc, ADN sintético u otro material genético extraño en una célula huésped (véase, por ejemplo, Sambrook *et al.*, arriba). Sólo es necesario que el procedimiento particular de ingeniería genética utilizado sea capaz de introducir con éxito al menos un gen en la célula huésped capaz de expresar la proteína de elección.

Se pueden utilizar métodos de transferencia de genes convencionales virales y no virales para introducir ácidos nucleicos en células de mamíferos o en tejidos diana. Estos métodos se pueden utilizar para administrar ácidos nucleicos que codifican polipéptidos de reprogramación de células *in vitro*. Preferiblemente, se administran los ácidos nucleicos para usos de terapia génica *in vivo* o *ex vivo*. Los sistemas de introducción de vectores virales incluyen plásmidos de ADN, ácido nucleico desnudo, y ácido nucleico complejado con un vehículo de introducción tal como un liposoma. Los sistemas de introducción de vectores virales incluyen virus de ADN y ARN, que tienen genomas ya sea episomales o integrados después de ser introducidos en la célula. Para revisiones de procedimientos de terapia génica, véase, por ejemplo, Anderson (1992) *Science* **256**:808-813; Nabel *et al.* (1993) *Trends Biotechnol.* **11**:211-217; Mitani *et al.* (1993) *Trends Biotechnol.* **11**:162-166; Dillon (1993) *Trends Biotechnol.* **11**:167-175; Miller (1992) *Nature* **357**:455-460; Van Brunt (1988) *Biotechnology* **6**(10):1149-1154, Vigne (1995) *Restorative Neurology and Neuroscience* **8**:35-36; Kremer *et al.* (1995) *British Medical Bulletin* **51** (1) :31-44; Haddada *et al.*, en *Current Topics in Microbiology and Immunology*, Doerfler and Bohm (eds.), 1995, y Yu *et al.* (1994) *Gene Therapy* **1**:13-26.

Los métodos de introducción no víricos de ácidos nucleicos incluyen lipofección, microinyección, balística, virosomas, liposomas, inmunoliposomas, policonjugados o conjugados de lípido:ácidos nucleicos, ADN desnudo, viriones artificiales, y agente potenciador de la captación de ADN. La lipofección se describe, por ejemplo, en EE.UU. Pat. N° 5.049.386; 4.946.787; y 4.897.355 y los reactivos de lipofección se venden comercialmente (por ejemplo, TransfectamTM y LipofectinaTM). Los lípidos catiónicos y neutros que son adecuados para la eficiente de reconocimiento de receptores de lipofección de polinucleótidos incluyen los de Felgner, WO91/17424 y WO91/16024. El ácido nucleico puede ser introducido en las células (administración *ex vivo*) o a tejidos diana (administración *in vivo*).

La preparación de los complejos lípido:ácido nucleico, incluyendo liposomas dirigidos tales como complejos de inmunolípidos, es bien conocido por los expertos en la técnica. Véase, por ejemplo, Crystal (1995) *Science* **270**:404-410; Blaese *et al.* (1995) *Cancer Gene Ther.* **2**:291-297; Behr *et al.* (1994) *Bioconjugate Chem.* **5**:382-389; Remy *et al.* (1994) *Bioconjugate Chem.* **5**:647-654; Gao *et al.* (1995) *Gene Therapy* **2**:710-722; Ahmad *et al.* (1992) *Cancer Res.* **52**:4817-4820, y las patentes de EE.UU n° 4.186.183; 4.217.344; 4.235.871; 4.261.975; 4.485.054; 4.501.728; 4.774.085; 4.837.028 y 4.946.787.

El uso sistemas basados en ARN o ADN vírico para la introducción de ácidos se aprovecha de los procesos altamente evolucionados para dirigir virus a células específicas en el cuerpo y el tráfico de carga viral al núcleo. Los vectores víricos pueden administrarse directamente a pacientes (*in vivo*) o pueden utilizarse para tratar células *in vitro*, donde las células modificadas se administran a los pacientes (*ex vivo*). Los sistemas convencionales basados en virus para suministro de ZFP incluyen vectores retrovirales, lentivirales, poxvirales, adenovirales, virales adeno-asociados, de virus de estomatitis vesicular y de herpesvirus. La integración en el genoma hospedador es posible con ciertos vectores virales, incluyendo métodos de transferencia génica de retrovirus, lentivirus, y virus adeno-asociados, teniendo como resultando frecuentemente la expresión a largo plazo del transgén insertado. Además, se ha observado alta eficiencia de la transducción en muchos tipos diferentes de células y tejidos diana.

El tropismo de un retrovirus puede alterarse mediante la incorporación de proteínas de la envoltura ajenas, permitiendo la modificación y/o expansión de la población potencial de células diana. Los vectores de Lentivirus son vectores retrovirales que son capaces de transducir o infectar células que no se dividen y producen normalmente títulos virales altos. La selección de un sistema de transferencia génica retroviral dependería por consiguiente del tejido diana. Los vectores retrovirales tienen una capacidad de empaquetamiento de hasta 6-10 kb de secuencias extrañas y se componen de repeticiones terminales largas (LTRs) que actúan en *cis*. El mínimo de LTRs que actúan en *cis* es suficiente para la replicación y el empaquetamiento de los vectores, que se utilizan entonces para integrar el gen terapéutico en la célula diana para proporcionar la expresión permanente del transgen. Los vectores retrovirales utilizados más ampliamente incluyen los basados en virus de la leucemia murina (MuLV), virus de leucemia de mono gibbon (GaLV), virus de inmunodeficiencia de simios (SIV), el virus de la inmunodeficiencia humana (VIH), y combinaciones de los mismos. Buchscher *et al.* (1992) *J. Virol.* **66**:2731-2739; Johann *et al.* (1992) *J. Virol.* **66**:1635-1640; Sommerfelt *et al.* (1990) *Virol.* **176**:58-59; Wilson *et al.* (1989) *J. Virol.* **63**:2374-2378; Miller *et al.* (1991) *J. Virol.* **65**:2220-2224, y PCT/US94/05700).

Los vectores de virus adeno-asociados (AAV) también se usan para transducir células con ácidos nucleicos diana, por ejemplo, en la producción *in vitro* de ácidos nucleicos y péptidos, y para procedimientos de terapia génica *ex vivo* e *in vivo*. Véase, por ejemplo, West *et al.* (1987) *Virology* **160**:38-47; EE.UU. Pat. N° 4.797.368, WO 93/24641; Kotin (1994) *Hum. Gene Ther.* **5**:793-801, y Muzyczka (1994) *J. Clin. Invest.* **94**:1351. La construcción de vectores de AAV recombinantes se describe en una serie de publicaciones, que incluyen EE.UU. Pat. N° 5.173.414; Tratschin *et al.* (1985) *Mol. Cell. Biol.* **5**:3251-3260; Tratschin, *et al.* (1984) *Mol. Cell. Biol.* **4**:2072-2081; Hermonat *et al.* (1984) *Proc. Natl. Acad. Sci.* EE.UU. **81**:6466-6470, y Samulski *et al.* (1989) *J. Virol.* **63**:3822-3828.

Los vectores recombinantes de virus adeno-asociados basados en el parvovirus adeno-asociado tipo 2 (AAV-2) defectuoso y no patógeno son un sistema de introducción de genes prometedor.

Los ejemplos de vectores de AAV se derivan de un plásmido que contiene repeticiones terminales invertidas de 145 pares de bases del AAV que flanquean un casete de expresión del transgén. La transferencia génica eficiente y la introducción estable del transgén debida a la integración en el genoma de la célula transducida son características clave para este sistema de vectores. Wagner *et al.* (1998) *Lancet* **351** (9117) :1702-3, y Kearns *et al.* (1996) *Gene Ther.* **9**:748-55.

pLASN y MFG-S son ejemplos de los vectores retrovirales que se han usado en ensayos clínicos.

Dunbar *et al.* (1995) *Blood* **85**:3048-305; Kohn *et al.* (1995) *Nature Med.* **1**:1017-102; Malech *et al.* (1997) *Proc. Natl. Acad. Sci. EE.UU.* **94**:12133-12138. PA317/pLASN fue el primer vector terapéutico utilizado en una prueba de terapia génica. Blaese *et al.* (1995) *Science* **270**:475-480. Se han observado una eficacia de transducción de 50% o mayor para vectores empaquetados MFG-S. Ellem *et al.* (1997) *Immunol Immunother.* **44** (1): 10-20; Dranoff *et al.* (1997) *Hum. Gene Ther.* **1**: 111-2.

En aplicaciones para las que se prefiere la expresión transitoria, son útiles los sistemas basados en adenovirus. Los vectores basados en adenovirus tienen la capacidad de eficacia de transducción muy alta en muchos tipos de células y son capaces de infectar, y por lo tanto de introducir ácidos nucleicos, tanto en células que se dividen y como en las que no se dividen. Con estos vectores, se han obtenido altos títulos y niveles de expresión. Los vectores de adenovirus pueden ser producidos en grandes cantidades en un sistema relativamente simple.

Los vectores de adenovirus (Ad) recombinantes de replicación deficiente se pueden producir a títulos elevados e infectan fácilmente varios tipos de células diferentes. La mayoría de los vectores de adenovirus se han diseñado de tal manera que un transgén reemplaza a los genes Ad E1a, E1b, y/o E3; el vector que produce el defecto de replicación se propaga en las células 293 humanas que proporcionan las funciones E1 requeridas en *trans*. Los vectores Ad pueden transducir múltiples tipos de tejidos *in vivo*, incluyendo células que no se dividen, diferenciadas, tales como las que se encuentran en el hígado, riñón y músculo. Los vectores Ad convencionales tienen una gran capacidad de carga de ADN insertado. Un ejemplo del uso de un vector Ad en una prueba clínica implicaba terapia polinucleotídica para la inmunización antitumoral con inyección intramuscular. Sterman *et al.* (1998) *Hum. Gene Ther.* **7**:1083-1089. Otros ejemplos del uso de vectores de adenovirus para la transferencia génica en ensayos clínicos incluyen Rosenecker *et al.* (1996) *Infección* **24**:5-10; Sterman *et al.*, arriba; Welsh *et al.* (1995) *Hum. Gene Ther.* **2**:205-218; Alvarez *et al.* (1997) *Hum. Gene Ther.* **5**:597-613, y Topf *et al.* (1998) *Gene Ther.* **5**:507-513.

Las células de empaquetamiento se usan para formar partículas de virus que sean capaces de infectar una célula huésped. Estas células incluyen células 293, que empaquetan adenovirus, y células Ψ 2 o células PA317, que empaquetan retrovirus. Los vectores víricos utilizados en terapia génica normalmente se generan por una línea celular productora que empaqueta un vector de ácido nucleico en una partícula vírica. Los vectores contienen normalmente las secuencias víricas mínimas requeridas para el empaquetamiento e integración inmediata en un hospedador, siendo otras secuencias virales sustituidas por un casete de expresión para que se exprese la proteína. Las funciones virales que faltan se proporcionan *in trans*, en caso necesario, por la línea celular de empaquetamiento. Por ejemplo, los vectores AAV utilizados en terapia génica normalmente poseen sólo secuencias ITR del genoma de AAV, que se requieren para el empaquetamiento e integración en el genoma hospedador. El ADN viral se empaqueta en una línea celular, que contiene un plásmido auxiliar que codifica los otros genes de AAV, concretamente *rep* y *cap*, pero que carecen de secuencias ITR. La línea celular también se infecta con adenovirus auxiliar. El virus auxiliar promueve la replicación del vector AAV y la expresión de genes de AAV del plásmido auxiliar. El plásmido auxiliar no se empaqueta en cantidades significativas debido a una carencia de secuencias ITR. La contaminación con adenovirus puede reducirse mediante, por ejemplo, tratamiento térmico, que preferentemente inactiva adenovirus.

En muchas aplicaciones de terapia génica, es deseable que el vector de terapia génica pueda ser introducido con un alto grado de especificidad para un tipo de tejido particular. Un vector viral se puede modificar para que tenga especificidad para un tipo de célula dado expresando un ligando como una proteína de fusión con una proteína de la cubierta viral en la superficie exterior del virus. Se elige el ligando para que tenga afinidad por un receptor conocido que esté presente en el tipo celular de interés. Por ejemplo, Han *et al.* (1995) *Proc. Natl. Acad. Sci. EE.UU.* **92**:9747-9751 indica que el virus de la leucemia murina de Moloney puede modificarse para expresar la herregulina humana fusionada con gp70, y el virus recombinante infecta determinadas células de cáncer de mama humano que expresan el receptor del factor de crecimiento epidérmico humano. Este principio puede extenderse a otros pares de virus que expresan una proteína de fusión de ligando y a la célula diana que expresa un receptor. Por ejemplo, el fago filamentoso se puede diseñar para exponer los fragmentos de anticuerpo (por ejemplo, F_{ab} o F_v) que tienen afinidad de unión específica virtualmente para cualquier receptor celular seleccionado. Aunque la descripción anterior se aplica fundamentalmente a vectores víricos, los mismos principios pueden aplicarse a vectores no víricos. Tales vectores se pueden diseñar para contener secuencias específicas de captación pensadas para favorecer la captación por células diana específicas.

Los vectores de terapia génica se pueden introducir *in vivo* mediante la administración a un paciente individual, normalmente por administración sistémica (por ejemplo, infusión intravenosa, intraperitoneal, intramuscular, subdérmica, o intracraneal) o aplicación tópica, como se describe *abajo*. Alternativamente, los vectores se pueden administrar a células *ex vivo*, tales como células explantadas de un paciente específico (por ejemplo, linfocitos, aspirados de médula ósea, tejido de biopsia) o células madre hematopoyéticas de donantes universales, seguido de reimplante de las células a un paciente, normalmente después de la selección de células que hayan incorporado el vector.

La transfección de células *ex vivo* para diagnósticos, investigación, o para terapia génica (por ejemplo, a través de la re-infusión de las células transfectadas en el organismo hospedador) es bien conocida por los expertos en la técnica. En una realización preferida, las células se aíslan del organismo específico, se transfectan con un ácido nucleico (gen o ADNc), y se vuelven a infundir en el organismo específico (por ejemplo, un paciente). Varios tipos celulares adecuados para la transfección *ex vivo* son bien conocidos por los expertos en la técnica. Véase, por ejemplo, Freshney *et al.*, *Culture of Animal Cells, A Manual of Basic Technique*, 3^o ed., 1994, y las referencias allí citadas, para una discusión del aislamiento y cultivo de células de pacientes.

En una realización, se utilizan las células madre hematopoyéticas en procedimientos *ex vivo* para la transfección celular y la terapia génica. La ventaja de usar células madre es que se pueden diferenciar en otros tipos de células *in vitro*, o pueden introducirse en un mamífero (tal como el donante de las células) donde se implantan en la médula ósea. Se conocen métodos para diferenciar células madre CD34 + *in vitro* en tipos de células inmunitarias de importancia clínica utilizando citoquinas como GM-CSF, IFN- γ y TNF- α . Inaba *et al.* (1992) *J. Exp. Med.* **176**:1693-1702.

Las células madre se aíslan para la transducción y diferenciación utilizando métodos conocidos. Por ejemplo, las células madre se aíslan de células de médula ósea cribado de las células de médula ósea con anticuerpos que se unen a células no deseadas, tales como (células T) CD4+ y CD8+, (células panB) CD45+, GR-1 (granulocitos), y lad (células presentadoras de antígeno diferenciadas). Véase Inaba *et al.*, *arriba*.

Los vectores (por ejemplo, retrovirus, adenovirus, liposomas, etc) que contienen ácidos nucleicos terapéuticos se pueden administrar también directamente al organismo para transducción de células *in vivo*. Alternativamente, se puede administrar ADN desnudo. La administración se realiza mediante cualquiera de las rutas utilizadas normalmente para introducir una molécula en contacto final con la sangre o células de tejidos. Los métodos adecuados de administración de tales ácidos nucleicos están disponibles y son bien conocidos por los expertos en la técnica, y, aunque se puede utilizar más de una ruta para administrar una composición específica, una ruta concreta puede a menudo ofrecer una reacción más directa y más eficaz que otra.

Los vehículos farmacéuticamente aceptables están determinados en parte por la composición específica que se administra, así como por el método particular usado para administrar la composición. En consecuencia, hay una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas descritas en este documento. Véase, por ejemplo, *Remington Pharmaceutical Sciences*, 17^a ed., 1989.

5 B. Introducción de polipéptidos

En otras realizaciones, las proteínas de fusión se administran directamente a las células diana. En determinadas situaciones *in vitro*, las células diana se cultivan en un medio que contiene una proteína de fusión que comprende uno o más dominios funcionales fusionados con una o más ZFP modificadas descritas en la presente memoria.

10 Un factor importante en la administración de compuestos polipeptídicos es garantizar que el polipéptido tiene la capacidad de atravesar la membrana plasmática de una célula, o la membrana de un compartimento intracelular como el núcleo. Las membranas celulares están compuestas de bicapas de lípidos y proteínas que libremente permeables para compuestos lipófilos no iónicos pequeños, y son intrínsecamente impermeables para los compuestos polares, macromoléculas, y agentes terapéuticos o de diagnóstico. Sin embargo, se han descrito proteínas, lípidos y otros compuestos, que tienen la capacidad de translocar polipéptidos a través de una membrana celular.

15 Por ejemplo, los "polipéptidos de translocación de membrana" tienen subsecuencias de aminoácidos anfífilos o hidrófobos que tienen la capacidad de actuar como transportadores de translocación a través de la membrana. En una realización, las proteínas con homeodominio tienen la capacidad de translocarse a través de las membranas celulares. El péptido más corto internalizable de una proteína con homeodominio, *Antennapedia*, resultó ser la tercera hélice de la proteína, desde el aminoácido de posición 43 a 58. Prochiantz (1996) *Curr. Opin. Neurobiol.* 6:629-634. Otra subsecuencia, el dominio h (hidrófobo) de los péptidos señal, resultó tener características similares de translocación de la membrana celular. Lin *et al.* (1995) *J. Biol. Chem.* 270:14255-14258.

25 Ejemplos de secuencias de péptidos que pueden ser vinculados a polipéptidos con dedos de zinc no canónicos (o de fusión que las contengan) para facilitar su absorción en las células incluyen, pero no se limitan a: un péptido de 11 aminoácidos de la proteína *tat* de VIH; una secuencia peptídica de 20 restos que corresponde a los aminoácidos 84-103 de la proteína p16 (véase Fahraeus *et al.* (1996) *Curr. Biol.* 6:84), la tercera hélice de 60 aminoácidos del homeodominio largo de *Antennapedia* (Derossi *et al.* (1994) *J. Biol. Chem.* 269:10444), la región h de un péptido señal, tal como la región h del factor de crecimiento de fibroblastos de Kaposi (K-FGF) (Lin *et al.*, *arriba*); y el dominio de translocación VP22 de HSV (Elliot *et al.* (1997) *Cell* 88:223-233). Otras fracciones químicas adecuadas que proporcionan una absorción celular mejorada también pueden estar unidas, ya sea covalentemente o no covalentemente, a las ZFP.

35 Las moléculas de toxinas tienen también la capacidad de transportar polipéptidos a través de membranas celulares. A menudo, tales moléculas (llamadas "toxinas binarias") se componen de al menos dos partes: un dominio de translocación o de unión y un dominio de toxina separado. Normalmente, el dominio de translocación, que opcionalmente puede ser un polipéptido, se une a un receptor celular, facilitando el transporte de la toxina dentro de la célula. Varias toxinas bacterianas, incluyendo la toxina de *Clostridium perfringens* *iota*, toxina de la difteria (DT), *exotoxina A* de *Pseudomonas* (PE), la toxina de *Bacillus anthracis* (PT), la toxina de *Bacillus anthracis*, y la adenilato ciclasa pertusis (CYA), han sido utilizadas para introducir péptidos al citosol de la célula con fusiones internas o amino-terminales. Arora *et al.* (1993) *J. Biol. Chem.* 268:3334-3341; Perelle *et al.* (1993) *Infect. Immun.* 61:5147-5156; Stenmark *et al.* (1991) *J. Cell Biol.* 113:1025-1032; Donnelly *et al.* (1993) *Proc. Natl. Acad. Sci. EE.UU.* 90:3530-3534; Carbonetti *et al.* (1995) *Abstr. Annu. Meet. Soc. Microbiol.* 95:295; Sebo *et al.* (1995) *Infect. Immun.* 63:3851-3857; Klimpel *et al.* (1992) *Proc. Natl. Acad. Sci. EE.UU.* 89:10277-10281, y Novak *et al.* (1992) *J. Biol. Chem.* 267:17186-17193.

45 Tales subsecuencias se pueden utilizar para la translocación de polipéptidos, incluyendo los polipéptidos que se describen en la presente memoria, a través de una membrana celular. Esto se logra, por ejemplo, por derivatización del polipéptido de fusión con una de estas secuencias de translocación, o mediante la formación de una fusión adicional de la secuencia de translocación con el polipéptido de fusión. Opcionalmente, se puede utilizar un enlazador para conectar el polipéptido de fusión y la secuencia de translocación. Se puede utilizar cualquier enlazador adecuado, por ejemplo, un enlazador peptídico.

50 También se puede introducir un polipéptido adecuado en una célula animal, preferiblemente una célula de mamífero, a través de liposomas y derivados de liposomas tales como inmunoliposomas. El término "liposoma" se refiere a vesículas que comprenden una o más bicapas lipídicas ordenadas concéntricamente, que encapsulan una fase acuosa. La fase acuosa contiene normalmente el compuesto que se va a introducir en la célula.

55 El liposoma se fusiona con la membrana plasmática, liberando así el compuesto en el citosol. Alternativamente, el liposoma es fagocitado o absorbido por la célula en una vesícula de transporte. Una vez en el endosoma o fagosoma, el liposoma es degradado o se fusiona con la membrana de la vesícula de transporte y libera su contenido.

En los métodos actuales de administración de fármacos mediante liposomas, el liposoma finalmente se hace permeable y libera el compuesto encapsulado en el tejido o célula diana. Para la introducción sistémica o específica para un tejido, ésta se puede lograr, por ejemplo, de una manera pasiva cuando la bicapa del liposoma se degrada con el tiempo por la acción de diversos agentes del cuerpo. Alternativamente, la liberación del fármaco activo implica el uso de un agente que induzca un cambio de permeabilidad en la vesícula del liposoma. Se pueden construir las membranas de los liposomas de manera que se desestabilicen cuando el entorno se vuelve ácido cerca de la membrana del liposoma. Véase, por ejemplo, *Proc. Natl. Acad. Sci. EE.UU.* **84**:7851 (1987); *Biochemistry* **28**:908 (1989). Cuando los liposomas se introducen por endocitosis en célula diana, por ejemplo, se desestabilizan y liberan sus contenidos. Esta desestabilización se denomina fusogénesis.

La dioleoilfosfatidiletanolamina (DOPE) es la base de muchos sistemas "fusogénicos".

Para el uso con los métodos y composiciones descritos en este documento, los liposomas comprenden normalmente un polipéptido de fusión como se describe en la presente memoria, un componente lipídico, por ejemplo, un lípido neutro y/o catiónico y, opcionalmente, incluyen una molécula de reconocimiento de receptores como un anticuerpo que se une a un receptor de la superficie celular o ligando predeterminado (por ejemplo, un antígeno). Se encuentran disponibles varios de métodos para preparar liposomas como se describen en, por ejemplo,;

EE.UU. Pat. N° 4.186.183; 4.217.344; 4.235.871; 4.261.975; 4.485.054; 4.501.728; 4.774.085; 4.837.028; 4.235.871; 4.261.975; 4.485.054; 4.501.728; 4.774.085; 4.837.028; 4.946.787; Publicación PCT N° WO 91/17424; Szoka et al. (1980) *Ann. Rev. Biophys. Bioeng.* **9**:467; Deamer et al. (1976) *Biochim. Biophys. Acta* **443**:629-634; Fraley, et al. (1979) *Proc. Natl. Acad. Sci. EE.UU.* **76**:3348-3352; Hope et al. (1985) *Biochim. Biophys. Acta* **812**:55-65; Mayer et al. (1986) *Biochim. Biophys. Acta* **858**:161-168; Williams et al. (1988) *Proc. Natl. Acad. Sci. EE.UU.* **85**:242-246; *Liposomes*, Ostro (ed.), 1983, capítulo 1; Hope et al. (1986) *Chem. Phys. Lip.* **40**:89; Gregoriadis, *Liposomes Technology* (1984) y Lasic, *Liposomes: from Physics to Applications* (1993). Los métodos adecuados incluyen, por ejemplo, sonicación, extrusión, homogenización/alta presión, microfluidización, diálisis con detergente, fusión inducida por calcio de pequeñas vesículas de liposomas y métodos de fusión-éter, todos los cuales son bien conocidos en la técnica.

En determinadas realizaciones, puede ser deseable dirigir un liposoma utilizando restos de direccionamiento que sean específicos para un tipo particular de célula, tejido, y similares. La orientación de liposomas usando diversos de restos de direccionamiento (por ejemplo, ligandos, receptores, y anticuerpos monoclonales) se ha descrito previamente. Véase, por ejemplo, las patentes de EE.UU n° 4.957.773 y 4.603.044.

Ejemplos de restos de direccionamiento incluyen anticuerpos monoclonales específicos para antígenos asociados con neoplasmas, tales como el antígeno prostático específico de cáncer y MAGE. Los tumores también pueden ser diagnosticados mediante la detección de los productos génicos resultantes de la activación o sobreexpresión de oncogenes, tales como *ras* o *c-erbB2*. Además, muchos tumores expresan antígenos expresados normalmente por el tejido fetal, tales como la alfafetoproteína (AFP) y el antígeno carcinoembrionario (CEA). Se pueden diagnosticar los sitios de infección viral utilizando diversos antígenos virales tales como los antígenos del núcleo de la hepatitis B y antígenos de superficie de la hepatitis C (HBVc, VHBS), antígenos del virus de Epstein-Barr, antígenos del virus tipo-1 de inmunodeficiencia humana (VIH-1) y antígenos del virus del papiloma. La inflamación puede detectarse utilizando moléculas reconocidas específicamente por moléculas de superficie que se expresan en sitios de inflamación tales como integrinas (por ejemplo, VCAM-1), receptores de selectina (por ejemplo, ELAM-1) y similares.

Se utilizan métodos estándar para acoplamiento de agentes de direccionamiento a liposomas. Estos métodos implican generalmente la incorporación en los liposomas de componentes lipídicos, por ejemplo, fosfatidiletanolamina, que puede activarse para la fijación de agentes de direccionamiento, o incorporación de compuestos lipófilos derivatizados, tales como bleomicina lipídica derivatizada. Se pueden construir liposomas dirigidos a anticuerpos usando, por ejemplo, liposomas que incorporan la proteína A. Veas Renneisen et al. (1990) *J. Biol. Chem.* **265**:16337-16342 y Leonetti et al. (1990) *Proc. Natl. Acad. Sci. EE.UU* **87**:2448-2451.

Composiciones farmacéuticas y administración

Las proteínas con dedos de zinc modificadas y las moléculas de fusión descritas en la presente memoria, y los vectores de expresión que codifican estos polipéptidos, se pueden utilizar en conjunto con diversos métodos de terapia génica para facilitar la acción de un producto génico terapéutico. En tales aplicaciones, las composiciones que contienen ZFP se pueden administrar directamente a un paciente, por ejemplo, para facilitar la modulación de la expresión génica y para aplicaciones terapéuticas o profilácticas, por ejemplo, cáncer (incluyendo tumores asociados con el tercer gen del tumor de Wilms), isquemia, retinopatía diabética, degeneración macular, artritis reumatoide, psoriasis, infección por VIH, anemia de células falciformes, enfermedad de Alzheimer, distrofia muscular, enfermedades neurodegenerativas, enfermedad vascular, fibrosis quística, derrame cerebral, y similares. Ejemplos de microorganismos cuya inhibición puede facilitarse mediante el uso de los métodos y composiciones descritos en la presente memoria incluyen bacterias patógenas, por ejemplo, *Clamidia*, bacterias *Rickettsia*, *Micobacterias*, *Estafilococos*, *Streptococos*, *Neumococos*, *Meningococos* y *Conococos*, *Klebsiella*, *Proteus*, *Serratia*, *Pseudomonas*, *Legionella*, *Difteria*, *Salmonella*, *Bacilos* (por ejemplo, ántrax), *Vibrio* (por ejemplo, cólera), *Clostridium* (por ejemplo, el tétanos, botulismo), *Yersinia* (por ejemplo, peste), *Leptospirosis*, y *Borrellia* (por ejemplo,

bacterias de la enfermedad de Lyme); hongos infecciosos, por ejemplo, *Aspergillus*, especies de *Candida*; protozoos tales como esporozoos (por ejemplo, *Plasmodia*), rizópodos (por ejemplo, *Entamoeba*) y flagelados (*Trypanosoma*, *Leishmania*, *Trichomonas*, *Giardia*, etc), virus, por ejemplo, hepatitis (A, B, o C), herpes virus (por ejemplo, VZV, HSV-1, HHV-6, HSV-II, CMV, y EBV), HIV, Ebola, Marburg y virus relacionados con los que causan la fiebre hemorrágica, adenovirus, virus de la influenza, flavivirus, echovirus, rinovirus, virus coxsackie, coronaviruses, virus sincitial respiratorio, virus de la parotiditis epidémica, rotavirus, virus del sarampión, rubéola, parvovirus, virus vaccinia, virus HTLV, retrovirus, lentivirus, virus del dengue, virus del papiloma, virus de la polio, virus de rabia, y los virus de encefalitis por arbovirus, etc.

La administración de cantidades terapéuticamente eficaces de ZFP modificadas descritos en la presente memoria, las moléculas de fusión que incluyen estas ZFP, o los ácidos nucleicos que codifican estos polipéptidos, se realiza por cualquiera de las rutas utilizadas normalmente para introducir polipéptidos o ácidos nucleicos en contacto final con el tejido a tratar. Los polipéptidos o ácidos nucleicos se administran de cualquier manera adecuada, preferiblemente con vehículos farmacéuticamente aceptables. Los métodos adecuados de administración de tales moduladores están disponibles y son bien conocidos por los expertos en la técnica, y, aunque se puede utilizar más de una ruta para administrar una composición particular, una ruta específica puede proporcionar a menudo una reacción más inmediata y más eficaz que otra.

Los vehículos farmacéuticamente aceptables vienen determinados en parte por la composición particular que se administra, así como por el método particular usado para administrar la composición. En consecuencia, hay una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas. Véase, por ejemplo, *Remington's Pharmaceutical Sciences*, 7^o ed., 1985.

Las ZFP y los polipéptidos de fusión de ZFP o los ácidos nucleicos, solos o en combinación con otros componentes adecuados, se pueden preparar en formulaciones de aerosol (es decir, se pueden "nebulizar") para ser administrados por inhalación. Las formulaciones en aerosol se pueden emplear con propelentes presurizados aceptables, tales como diclorodifluorometano, propano, nitrógeno y similares.

Las formulaciones adecuadas para administración parenteral, tales como, por ejemplo, por vía intravenosa, intramuscular, intradérmica y subcutánea, incluyen, soluciones inyectables estériles isotónicas acuosas y no acuosas, que pueden contener antioxidantes, amortiguadores, bacteriostáticos y solutos que hacen que la formulación sea isotónica con la sangre del receptor previsto, y suspensiones estériles acuosas y no acuosas que pueden incluir agentes en suspensión, solubilizantes, agentes espesantes, estabilizadores y conservantes. Las composiciones se pueden administrar, por ejemplo, por infusión intravenosa, por vía oral, tópica, intraperitoneal, intravesical o intratecal. Las formulaciones de compuestos se pueden presentar en dosis unitarias o en recipientes de dosis múltiples sellados, tales como ampollas y viales. Las soluciones y suspensiones inyectables se pueden preparar a partir de polvos estériles, gránulos y comprimidos del tipo de los conocidos por los expertos en la técnica.

Aplicaciones

Las composiciones y métodos descritos aquí se pueden utilizar para facilitar una serie de procesos que implican la regulación transcripcional. Estos procesos incluyen, pero no están limitados a, la transcripción, la replicación, la recombinación, la reparación, la integración, el mantenimiento de los telómeros, los procesos implicados en la estabilidad y disyunción de los cromosomas, y el mantenimiento y la propagación de las estructuras de cromatina. En consecuencia, los métodos y composiciones descritos en la presente memoria pueden usarse para afectar a cualquiera de estos procesos, así como a cualquier otro proceso que pueda ser influenciado por ZFP o ZFP de fusión.

En realizaciones preferidas, se utilizan una o más de las moléculas descritas en la presente memoria para lograr la activación o represión dirigida de la expresión génica, por ejemplo, en base a la especificidad de la ZFP modificada. En otra realización, una o más de las moléculas descritas en la presente memoria se utilizan para conseguir la reactivación de un gen, por ejemplo un gen silenciado de desarrollo, o para lograr la activación mantenida de un transgén. La ZFP modificada se puede dirigir a una región fuera de la región de codificación del gen de interés y, en determinadas realizaciones, se dirige a una región fuera de la(s) región(es) reguladora(s) del gen. En estas formas de realización, se puede utilizar moléculas adicionales, exógenas y/o endógenas, para facilitar la represión o activación de la expresión génica. Las moléculas adicionales también pueden ser moléculas de fusión, por ejemplo, fusiones entre una ZFP y un dominio funcional tales como un dominio de activación o represión. Véase, por ejemplo, en la solicitud de propiedad conjunta WO 00/41566.

En consecuencia, la expresión de cualquier gen de cualquier organismo puede ser modulada utilizando los métodos y composiciones descritos aquí, incluyendo genes relevantes terapéuticamente, genes de microorganismos infecciosos, genes de virus, y los genes cuya expresión se modula en los procesos de descubrimiento de fármacos y/o de validación de dianas. Tales genes incluyen, pero no se limitan al, tercer gen del tumor de Wilms (WT3), a factores de crecimiento de endotelio vascular (VEGF), a receptores de VEGF CCR-5 (por ejemplo, *flt* y *flk*), al receptor de lipoproteínas de baja densidad (LDLR), al receptor de estrógeno, HER-2/neu, BRCA-1, BRCA-2, a carboxiquinasa fosfoenolpiruvato (PEPCK), CYP7, a fibrinógeno, a la apolipoproteína A (ApoA), a la apolipoproteína B (ApoB), a renina, a carboxiquinasa fosfoenolpiruvato (PEPCK), CYP7, a fibrinógeno, al factor nuclear κ B (NF- κ B),

al inhibidor de NF- κ B (I- κ B), a los factores de necrosis tumoral (por ejemplo, TNF- α , TNF- β), a la interleuquina-1 (IL-1), FAS (CD95), al ligando de FAS (CD95L), al factor natriurético atrial, al factor derivado de plaquetas (PDF), a la proteína precursora amiloidea (APP), a la tirosinasa, a la tirosina hidroxilasa, a la β -aspartil hidroxilasa, a la fosfatasa alcalina, a calpaínas (por ejemplo, CAPN10), al receptor de pentraxina neuronal, a la proteína de respuesta adriamycin, a la apolipoproteína E (apoE), a leptina, al receptor de leptina, UCP-1, IL-1, a los receptores de IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-12, IL-15, a receptores de interleuquinas, G-CSF, GM-CSF, al factor estimulante de colonias, a eritropoyetina (EPO), al factor de crecimiento derivado de plaquetas (PDGF), al receptor de PDGF, al factor de crecimiento de fibroblastos (FGF), al receptor de FGF, PAF, p16, p19, p53, Rb, p21, *myc*, *myb*, globina, distrofina, eutrofina, regulador de la conductancia transmembrana de la fibrosis quística (CFTR), GDNF, al factor de crecimiento nervioso (NGF), al receptor de NGF, al factor de crecimiento epidérmico (EGF), al receptor de EGF, a los factores de crecimiento transformante (por ejemplo, TGF- α , TGF- β), al factor de crecimiento de fibroblastos (FGF), a los interferones (por ejemplo, IFN- α , IFN- β y IFN- γ), al factor de crecimiento-1 similar a la insulina del (IGF-1), a la angiostatina, ICAM-1, al transductor de señales y activador de la transcripción (STAT), a los receptores de andrógenos, a e-cadherina, catepsinas (por ejemplo, la catepsina W), topoisomerasa, telomerasa, *bcl*, *bcl-2*, *Bax*, a la tirosina quinasa específica de células T (LCK), a la proteína quinasa activada por mitógeno p38, a la proteína tirosina fosfatasa (HPTP), a la adenilato ciclasa, a la guanilato ciclasa, al receptor de acetilcolina nicotínico neuronal $\alpha 7$, al receptor de 5-hidroxitriptamina (serotonina) 2A, al factor alargamiento de la transcripción-3 (TEF-3), a la fosfatidilcolina transferasa, *fts*, PTI-1, a la poligalacturonasa, EPSP sintasa, FAD2-1, desaturasa Δ -9, desaturasa Δ -12, desaturasa Δ -15, a la carboxilasa acetil-coenzima A, a la acil-ACP tioesterasa, a la ADP-glucosa pirofosforilasa, a la almidón sintasa, a la celulosa sintasa, a la sacarosa sintasa, a la liasa de hidroperóxido de ácidos grasos, y a los peroxisoma activados por el proliferador receptores, tales como PPAR- γ 2.

Se puede modular la expresión de los genes humanos, de mamífero, bacterianas, fúngicas, protozoarias, de plantas, de Arquea, y virales; los genes virales incluyen, pero no están limitados a, genes de virus de la hepatitis, como, por ejemplo, VHB-C, VHB-S, VHBX- y VHB-P, y los genes del VIH, tales como, por ejemplo, *tat* y *rev*. Se puede lograr la modulación de la expresión de genes que codifican antígenos de un organismo patógeno utilizando los métodos y composiciones descritas.

Otros genes incluyen aquellos que codifican citoquinas, linfoquinas, interleuquinas, factores de crecimiento, factores mitógenos, factores apoptóticos, citocromos, factores quimiotácticos, receptores de quimioquinas (por ejemplo, CCR-2, CCR-3, CCR-5, CXCR-4), fosfolipasas (por ejemplo, la fosfolipasa C), receptores nucleares, los receptores retinoides, receptores de orgánulos, hormonas, receptores hormonales, oncogenes, supresores tumorales, ciclinas, proteínas de control del ciclo celular (por ejemplo, Chk1, Chk2), genes asociados a la senescencia, genes que codifican inmunoglobulinas, quelantes de metales pesados, proteínas tirosina quinasas, proteínas tirosina fosfatasa, factores asociados a receptores de factores de necrosis tumoral (por ejemplo, Traf-3, Traf-6), apolipoproteínas, factores trombóticos, factores vasoactivos, neuro-receptores, receptores de superficie celular, proteínas G, receptores acoplados a proteínas-G (por ejemplo, receptor de la sustancia K, receptor de la angiotensina, receptores α - y β -adrenérgicos, receptores de serotonina, y receptores de PAF), receptores muscarínicos, receptores de acetilcolina, receptores de GABA, receptores de glutamato, receptores de dopamina, proteínas de adhesión (por ejemplo, CAM, selectinas, integrinas y miembros de la superfamilia de las inmunoglobulinas), canales iónicos, factores asociados receptores, factores hematopoyéticos, factores de transcripción, y moléculas participan en la transducción de la señal. También se puede modular la expresión de genes relacionados con enfermedades, y/o de uno o más genes específicos de un tejido particular o de un tipo de célula, tal como, por ejemplo, cerebro, músculo, corazón, sistema digestivo, sistema nervioso, sistema circulatorio, sistema reproductor, sistema genitourinario, y sistema respiratorio.

Otras aplicaciones incluyen métodos terapéuticos en los que se administra una ZFP modificada, un polipéptido de fusión de ZFP, o un ácido nucleico que codifica una ZFP modificada o una ZFP de fusión a un sujeto y se utiliza para modular la expresión de un gen diana en el sujeto (como se describe, por ejemplo, en la solicitud de propiedad conjunta PCT WO 00/41566). La modulación puede ser en forma de represión, por ejemplo, cuando el gen diana reside en un microorganismo infectante patológico, o en un gen endógeno del paciente, tales como un oncogén o receptor viral, que favorece un estado de enfermedad. Alternativamente, la modulación puede ser en forma de activación, cuando la activación de la expresión o el aumento de expresión de un gen celular endógeno (tal como, por ejemplo, un gen supresor de tumor) pueden mejorar un estado de la enfermedad. Ejemplos de polipéptidos de fusión de ZFP, tanto para la activación y la represión de la expresión génica se describen *arriba*. Para tales aplicaciones, las ZFP modificadas, los polipéptidos de fusión de ZFP o, más especialmente, los ácidos nucleicos que los codifican se formulan con un vehículo farmacéuticamente aceptable como una composición farmacéutica.

Los vehículos farmacéuticamente aceptables y excipientes vienen determinados en parte por la composición particular que se administra, así como por el método particular utilizado para administrar la composición. Véase, por ejemplo, *Remington's Pharmaceutical Sciences*, 17^a ed. 1985. Las ZFP, polipéptidos de fusión de ZFP, o polinucleótidos que codifican polipéptidos de fusión de ZFP, solos o en combinación con otros componentes adecuados, pueden prepararse en formulaciones aerosoles (es decir, pueden ser "nebulizadas") para administrarse por inhalación. Las formulaciones en aerosol se pueden emplear en propelentes presurizados aceptables, tales como diclorodifluorometano, propano, nitrógeno y similares. Las formulaciones adecuadas para administración parenteral, tales como, por ejemplo, por vía intravenosa, intramuscular, intradérmica y subcutánea, incluyen soluciones inyectables estériles isotónicas, acuosas y no acuosas, que pueden contener antioxidantes, tampones,

bacteriostáticos y solutos que hacen que la formulación sea isotónicas con la sangre del receptor pretendido, y con las soluciones estériles acuosas y no acuosas que pueden incluir agentes en suspensión, solubilizantes, agentes espesantes, estabilizantes y conservantes. Las composiciones se pueden administrar, por ejemplo, por infusión intravenosa, por vía oral, tópica, intraperitoneal, intravesical o intratecal. Las formulaciones de compuestos pueden presentarse en dosis unitarias o en recipientes de dosis múltiples sellados, tales como ampollas y viales. Las soluciones y suspensiones inyectables se pueden preparar a partir de polvos estériles, gránulos y comprimidos del tipo descrito previamente.

La dosis administrada a un paciente debería ser suficiente para causar en el paciente a una respuesta terapéutica beneficiosa con el tiempo. La dosis viene determinada por la afinidad y la eficacia de unión (K_d) de la ZFP particular empleada, de la célula diana, y del estado del paciente, así como del peso corporal o de la superficie a tratar del paciente. El tamaño de la dosis también está determinado por la existencia, naturaleza, y extensión de los efectos secundarios adversos que acompañan a la administración de un compuesto particular o de un vector a un paciente particular.

En otras aplicaciones, se utilizan las ZFP modificadas y otras proteínas de unión a ADN y/o ARN en métodos de diagnóstico para la detección específica de la secuencia de ácido nucleico diana en una muestra. Por ejemplo, los ZFP modificados pueden usarse para detectar alelos variantes asociados con una enfermedad o un fenotipo en muestras de pacientes. Como ejemplo, las ZFP modificadas pueden ser utilizadas para detectar la presencia de especies de ARNm o ADNc particulares en una mezcla compleja de ARNm o ADNc. Como un ejemplo adicional, la ZFP modificada puede ser utilizada para cuantificar el número de copias de un gen en una muestra. Por ejemplo, la detección de pérdida de una copia de un gen p53 en una muestra clínica es un indicador de la susceptibilidad al cáncer. En un ejemplo adicional, las ZFP modificadas se utilizan para detectar la presencia de microorganismos patógenos en muestras clínicas. Esto se logra mediante el uso de una o más ZFP modificadas, como se describe en la presente memoria, que se unen a una secuencia diana en uno o más genes dentro del microorganismo a detectar. Un formato adecuado para la realización de ensayos de diagnóstico emplea ZFP modificadas vinculadas a un dominio que permite la inmovilización de la ZFP en un soporte sólido tal como, por ejemplo, una placa de microtitulación o una placa de ELISA. La ZFP inmovilizada se pone en contacto con una muestra sospechosa de contener un ácido nucleico diana en condiciones en las que pueda tener lugar la unión entre la ZFP modificada y su secuencia diana. Normalmente, los ácidos nucleicos de la muestra están marcados (por ejemplo, durante la amplificación por PCR). Alternativamente, los ácidos nucleicos no marcados se pueden detectar usando una segunda sonda de ácido nucleico marcada. Después del lavado, se detectan los ácidos nucleicos marcados unidos. El marcaje puede ser directo (es decir, la sonda se une directamente con el ácido nucleico diana) o indirecto (es decir, la sonda se une a una o más moléculas que se unen a la diana). El marcaje puede ser, por ejemplo, radioactivo, quimioluminiscente fluorescente, y/o enzimático.

Las ZFP modificadas, tal como se describe en la presente memoria, también se pueden utilizar en ensayos que relacionan el fenotipo con la expresión de genes particulares. Las metodologías actuales para la determinación de la función génica se basan principalmente en ya sea la sobre-expresión de un gen de interés o la eliminación de un gen de interés de su entorno biológico natural, y la observación de los efectos. Los efectos fenotípicos resultantes de la sobre-expresión o del knockout se interpretan entonces como una indicación de la función del gen en el sistema biológico. Un ejemplo de sistema modelo animal para llevar a cabo este tipo de análisis es el ratón. Un ratón transgénico contiene generalmente un gen introducido o ha sido modificado genéticamente con el fin de incrementar un gen endógeno. Alternativamente, en un ratón "knock-out", se ha eliminado un gen endógeno o su expresión ha sido suprimida. Hay varios problemas con estos sistemas existentes, muchos de los cuales están relacionados con el hecho de que sólo es posible lograr una modulación de la expresión génica en estos sistemas de "todo o nada". El primero es la capacidad limitada de modular la expresión del gen objeto de estudio (por ejemplo, en ratones knock-out, el gen objeto de estudio está generalmente ausente en el genoma o es completamente no funcional, mientras que en ratones transgénicos que sobreexpresan un gen particular, generalmente hay un único nivel de sobreexpresión). El segundo es el requisito de encontrar determinados genes a menudo en múltiples etapas del desarrollo. Así pues, no es posible determinar la función de un gen adulto en particular, cuya actividad también es necesaria durante el desarrollo embrionario, mediante la generación de un knock-out de ese gen, ya que los animales que contienen el knock-out no sobrevivirán a la edad adulta.

Una ventaja de usar la regulación de un gen por medio de la ZFP para determinar su función, en relación con el análisis knockout convencional anteriormente mencionado, es que la expresión de una ZFP puede ser empleada bajo el control de moléculas pequeñas. Véase, por ejemplo, las patentes de EE.UU. nº 5.654.168; 5.789.156; 5.814.618; 5.888.981; 6.004.941; 6.087.166; 6.136.954, y la WO 00/41566 de propiedad conjunta. Mediante el control de los niveles de expresión de las ZFP, uno puede a su vez controlar los niveles de expresión de un gen regulado por la ZFP para determinar qué grado de represión o estimulación de la expresión se requiere para conseguir un efecto fenotípico o bioquímico dado. Este enfoque tiene un especial valor para el desarrollo de fármacos. Además, el uso de la expresión de ZFP controlada por moléculas pequeñas permite superar los problemas mencionados anteriormente de la letalidad embrionaria y compensación del desarrollo, mediante el inicio de la expresión de la ZFP en una etapa posterior en el desarrollo y la observación de los efectos en el animal adulto.

Se pueden producir ratones transgénicos que tienen genes diana regulados por una ZFP modificada o una proteína de fusión de ZFP por integración del ácido nucleico que codifica la ZFP modificada o la ZFP de fusión en cualquier

sitio en *trans* para el gen diana. En consecuencia, no es necesaria la recombinación homóloga para la integración del ácido nucleico que codifica la ZFP. Además, debido a que la actividad reguladora de la transcripción de una ZFP modificada o una ZFP de fusión es trans-dominante, sólo es necesaria una para obtener animales que tengan una copia cromosómica de un ácido nucleico que codifica una ZFP. Por lo tanto, se pueden producir los animales knock-out funcionales sin retrocruzamiento.

Los siguientes ejemplos se presentan como ilustrativos, pero no limitativos, de la materia reivindicada.

Ejemplos

Ejemplo 1. Producción de proteínas de unión con dedos de zinc no-canónicos

Se obtuvieron genes sintéticos que codifican proteínas de unión con dedos de zinc no-canónicos siguiendo el procedimiento descrito la solicitud de propiedad conjunta PCT WO 00/42219, con la excepción de que el oligonucleótido que codifica la hélice de reconocimiento que va a ser modificada incluye una secuencia de polinucleótidos que especifica la secuencia de aminoácidos modificada. Por ejemplo, para la modificación de dedo 3 (el dedo más C-terminal de una ZFP de tres dedos), la secuencia del oligonucleótido 6 está diseñada para codificar el (los) resto(s) de coordinación del zinc modificado(s).

Ejemplo 2. Modulación de la expresión del gene LCK con ZFP no canónica

En este experimento, la proteína con dedos de zinc "PTP2" de diseño que reconoce la secuencia diana GAGGGGGCG y regula la expresión del gen LCK, fue modificada por medio de la sustitución de la 2ª histidina en el tercer dedo con cisteína (para producir la proteína "PTP2 (H-> C)"). Se cambiaron dos restos adyacentes por glicina también para mejorar el potencial de la cisteína introducida para coordinar productivamente zinc. Las secuencias de las proteínas con dedos de zinc resultantes fueron las siguientes:

PTP2:

F1 PGKKKQHICHIQGCGKVYGRSDELTRHLRWHTGER (SEQ ID NO:112)
 F2 PFMCTWSYCGKRFTRSDHLTRHKRTHTGEK (SEQ ID NO:113)
 F3 KFACPE----CPKRFMRSDNLTRHIKTHQNKKGGS (SEQ ID NO:114)

PTP2(H->C):

F1 PGKKKQHICHIQGCGKVYGRSDELTRHLRWHTGER (SEQ ID NO:115)
 F2 PFMCTWSYCGKRFTRSDHLTRHKRTHTGEK (SEQ ID NO:116)
 F3 KFACPE---CPKRFMRSDNLTRHIGGCQNKKGGS (SEQ ID NO:117)

El subrayado y la negrita resalta los restos de coordinación de zinc, y las posiciones en cursiva resaltan que cambian en la conversión de PTP2 a PTP2 (H-> C).

Ambas ZFP se expresaron en células 293 como fusiones con una señal de localización nuclear (NLS), el dominio de activación de VP16, y una etiqueta FLAG. La estructura (por ejemplo, el orden) de las proteínas de fusión fueron los siguientes:

NLS	VZPF	VP16	FLAG
-----	------	------	------

Después de la expresión de cada proteína en las células 293, se determinaron los niveles celulares del ARNm de LCK en relación con el nivel de un ARN control (ARN 18S) usando un ensayo "Taqman" basado PCR. También se determinaron los niveles de ARN para una proteína de control (NVF) que carece de cualquier ZFP (y que contiene sólo las regiones NLS, VP16 y FLAG). Cada experimento se realizó por duplicado, y los ratios de ARN medidos se muestran en la figura 1. Estas ratios indican que la ZFP PTP2 activa la expresión del gen LCK, y que las ZFP PTP (H-> C) activan LCK a niveles aún más altos. Estos resultados ilustran el potencial de las sustituciones en posiciones de coordinación de zinc para proporcionar ZFP con la función celular mejorada. Como se ilustra en la Figura 1, la modificación de las posiciones de coordinación de zinc puede mejorar la actividad celular de los factores de transcripción de proteínas con dedos de zinc de diseño.

Ejemplo 3. La modulación de la expresión de un gen de VEGF humano con las ZFP modificadas

Este ejemplo describe la modificación de dos ZFP reguladoras de VEGF. Para cada una de las dos ZFP, se construyeron varias ZFP modificadas de no canónicas. Se analizó la capacidad de las proteínas para regular la expresión de VEGF y se comparó con las dos proteínas parentales C2H2.

Se han descrito proteínas con dedos de zinc que comprenden una serie de dedos de zinc C₂H₂, y diseñadas para unirse al VEGF humano de un gen y regulan su expresión. Liu *et al.* (2001) *J. Biol. Chem.* **276**:11323-11334. Dos de estas ZFP (denominadas VOP30A y VOP32B), cada una con tres dedos de zinc, se convirtieron en ZFP no canónicas. VOP30A corresponde a VZ+42/+530 y VOP32B corresponde a VZ+434a en la referencia de Liu *et al.* Esto se logró mediante la modificación del tercer dedo de cada proteína. Se realizaron siete versiones no canónicas de cada proteína, comprendiendo cada una un tercer dedo C2HC no canónico diferente. En la Tabla 1 se muestran las secuencias de aminoácidos de las partes de las ZFP canónicas parentales y cada una de las ZFP no canónicas, comenzando por histidina +7 (con respecto al inicio de la hélice alfa) del tercer dedo.

TABLA1

NOMBRE	SECUENCIA	SEQ ID NO.
C2H2	<u>H</u> IK <u>T</u> HQNKKGGS	11
S	<u>H</u> SETG <u>C</u> TKKGGS	12
E	<u>H</u> LKSLTP <u>C</u> TGGS	13
K	<u>H</u> K <u>C</u> G IQNKKGGS	14
CT	<u>H</u> SE <u>N</u> CQGGKGGS	15
C	<u>H</u> IK <u>T</u> CQNKKGGS	16
GC	<u>H</u> IKG <u>C</u> QNKKGGS	17
GGC	<u>H</u> IGG <u>C</u> QNKKGGS	18
Notas 1. secuencias que comienzan en +7 de la hélice alfa del tercer dedo de zinc 2. los restos que participan en la coordinación de metal están en negrita y subrayados 3. la primera fila (proteína designada C2H2) muestra la secuencia de las ZFP parentales		

Las células de riñón embrionario humano (HEK 293) se transfectaron con los ácidos nucleicos que codifican derivados no canónicos de la VOP30A y de proteínas VOP32B de fusión, así como de las proteínas de fusión parentales (canónicas). Las proteínas de fusión también comprenden un dominio de activación de transcripcional VP16, una secuencia de localización nuclear y una etiqueta de epítipo.

Las células se cultivaron en DMEM (medio de Eagle modificado por Dulbecco), suplementado con suero fetal bovino al 10%, en un incubador de 5% de CO₂ a 37°C. Las células se sembraron en placas de 24-pocillos a una densidad de 160.000 células por pocillo. Un día después, cuando las células alcanzaban una confluencia de aproximadamente 70%, los plásmidos que codifican las fusiones ZFP-VP16 se introdujeron en las células utilizando el reactivo Lipofectamine 2000TM (Gibco Life Technologies, Rockville, MD) de acuerdo con las recomendaciones del fabricante, usando 2 µl de Lipofectamine 2000 (TM) y 1 µg de plásmido de ADN por pocillo. Se retiró el medio y se substituyó por medio fresco 16 horas después de la transfección. Cuarenta horas después de la transfección, el medio de cultivo se recolectó y se analizó la expresión de VEGF-A. El contenido de la proteína VEGF-A en el medio de cultivo se analizó usando un kit ELISA de VEGF humano (Quanti-Glo, R & D Systems, Minneapolis, MN) de acuerdo con las instrucciones del fabricante.

Los resultados, mostrados en la figura. 2, indican que los derivados de C2HC de ambas VOP 30A y 32B VOP activan la expresión de VEGF y son por tanto útiles como moléculas reguladoras exógenas específicas.

Ejemplo 4. Producción de proteínas de unión con dedos de zinc modificadas

Este ejemplo describe una estrategia para seleccionar secuencias de aminoácidos de los esqueletos de plantas con dedos de zinc de entre secuencias existentes de plantas con dedos de zinc, y la posterior modificación conceptual de las secuencias de aminoácidos de los dedos de zinc de la planta seleccionadas para optimizar su capacidad de unión a ADN. Se describen a continuación los oligonucleótidos utilizados en la preparación de polinucleótidos que codifican las proteínas que contienen estos dedos de zinc en dispuestas en tándem.

A. Selección de los esqueletos de los dedos de zinc de plantas.

Se realizó una búsqueda de dedos de zinc de plantas en los que las secuencias del esqueleto (es decir, de la parte del dedo de zinc exterior a la partes -1 hasta +6 de la hélice de reconocimiento) se parecía a la de la secuencia consenso SP-1 descrita por Berg (1992) *Proc. Natl. Acad. Sci. EE.UU.* **89**:11109-11110. Las secuencias seleccionadas incluyeron los dos restos de cisteína conservados, un resto conservado básico (lisina o arginina)

situado a dos restos del extremo C-terminal de la segunda cisteína (es decir, C-terminal) de cisteína, un resto de fenilalanina conservado que se sitúa a dos restos del extremo C-terminal del resto básico, los dos restos conservados de histidina, y un resto de arginina conservado que se sitúa a dos restos del extremo C-terminal de la primera (es decir, N-terminal) histidina conservada. A continuación se muestran las secuencias de aminoácidos de estos esqueletos con dedos de zinc de plantas seleccionadas (en comparación con la secuencia consenso de SP-1), con los restos conservados representados en negrita y X haciendo referencia a los restos situados en las posiciones de -1 a +6 en la hélice de reconocimiento (que será diferente entre proteínas diferentes dependiendo de la secuencia diana):

SP-1 consenso:	YKC PECGKS FSXXXXXXXX HQ RTH TG EKP	(SEQ ID NO: 19)
F1	KKK SKGHEC PI CFR VFKXXXXXXXX HK RSHTGEKP	(SEQ ID NO:20)
F2	YK CTVCGK SFSXXXXXXXX HK RLHTGEKP	(SEQ ID NO:21)
F3	FSCNYC Q RK FG XXXXXXXX HV RIH	(SEQ ID NO:22)

-5 -1 5

Se seleccionó el primer dedo (F1) debido a que contenía una secuencia básica N-terminal para el dedo que también se encuentra adyacente al primer dedo de SP-1. El dedo llamado F1 se refiere a una secuencia de Petunia, los dedos F2 y F3 son secuencias de Arabidopsis.

B. Modificación de las estructuras principales de dedos de zinc de plantas

Se modificaron dos de los tres dedos de zinc de plantas (F1 y F3, arriba) de manera que sus secuencias de aminoácidos se parecían más a la secuencia de SP-1, de la siguiente manera.

(Nótese que la secuencia de SP-1 es diferente de la secuencia denominada "SP-1 consenso."). En F3, el resto Y en la posición -2 se convirtió a una G, y la secuencia QNKK (SEQ ID NO: 23) se añadió a la del C-terminal de F3. La secuencia QNKK (SEQ ID NO: 23) está presente en C-terminal para el tercer dedo de SP-1, y permite una mayor flexibilidad de ese dedo, en comparación con los dedos 1 y 2, que están flanqueados por la secuencia de terminación de hélice TGEK / RK / P (SEQ ID NO: 24). Esta flexibilidad puede ser particularmente beneficiosa cuando el tercer dedo se modifica para contener un esqueleto distinto de C2H2, como se describe en la presente memoria. Finalmente, varios aminoácidos se han eliminan del extremo N-terminal de F1. Los esqueletos de los dedos de zinc resultantes tenían las siguientes secuencias:

KSKGHEC**PICFR**VFKXXXXXXXX**HKRS**HTGEKP (SEQ ID NO: 25)

YK**CTVCGK**SFSXXXXXXXX**HK**RLHTGEKP (SEQ ID NO: 26)

FSCNYC**Q**RK**FG**XXXXXXXX**HV**RIHQNKK (SEQ ID NO: 27)

Los restos aminoácidos indicados con X, presentes en la zona de reconocimiento de estos dedos de zinc, se han diseñado o seleccionado dependiendo del sitio diana deseado, de acuerdo con los métodos descritos, por ejemplo, en las solicitudes de propiedad conjunta WO00/41566 y WO00/42219, y/o en las referencias citadas arriba.

C. Secuencias de ácidos nucleicos que codifican esqueletos de ZFP modificadas de plantas

Se utilizaron las secuencias de polinucleótidos siguientes para el diseño de las ZFP de tres dedos de plantas que contienen los esqueletos F1, F2 y F3 descritas anteriormente. Los polinucleótidos que codifican las ZFP de múltiples dedos se diseñaron de acuerdo con un método de solapamiento de oligonucleótidos como se describe en, por ejemplo, las solicitudes de propiedad conjunta WO 00/41566 y WO 00/42219. Los oligonucleótidos H1, H2 y H3 (abajo) comprenden secuencias correspondientes a las secuencias complementarias inversas de las hélices de reconocimiento de los dedos 1-3, respectivamente, en consecuencia, los nucleótidos indicados con una N varían dependiendo de las secuencias de aminoácidos de las hélices de reconocimiento deseadas, lo que, a su vez, depende de la secuencia de nucleótidos del sitio diana. Los oligonucleótidos PB1, PB2 y PB3 codifican las zonas de las láminas beta de los dedos de zinc, que son comunes a todas las construcciones. Se seleccionaron los codones utilizados frecuentemente en Arabidopsis y E. coli para su uso en estos oligonucleótidos.

H1:

5'-CTC ACC GGT GTG AGA ACG CTT GTG NNN NNN NNN NNN NNN NNN NNN
CTT GAA AAC ACG GAA-3'
(SEQ ID NO:28)

H2:

5'-TTC ACC AGT ATG AAG ACG CTT ATG NNN NNN NNN NNN NNN NNN
AGA AAA AGA CTT ACC-3'

(SEQ ID NO:29)

H3:

5'-CTT CTT GTT CTG GTG GAT ACG CAC GTG NNN NNN NNN NNN NNN NNN
NNN ACC GAA CTT ACG CTG-3'

(SEQ ID NO:30)

PB1:

5'-AAGTCTAAGGGTCACGAGTGCCCAATCTGCTTCCGTGTTTCAAG-3'

(SEQ ID NO:31)

PB2:

5'-

TCTCACACCGGTGAGAAGCCATACAAGTGCACCTGTTTGTGGTAAGTCTTTTTCT-3'

(SEQ ID NO:32)

PB3:

5'-

CTTCATACTGGTGAAAAGCCATTCTCTTGCAACTACTGCCAGCGTAAGTTCGGT-3'

(SEQ ID NO:33)

Brevemente, estos seis oligonucleótidos hibridan y se amplifican por medio de la reacción en cadena de la polimerasa. El producto de amplificación inicial se volvió a amplificar usando los cebadores que son complementarios al producto de amplificación inicial y que también contienen las extensiones 5' que contienen los sitios de reconocimiento de las enzimas de restricción, para facilitar la clonación. El segundo producto de amplificación se inserta en un vector que contiene, por ejemplo, uno o más dominios funcionales, secuencias de localización nuclear, y/o etiquetas de epítomos. Véase, por ejemplo, en las solicitudes de propiedad conjunta WO 00/41566 y WO 00/42219.

Ejemplo 5. Construcción de un polinucleótido que codifica una proteína con dedos de zinc modificada para unirse a una secuencia diana predeterminada

Se diseñó una proteína con dedos de zinc modificada de plantas para reconocer la secuencia diana 5'-GAGGGGGCG-3'. Se determinaron las secuencias de la hélice de reconocimiento para F1, F2 y F3, como se muestra en la Tabla 2, y se utilizaron los oligonucleótidos correspondientes a H1, H2 y H3 anteriormente citados, que incluyen también las secuencias que codifican estas hélices de reconocimiento, para el montaje de la PCR como se describió anteriormente.

TABLA 2

Dedo	Diana	Secuencia Hélice	Secuencia nucleótidos para el montaje de la PCR
F1	GCG	RSDELTR SEQ ID NO:109	5'CTCACCGGTGTGAGAACGCTTGTGACGGGTCAACT CGTCAGAACGCTTGAAAACACGGAA-3' (SEQ ID NO:34)
F2	GGG	RSDHLTR SEQ ID NO:110	5'TTCACCAAGTATGAAGACGCTTATGACGGGTCAAGT GGTCAGAACGAGAAAAAGACTTACC-3' (SEQ ID NO:35)
F3	GAG	RSDNLTR SEQ ID NO:111	5'CTTCTTGTTCTGGTGGATACGCACGTGACGGGTCA AGTTGTCAGAACGACCGAACTTACGCTG-3' (SEQ ID NO:36)

A continuación de la amplificación inicial, se llevó a cabo una amplificación secundaria, como se describe anteriormente, utilizando los siguientes cebadores:

PZF: 5'-CGGGGTACCAAGTAAGTCTAAGGGTCAC (SEQ ID NO: 37)

PZR: 5'-GCGCGGATCCACCCTTCTTGTCTGGTGGATACG. (SEQ ID NO: 38)

La PZF incluye un sitio KpnI (subrayado) y se solapa con la secuencia de PB1 (solapamiento indicado en negrita). La PZR incluye un sitio BamHI (subrayado) y se solapa con H3 (indicadas en negrita).

El producto de amplificación secundario se digiere con Kpn I y Bam HI y se inserta en un vector apropiado (por ejemplo, YCF3, cuya construcción se describe más adelante) para construir un vector de expresión que codifique una ZFP de plantas modificada fusionada a un dominio funcional, para la modulación de la expresión génica en células vegetales.

Ejemplo 6. Construcción de vectores para la expresión de ZFP de plantas modificadas

Se genera YCF3 como se muestra en la Figura 3. La construcción de partida era un plásmido que contiene un promotor de CMV, una secuencia de localización nuclear de SV40 (NLS), un dominio de unión a ADN de ZFP, un dominio de activación transcripcional VP 16 de Herpesvirus y una etiqueta de epítipo FLAG (pSB5186-NVF). Esta construcción se digirió con SpeI para eliminar el promotor CMV. El fragmento más grande se purificó en gel y se autiligó para crear un plásmido denominado GF1. GF1 fue digerido con KpnI y HindIII, liberando las secuencias que codifican el dominio de ZFP, el dominio de activación de VP16, y la etiqueta de epítipo FLAG, a continuación el fragmento más grande se ligó a un fragmento KpnI/HindIII que contiene las secuencias que codifican un dominio de unión de ZFP y un dominio de activación de VP16, denominado GF2. Esto dio como resultado la delección de las secuencias que codifican la etiqueta FLAG de la construcción.

GF2 se digirió con BamHI y HindIII, liberando un pequeño fragmento que codifica el dominio de activación de VP16, y el fragmento más grande se purificó y se ligó a un fragmento de PCR digerido con BamHI/HindIII que contiene el dominio de activación del maíz C1 (Goff et al. (1990) *EMBO J.* 9:2517-2522) (los sitios KpnI y HindIII se introdujeron en el fragmento de PCR por medio de cebadores que contienen los sitios KpnI y HindIII) para generar NCF1. Un fragmento de PCR que contiene un NLS de maíz opaco-2 fue digerido con SpeI/KpnI y se ligó al fragmento más grande de NCF1 digerido con KpnI/SpeI para producir YCF2. YCF2 se digirió con MluI y SpeI y el fragmento más grande se ligó a un fragmento de PCR digerido con MluI y SpeI que contiene el promotor 35S CaMV derivado de plantas (los sitios MluI y SpeI se introdujeron en el fragmento de PCR por medio de los sitios MluI o SpeI que contienen los cebadores) para generar el vector YCF3.

Las secuencias que codifican dominios ZFP de unión modificadas de plantas se pueden insertar, como fragmentos de KpnI/BamHI, en YCF3 digerido con KpnI/BamHI para generar construcciones que codifican proteínas de fusión con dominios funcionales ZFP para modular la expresión génica en células vegetales. Por ejemplo, se insertaron una serie de dominios de ZFP modificados de plantas, que se describen en el Ejemplo 5 *abajo*, en YCF3 digerido con KpnI/BamHI para generar los vectores de expresión que codifican los polipéptidos de fusión con dominio de activación de ZFP modificadas de plantas que potencian la expresión del gen GMT de *Arabidopsis thaliana*.

Ejemplo 7. Diseños de ZFP modificadas para la regulación de un gen de la metiltransferasa del tocoferol gamma (GMT) de *Arabidopsis thaliana*

Se diseñaron proteínas con dedos de zinc modificadas para reconocer diferentes secuencias diana en el gen GMT de *Arabidopsis* (Número de Acceso de GenBank AAD38271). Estas proteínas se han modificado de dos maneras. En primer lugar, aquellas que contenían un esqueleto de plantas como se describe en el Ejemplo 4. En segundo lugar, aquellas que contenían un tercer dedo de zinc no canónico (C₂HC) en el que se convierte el donde la segunda histidina de coordinación de zinc de una estructura C₂H₂ canónica se transforma en una cisteína. La Tabla 3 muestra las secuencias de nucleótidos de los diferentes sitios diana de GMT, y las secuencias de aminoácidos de los dedos de zinc que reconocen los sitios diana. Las secuencias que codifican estos dominios de unión se prepararon como se describe en el Ejemplo 4 y se insertan en YCF3 como se describe en el Ejemplo 6.

TABLA 3

ZFP #	Diana	F1	F2	F3
1	GTGGACGAGT (SEQ ID NO:39)	RSDNLAR (SEQ ID NO:40)	DRSNLTR (SEQ ID NO:41)	RSDALTR (SEQ ID NO:42)
2	CGGGATGGGT (SEQ ID NO:43)	RSDHLAR (SEQ ID NO:44)	TSGNLVR (SEQ ID NO:45)	RSDHLRE (SEQ ID NO:46)
3	TGGTGGGTGT (SEQ ID NO:47)	RSDALTR (SEQ ID NO:48)	RSDHLTT (SEQ ID NO:49)	RSDHLTT (SEQ ID NO:50)
4	GAAGAGGATT (SEQ ID NO:51)	QSSNLAR (SEQ ID NO:52)	RSDNLAR (SEQ ID NO:53)	QSGNLTR (SEQ ID NO:54)
5	GAGGAAGGGG (SEQ ID NO:55)	RSDHLAR (SEQ ID NO:56)	QSGNLAR (SEQ ID NO:57)	RSDNLTR (SEQ ID NO:58)

ZFP #	Diana	F1	F2	F3
6	TGGGTAAGTC (SEQ ID NO:59)	ERGLAR (SEQ ID NO:60)	QSGSLTR (SEQ ID NO:61)	RSDHLTT (SEQ ID NO:62)
7	GGGGAAGGG (SEQ ID NO:63)	RSDHLTQ (SEQ ID NO:64)	QSGNLAR (SEQ ID NO:65)	RSDHLR (SEQ ID NO:66)
8	GAAGAGGGTG (SEQ ID NO:67)	QSSHLAR (SEQ ID NO:68)	RSDNLAR (SEQ ID NO:69)	QSGNLAR (SEQ ID NO:70)
9	GAGGAGGATG (SEQ ID NO:71)	QSSNLQR (SEQ ID NO:72)	RSDNALR (SEQ ID NO:73)	RSDNLQR (SEQ ID NO:74)
10	GAGGAGGAGG (SEQ ID NO:75)	RSDNALR (SEQ ID NO:76)	RSDNLAR (SEQ ID NO:77)	RSDNLTR (SEQ ID NO:78)
11	GTGGCGGCTG (SEQ ID NO:79)	QSSDLRR (SEQ ID NO:80)	RSDELQR (SEQ ID NO:81)	RSDALTR (SEQ ID NO:82)
12	TGGGGAGAT (SEQ ID NO:83)	QSSNLAR (SEQ ID NO:84)	QSGHLQR (SEQ ID NO:85)	RSDHLTT (SEQ ID NO:86)
13	GAGGAAGCT (SEQ ID NO:87)	QSSDLRR (SEQ ID NO:88)	QSGNLAR (SEQ ID NO:89)	RSDNLTR (SEQ ID NO:90)
14	GCTTGTGGCT (SEQ ID NO:91)	DRSHLTR (SEQ ID NO:92)	TSGHLTT (SEQ ID NO:93)	QSSDLTR (SEQ ID NO:94)
15	GTAGTGGATG (SEQ ID NO:95)	QSSNLAR (SEQ ID NO:96)	RSDALSR (SEQ ID NO:97)	QSGSLTR (SEQ ID NO:98)
16	GTGTGGGATT (SEQ ID NO:99)	QSSNLAR (SEQ ID NO:100)	RSDHLTT (SEQ ID NO:101)	RSDALTR (SEQ ID NO:102)

Ejemplo 8 Modulación de la expresión del gen de la metiltransferasa del tocoferol gamma (GMT) de *Arabidopsis thaliana*

Se prepararon protoplastos de *Arabidopsis thaliana* y se transfectaron con plásmidos que codifican polipéptidos de fusión del dominio de activación de ZFP modificadas. Se realizó la preparación de los protoplastos y la transfección mediada por polietilenglicol como se describe. Abel *et al.* (1994) *Plant Journal* 5:421-427. Los diferentes plásmidos contenían los dominios de unión de las ZFP modificadas de plantas descritos en la Tabla 3, insertados como fragmentos de KpnI/BamHI en YCF3.

18 horas después de la transfección, se aisló el ARN a partir de los protoplastos transfectados, usando un kit de extracción de ARN de Qiagen (Valencia, CA) según las instrucciones del fabricante. Posteriormente el ARN se trató con ADNasa (libre de ARNasa), y se analizó el contenido de ARNm de GMT mediante PCR en tiempo real (TaqMan (R)). La Tabla 4 muestra las secuencias de los cebadores y sondas utilizados por análisis TaqMan®. Los resultados para los niveles de mRNA se normalizaron GMT con los niveles de rRNA 18S. Estos resultados normalizados se muestran en la Figura 4 como pliegues de activación de los niveles de mRNA de GMT, frente a los protoplastos transfectados con ADN portador (denominado "No ZFP" en la fig. 4). Los resultados indican que la expresión del gen GMT se incremento en los protoplastos que fueron transfectadas con plásmidos que codifican las fusiones entre un dominio de activación transcripcional y un dominio de unión de una ZFP modificada de plantas específica para el gen GMT.

TABLA 4

	SECUENCIA
Cebador sentido GMT	5'-AATGATCTCGCGGCTGCT-3' (SEQ ID NO: 103)
Cebador antisentido GMT	5'-GAATGGCTGATCCAACGCAT-3' (SEQ ID NO: 104)
Sonda GMT	5'-TCACTCGCTCATAAGGCTTCCTTCCAAGT-3' (SEQ ID NO: 105)
Cebador sentido 18S	5'-TGCAACAAACCCCGACTTATG-3' (SEQ ID NO: 106)
Cebador antisentido 18S	5'-CCCGCGTCGACCTTTTATC-3' (SEQ ID NO: 107)
Sonda 18S	5'-AATAAATGCGTCCCTT-3' (SEQ ID NO: 108)

Aunque los métodos y composiciones anteriores se han descrito en detalle con el fin su clara comprensión, se pueden realizar ciertas modificaciones, como es conocido por los expertos en la técnica, dentro del alcance de las reivindicaciones adjuntas.

LISTADO DE SECUENCIAS

<110> Sangamo BioSciences, Inc.

<120> PROTEÍNAS DE UNIÓN CON DEDOS DE ZINC MODIFICADAS

<130> 8325-0025.43 (S25-EP)

5 <140> 02707545.6
<141> 22-01-2002

<150> 10/055.711
<151> 22-01-2002

<160> 147

10 <170> PatentIn Ver. 2.0

<210> 1
<211> 25
<212> PRT
<213> Secuencia Artificial

15 <220>
<223> Descripción de la Secuencia Artificial: Secuencia general

<220>
<221> SITIO
<222> (2)..(5)

20 <223> donde Xaa es cualquier aminoácido

<220>
<221> SITIO
<222> (4)..(5)

<223> donde Xaa puede estar presente o ausente

25 <220>
<221> SITIO
<222> (7)..(18)

<223> donde Xaa es cualquier aminoácido

<220>
<221> SITIO
<222> (20)..(24)

30 <223> donde Xaa es cualquier aminoácido

<220>
<221> SITIO
<222> (23)..(24)

35 <223> donde Xaa puede estar presente o ausente

<400> 1

Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His

20 25

<210> 2
<211> 34
<212> PRT
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: secuencia consenso canónico

45 <220>
<221> SITIO

<222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 5 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 10 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 15 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 20 <220>
 <221> SITIO
 <222> (24) .. (29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 25 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

 <400> 2

 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

 Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
 20 25 30

 Xaa Xaa

 30 <210> 3
 <211> 9
 <212> ADN
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: secuencia diana

 35 <400> 3
 ggcgtagac 9

 <210> 4
 <211> 9
 <212> ADN
 40 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: segmento diana

 <400> 4
 ggcgacgta 9

 45 <210> 5
 <211> 5

<212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: enlazador peptídico
 5 <400> 5
 Thr Gly Glu Lys Pro
 1 5

 <210> 6
 <211> 5
 <212> PRT
 10 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: enlazador peptídico
 <400> 6
 Gly Gly Gly Gly Ser
 1 5

 15 <210> 7
 <211> 8
 <212> PRT
 <213> Secuencia Artificial

 <220>
 20 <223> Descripción de la Secuencia Artificial: enlazador peptídico
 <400> 7
 Gly Gly Arg Arg Gly Gly Gly Ser
 1 5

 <210> 8
 <211> 9
 25 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: enlazador peptídico
 <400> 8
 Leu Arg Gln Arg Asp Gly Glu Arg Pro
 30 1 5

 <210> 9
 <211> 12
 <212> PRT
 <213> Secuencia Artificial

 35 <220>
 <223> Descripción de la Secuencia Artificial: enlazador peptídico
 <400> 9
 Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro
 1 5 10

 40 <210> 10
 <211> 16
 <212> PRT
 <213> Secuencia Artificial

 <220>

<223> Descripción de la Secuencia Artificial: enlazador peptídico

<400> 10

Leu Arg Gln Lys Asp Gly Gly Gly Ser Gly Gly Gly Ser Glu Arg Pro
1 5 10 15

5 <210> 11
<211> 12
<212> PRT
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: porción C2H2 de ZFP

10 <400> 11

His Ile Lys Thr His Gln Asn Lys Lys Gly Gly Ser
1 5 10

<210> 12
<211> 12
<212> PRT
15 <213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: porción S de ZFP

<400> 12

His Ser Glu Thr Gly Cys Thr Lys Lys Gly Gly Ser
1 5 10

20 <210> 13
<211> 12
<212> PRT
<213> Secuencia Artificial

<220>
25 <223> Descripción de la Secuencia Artificial: porción E de ZFP

<400> 13

His Leu Lys Ser Leu Thr Pro Cys Thr Gly Gly Ser
1 5 10

<210> 14
<211> 12
30 <212> PRT
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: porción K de ZFP

<400> 14

His Lys Cys Gly Ile Gln Asn Lys Lys Gly Gly Ser
35 1 5 10

<210> 15
<211> 12
<212> PRT
<213> Secuencia Artificial

40 <220>
<223> Descripción de la Secuencia Artificial: porción CT de ZFP

<400> 15

His Ser Glu Asn Cys Gln Gly Lys Lys Gly Gly Ser
1 5 10

<210> 16
<211> 12
<212> PRT
5 <213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: porción C de ZFP

<400> 16

His Ile Lys Thr Cys Gln Asn Lys Lys Gly Gly Ser
1 5 10

10 <210> 17
<211> 12
<212> PRT
<213> Secuencia Artificial

<220>
15 <223> Descripción de la Secuencia Artificial: porción GC de ZFP

<400> 17

His Ile Lys Gly Cys Gln Asn Lys Lys Gly Gly Ser
1 5 10

20 <210> 18
<211> 12
<212> PRT
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: porción GGC de ZFP

<400> 18

His Ile Gly Gly Cys Gln Asn Lys Lys Gly Gly Ser
25 1 5 10

<210> 19
<211> 28
<212> PRT
<213> Secuencia Artificial

30 <220>
<223> Descripción de la Secuencia Artificial: consenso de SP-1

<220>
<221> SITIO
<222> (12)..(18)
35 <223> donde Xaa es cualquier aminoácido

<400> 19

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa His Gln Arg Thr His Thr Gly Glu Lys Pro
20 25

40 <210> 20
<211> 34
<212> PRT
<213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: esqueleto F1

<220>
 <221> SITIO
 5 <222> (18) .. (24)
 <223> donde Xaa es cualquier aminoácido

<400> 20

Lys Lys Lys Ser Lys Gly His Glu Cys Pro Ile Cys Phe Arg Val Phe
 1 5 10 15

Lys Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Lys Arg Ser His Thr Gly Glu
 20 25 30

Lys Pro

10 <210> 21
 <211> 28
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: esqueleto F2

15 <220>
 <221> SITIO
 <222> (12)..(18)
 <223> donde Xaa es cualquier aminoácido

<400> 21

Tyr Lys Cys Thr Val Cys Gly Lys Ser Phe Ser Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

Xaa Xaa His Lys Arg Leu His Thr Gly Glu Lys Pro
 20 25

<210> 22
 <211> 23
 <212> PRT
 <213> Secuencia Artificial

25 <220>
 <223> Descripción de la Secuencia Artificial: esqueleto F3

<220>
 <221> SITIO
 <222> (12)..(18)
 30 <223> donde Xaa es cualquier aminoácido

<400> 22

Phe Ser Cys Asn Tyr Cys Gln Arg Lys Phe Tyr Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

Xaa Xaa His Val Arg Ile His
 20

35 <210> 23
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: secuencia añadida a la del C-terminal de F3

<400> 23
 Gln Asn Lys Lys
 1
 <210> 24
 <211> 5
 5 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: secuencia de terminación de hélice
 <220>
 10 <221> SITIO
 <222> (4)
 <223> donde Xaa es 'Lys' o 'Arg'
 <220>
 15 <221> SITIO
 <222> (5)
 <223> donde Xaa es 'Lys' o 'Pro'
 <400> 24
 Thr Gly Glu Xaa Xaa
 1 5
 <210> 25
 20 <211> 32
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: esqueleto de dedo de zinc
 25 <220>
 <221> SITIO
 <222> (16)..(22)
 <223> donde Xaa es cualquier aminoácido
 <400> 25
 Lys Ser Lys Gly His Glu Cys Pro Ile Cys Phe Arg Val Phe Lys Xaa
 1 5 10 15
 30 Xaa Xaa Xaa Xaa Xaa Xaa His Lys Arg Ser His Thr Gly Glu Lys Pro
 20 25 30
 <210> 26
 <211> 28
 <212> PRT
 35 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: esqueleto de dedo de zinc
 <220>
 40 <221> SITIO
 <222> (12)..(18)
 <223> donde Xaa es cualquier aminoácido
 <400> 26

ES 2 409 080 T3

Tyr Lys Cys Thr Val Cys Gly Lys Ser Phe Ser Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa His Lys Arg Leu His Thr Gly Glu Lys Pro
 20 25
 <210> 27
 <211> 27
 <212> PRT
 5 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: esqueleto de dedo de zinc
 <220>
 <221> SITIO
 10 <222> (12)..(18)
 <223> donde Xaa es cualquier aminoácido
 <400> 27
 Phe Ser Cys Asn Tyr Cys Gln Arg Lys Phe Gly Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa His Val Arg Ile His Gln Asn Lys Lys
 20 25
 <210> 28
 15 <211> 60
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: oligonucleótido H1
 20 <220>
 <221> característica_misclánea
 <222> (25) .. (45)
 <223> n = a, c, g o t
 <400> 28
 25 ctcaccggtg tgagaacgct tggtnnnnnn nnnnnnnnnn nnnnnctga aaacacggaa 60
 <210> 29
 <211> 60
 <212> ADN
 <213> Secuencia Artificial
 30 <220>
 <223> Descripción de la Secuencia Artificial: oligonucleótido H2
 <220>
 <221> característica misclánea
 <222> (25) .. (45)
 35 <223> n = a, c, g o t
 <400> 29
 ttcaccagta tgaagacgct tatgtnnnnn nnnnnnnnnn nnnnnagaaa aagacttacc 60
 <210> 30
 <211> 63
 40 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: oligonucleótido H3
 <220>

ES 2 409 080 T3

<221> característica_miscelánea
 <222> (28) .. (48)
 <223> n = a, c, g o t
 <400> 30
 5 cttcttgttc tgggtggatac gcacgtgnnn nnnnnnnnnn nnnnnnnnac cgaacttacg 60
 ctg 63
 <210> 31
 <211> 45
 <212> ADN
 <213> Secuencia Artificial
 10 <220>
 <223> Descripción de la Secuencia Artificial: oligonucleótido PB1
 <400> 31
 aagtctaagg gtcacgagtg cccaatctgc ttccgtgtt tcaag 45
 <210> 32
 15 <211> 54
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: oligonucleótido PB2
 20 <400> 32
 tctcacaccg gtgagaagcc atacaagtg actgtttgtg gtaagtctt ttct 54
 <210> 33
 <211> 54
 <212> ADN
 25 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: oligonucleótido PB3
 <400> 33
 cttcatactg gtgaaaagcc attctctgc aactactgcc agcgtaagtt cggt 54
 30 <210> 34
 <211> 60
 <212> ADN
 <213> Secuencia Artificial
 <220>
 35 <223> Descripción de la Secuencia Artificial: dedo F1 para el montaje de la PCR
 <400> 34
 ctcaccggtg tgagaacgct tgtgacgggt caactcgta gaacgctga aaacacggaa 60
 <210> 35
 <211> 60
 40 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: dedo F2 para el montaje de la PCR
 <400> 35
 45 ttcaccagta tgaagacgct tatgacgggt caagtgttca gaacgagaaa aagacttacc 60
 <210> 36

ES 2 409 080 T3

<211> 63
 <212> ADN
 <213> Secuencia Artificial
 <220>
 5 <223> Descripción de la Secuencia Artificial: dedo F3 para el montaje de la PCR
 <400> 36
 cttctttgttc tgggtggatac gcacgtgacg ggtcaagttg tcagaacgac cgaacttacg 60
 ctg 63
 <210> 37
 <211> 28
 10 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: cebador PZF
 <400> 37
 15 cggggtacca ggtaagtcta agggtcac 28
 <210> 38
 <211> 34
 <212> ADN
 <213> Secuencia Artificial
 20 <220>
 <223> Descripción de la Secuencia Artificial: cebador PZR
 <400> 38
 gcgcggatcc acccttcttg ttctggtgga tacg 34
 <210> 39
 25 <211> 10
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: diana ZFP #1
 30 <400> 39
 gtggacgagt 10
 <210> 40
 <211> 7
 <212> PRT
 35 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #1 F1
 <400> 40
 Arg Ser Asp Asn Leu Ala Arg
 1 5
 40 <210> 41
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 45 <223> Descripción de la Secuencia Artificial: ZFP #1 F2
 <400> 41

Asp Arg Ser Asn Leu Thr Arg
 1 5
 <210> 42
 <211> 7
 <212> PRT
 5 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #1 F3
 <400> 42
 Arg Ser Asp Ala Leu Thr Arg
 1 5
 10 <210> 43
 <211> 10
 <212> ADN
 <213> Secuencia Artificial
 <220>
 15 <223> Descripción de la Secuencia Artificial: diana ZFP #2
 <400> 43
 cgggatgggt 10
 <210> 44
 <211> 7
 20 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #2 F1
 <400> 44
 Arg Ser Asp His Leu Ala Arg
 25 1 5
 <210> 45
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 30 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #2 F2
 <400> 45
 Thr Ser Gly Asn Leu Val Arg
 1 5
 35 <210> 46
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #2 F3
 40 <400> 46
 Arg Ser Asp His Leu Arg Glu
 1 5
 <210> 47
 <211> 10
 <212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: diana ZFP #3

<400> 47

5 tgggtgggtgt 10

<210> 48

<211> 9

<212> PRT

<213> Secuencia Artificial

10 <220>

<223> Descripción de la Secuencia Artificial: ZFP #3 F1

<400> 48

Arg Ser Asp Ala Leu Thr Arg Met Ser

1 5

15 <210> 49

<211> 7

<212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: ZFP #3 F2

20 <400> 49

Arg Ser Asp His Leu Thr Thr

1 5

<210> 50

<211> 7

<212> PRT

25 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: ZFP #3 F3

<400> 50

Arg Ser Asp His Leu Thr Thr

1 5

30 <210> 51

<211> 10

<212> ADN

<213> Secuencia Artificial

<220>

35 <223> Descripción de la Secuencia Artificial: diana ZFP #4

<400> 51

gaagaggatt 10

<210> 52

<211> 7

<212> PRT

40 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: ZFP #4 F1

<400> 52

Gln Ser Ser Asn Leu Ala Arg
1 5

<210> 53
<211> 7
<212> PRT
5 <213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: ZFP #4 F2
<400> 53

Arg Ser Asp Asn Leu Ala Arg
1 5

10 <210> 54
<211> 7
<212> PRT
<213> Secuencia Artificial

<220>
15 <223> Descripción de la Secuencia Artificial: ZFP #4 F3
<400> 54

Gln Ser Gly Asn Leu Thr Arg
1 5

20 <210> 55
<211> 10
<212> ADN
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: diana ZFP #5
<400> 55

25 gaggaagggg 10

<210> 56
<211> 7
<212> PRT
<213> Secuencia Artificial

30 <220>
<223> Descripción de la Secuencia Artificial: ZFP #5 F1
<400> 56

Arg Ser Asp His Leu Ala Arg
1 5

35 <210> 57
<211> 7
<212> PRT
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: ZFP #5 F2

40 <400> 57

Gln Ser Gly Asn Leu Ala Arg
1 5

<210> 58
<211> 7
<212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: ZFP #5 F3

<400> 58

Arg Ser Asp Asn Leu Thr Arg
 1 5

<210> 59
 <211> 9
 <212> ADN
 <213> Secuencia Artificial

10 <220>
 <223> Descripción de la Secuencia Artificial: diana ZFP #6

<400> 59

tgggtagtc

15 <210> 60
 <211> 7
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: ZFP #6 F1

20 <400> 60

Glu Arg Gly Thr Leu Ala Arg
 1 5

<210> 61
 <211> 7
 <212> PRT
 25 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: ZFP #6 F2

<400> 61

Gln Ser Gly Ser Leu Thr Arg
 1 5

30 <210> 62
 <211> 7
 <212> PRT
 <213> Secuencia Artificial

<220>
 35 <223> Descripción de la Secuencia Artificial: ZFP #6 F3

<400> 62

Arg Ser Asp His Leu Thr Thr
 1 5

<210> 63
 <211> 10
 40 <212> ADN
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: diana ZFP #7

<400> 63

9

ggggaaaggg 10

<210> 64
 <211> 7
 <212> PRT
 5 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: ZFP #7 F1

<400> 64

Arg Ser Asp His Leu Thr Gln
 1 5

10 <210> 65
 <211> 7
 <212> PRT
 <213> Secuencia Artificial

<220>
 15 <223> Descripción de la Secuencia Artificial: ZFP #7 F2

<400> 65

Gln Ser Gly Asn Leu Ala Arg
 1 5

20 <210> 66
 <211> 7
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: ZFP #7 F3

<400> 66

Arg Ser Asp His Leu Ser Arg
 1 5

25 <210> 67
 <211> 10
 <212> ADN
 <213> Secuencia Artificial

<220>
 30 <223> Descripción de la Secuencia Artificial: diana ZFP #8

<400> 67

gaagagggtg 10

35 <210> 68
 <211> 7
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: ZFP #8 F1

<400> 68

Gln Ser Ser His Leu Ala Arg
 1 5

40 <210> 69
 <211> 7
 <212> PRT
 45 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: ZFP #8 F2
 <400> 69
 Arg Ser Asp Asn Leu Ala Arg
 1 5
 5 <210> 70
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 10 <223> Descripción de la Secuencia Artificial: ZFP #8 F3
 <400> 70
 Gln Ser Gly Asn Leu Ala Arg
 1 5
 15 <210> 71
 <211> 10
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: diana ZFP #9
 <400> 71
 20 gaggaggatg 10
 <210> 72
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 25 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #9 F1
 <400> 72
 Gln Ser Ser Asn Leu Gln Arg
 1 5
 30 <210> 73
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #9 F2
 35 <400> 73
 Arg Ser Asp Asn Ala Leu Arg
 1 5
 <210> 74
 <211> 7
 <212> PRT
 40 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #9 F3
 <400> 74

	Arg	Ser	Asp	Asn	Leu	Gln	Arg	
	1				5			
	<210>	75						
	<211>	10						
	<212>	ADN						
5	<213>	Secuencia Artificial						
	<220>							
	<223>	Descripción de la Secuencia Artificial: diana ZFP #10						
	<400>	75						
	gaggaggagg							10
10	<210>	76						
	<211>	7						
	<212>	PRT						
	<213>	Secuencia Artificial						
	<220>							
15	<223>	Descripción de la Secuencia Artificial: ZFP #10 F1						
	<400>	76						
	Arg	Ser	Asp	Asn	Ala	Leu	Arg	
	1				5			
	<210>	77						
	<211>	7						
20	<212>	PRT						
	<213>	Secuencia Artificial						
	<220>							
	<223>	Descripción de la Secuencia Artificial: ZFP #10 F2						
	<400>	77						
	Arg	Ser	Asp	Asn	Leu	Ala	Arg	
25	1				5			
	<210>	78						
	<211>	7						
	<212>	PRT						
	<213>	Secuencia Artificial						
	<220>							
30	<223>	Descripción de la Secuencia Artificial: ZFP #10 F3						
	<400>	78						
	Arg	Ser	Asp	Asn	Leu	Thr	Arg	
	1				5			
	<210>	79						
35	<211>	10						
	<212>	ADN						
	<213>	Secuencia Artificial						
	<220>							
	<223>	Descripción de la Secuencia Artificial: diana ZFP #11						
40	<400>	79						
	gtggcggctg							10
	<210>	80						
	<211>	7						
	<212>	PRT						
45	<213>	Secuencia Artificial						

<220>
 <223> Descripción de la Secuencia Artificial: ZFP #11 F1
 <400> 80
 Gln Ser Ser Asp Leu Arg Arg
 1 5
 5 <210> 81
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 10 <223> Descripción de la Secuencia Artificial: ZFP #11 F2
 <400> 81
 Arg Ser Asp Glu Leu Gln Arg
 1 5
 15 <210> 82
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #11 F3
 <400> 82
 Arg Ser Asp Ala Leu Thr Arg
 20 1 5
 <210> 83
 <211> 9
 <212> ADN
 <213> Secuencia Artificial
 25 <220>
 <223> Descripción de la Secuencia Artificial: diana ZFP #12
 <400> 83
 tggggagat
 30 <210> 84
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #12 F1
 35 <400> 84
 Gln Ser Ser Asn Leu Ala Arg
 1 5
 <210> 85
 <211> 7
 <212> PRT
 40 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #12 F2
 <400> 85
 Gln Ser Gly His Leu Gln Arg
 1 5

9

<210> 86
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 5 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #12 F3
 <400> 86
 Arg Ser Asp His Leu Thr Thr
 1 5
 10 <210> 87
 <211> 9
 <212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: diana ZFP #13
 15 <400> 87
 gaggaagct
 <210> 88
 <211> 7
 <212> PRT
 20 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #13 F1
 <400> 88
 Gln Ser Ser Asp Leu Arg Arg
 1 5
 25 <210> 89
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 30 <223> Descripción de la Secuencia Artificial: ZFP #13 F2
 <400> 89
 Gln Ser Gly Asn Leu Ala Arg
 1 5
 <210> 90
 <211> 7
 35 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #13 F3
 <400> 90
 Arg Ser Asp Asn Leu Thr Arg
 40 1 5
 <210> 91
 <211> 10
 <212> ADN
 <213> Secuencia Artificial
 45 <220>

9

<223> Descripción de la Secuencia Artificial: diana ZFP #14

<400> 91

gcttggtggct 10

<210> 92

5 <211> 7

<212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: ZFP #14 F1

10 <400> 92

Asp Arg Ser His Leu Thr Arg

1 5

<210> 93

<211> 7

<212> PRT

15 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: ZFP #14 F2

<400> 93

Thr Ser Gly His Leu Thr Thr

1 5

20 <210> 94

<211> 7

<212> PRT

<213> Secuencia Artificial

<220>

25 <223> Descripción de la Secuencia Artificial: ZFP #14 F3

<400> 94

Gln Ser Ser Asp Leu Thr Arg

1 5

<210> 95

<211> 10

30 <212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: diana ZFP #15

<400> 95

35 gtagtggatg 10

<210> 96

<211> 7

<212> PRT

<213> Secuencia Artificial

40 <220>

<223> Descripción de la Secuencia Artificial: ZFP #15 F1

<400> 96

Gln Ser Ser Asn Leu Ala Arg

1 5

<210> 97
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 5 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #15 F2
 <400> 97
 Arg Ser Asp Ala Leu Ser Arg
 1 5
 <210> 98
 10 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #15 F3
 15 <400> 98
 Gln Ser Gly Ser Leu Thr Arg
 1 5
 <210> 99
 <211> 10
 <212> ADN
 20 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: diana ZFP #16
 <400> 99
 gtgtgggatt
 25 <210> 100
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 30 <223> Descripción de la Secuencia Artificial: ZFP #16 F1
 <400> 100
 Gln Ser Ser Asn Leu Ala Arg
 1 5
 <210> 101
 <211> 7
 35 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: ZFP #16 F2
 <400> 101
 Arg Ser Asp His Leu Thr Thr
 40 1 5
 <210> 102
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 45 <220>

10

<223> Descripción de la Secuencia Artificial: ZFP #16 F3

<400> 102

Arg Ser Asp Ala Leu Thr Arg
1 5

5 <210> 103
<211> 18
<212> ADN
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: cebador sentido GMT

10 <400> 103

aatgatctcg cggtgct 18

<210> 104
<211> 20
<212> ADN
15 <213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: cebador antisentido GMT

<400> 104

gaatggctga tccaacgcat 20

20 <210> 105
<211> 29
<212> ADN
<213> Secuencia Artificial

<220>
25 <223> Descripción de la Secuencia Artificial: sonda GMT

<400> 105

tcactcgctc ataaggcttc cttccaagt 29

<210> 106
<211> 21
30 <212> ADN
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: cebador sentido 18S

<400> 106

35 tgcaacaaac cccgacttat g 21

<210> 107
<211> 19
<212> ADN
<213> Secuencia Artificial

40 <220>

<223> Descripción de la Secuencia Artificial: cebador sentido 18S

<400> 107

cccgcgtcga cctttatc 19

<210> 108
45 <211> 16

<212> ADN
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: sonda 18S
 5 <400> 108
 aataaatgcg tccctt 16
 <210> 109
 <211> 7
 <212> PRT
 10 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: secuencia hélice F1
 <400> 109
 Arg Ser Asp Glu Leu Thr Arg
 1 5
 15 <210> 110
 <211> 7
 <212> PRT
 <213> Secuencia Artificial
 <220>
 20 <223> Descripción de la Secuencia Artificial: secuencia hélice F2
 <400> 110
 Arg Ser Asp His Leu Thr Arg
 1 5
 <210> 111
 <211> 7
 25 <212> PRT
 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: secuencia hélice F3
 <400> 111
 Arg Ser Asp Asn Leu Thr Arg
 30 1 5
 <210> 112
 <211> 35
 <212> PRT
 <213> Secuencia Artificial
 35 <220>
 <223> Descripción de la Secuencia Artificial: PTP2 F1
 <400> 112
 Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly Lys
 1 5 10 15
 Val Tyr Gly Arg Ser Asp Glu Leu Thr Arg His Leu Arg Trp His Thr
 20 25 30
 Gly Glu Arg
 35
 <210> 113

<211> 30
 <212> PRT
 <213> Secuencia Artificial

5 <220>
 <223> Descripción de la Secuencia Artificial: PTP2 F2

<400> 113

Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg Phe Thr Arg Ser
 1 5 10 15

Asp His Leu Thr Arg His Lys Arg Thr His Thr Gly Glu Lys
 20 25 30

<210> 114
 <211> 31
 10 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: PTP2 F3

<400> 114

Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg Ser Asp Asn
 1 5 10 15

Leu Thr Arg His Ile Lys Thr His Gln Asn Lys Lys Gly Gly Ser
 15 20 25 30

<210> 115
 <211> 35
 <212> PRT
 <213> Secuencia Artificial

20 <220>
 <223> Descripción de la Secuencia Artificial: PTP2(H->C) F1

<400> 115

Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly Lys
 1 5 10 15

Val Tyr Gly Arg Ser Asp Glu Leu Thr Arg His Leu Arg Trp His Thr
 20 25 30

Gly Glu Arg
 35

25 <210> 116
 <211> 30
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: PTP2(H->C) F2

30 <400> 116

Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg Phe Thr Arg Ser
 1 5 10 15

Asp His Leu Thr Arg His Lys Arg Thr His Thr Gly Glu Lys
 20 25 30

<210> 117

<211> 31
 <212> PRT
 <213> Secuencia Artificial

 <220>
 5 <223> Descripción de la Secuencia Artificial: PTP2(H->C) F3
 <400> 117

 Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg Ser Asp Asn
 1 5 10 15

 Leu Thr Arg His Ile Gly Gly Cys Gln Asn Lys Lys Gly Gly Ser
 20 25 30

 <210> 118
 <211> 34
 10 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 <220>
 15 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 <220>
 20 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido exdepto cisteína

 <220>
 25 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 30 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 <220>
 35 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 <220>
 40 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 45 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

 <400> 118

ES 2 409 080 T3

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
20 25 30

Xaa Xaa

<210> 119

<211> 34

<212> PRT

5 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>

<221> SITIO

10 <222> (1)..(3)

<223> donde Xaa es cualquier aminoácido

<220>

<221> SITIO

<222> (5)..(8)

15 <223> donde Xaa es cualquier aminoácido

<220>

<221> SITIO

<222> (7)..(8)

<223> donde Xaa puede estar presente o ausente

20 <220>

<221> SITIO

<222> (9)

<223> donde Xaa es cualquier aminoácido excepto cisteína

<220>

25 <221> SITIO

<222> (10)..(21)

<223> donde Xaa es cualquier aminoácido

<220>

<221> SITIO

30 <222> (23)..(29)

<223> donde Xaa es cualquier aminoácido

<220>

<221> SITIO

<222> (24)..(29)

35 <223> donde Xaa puede estar presente o ausente

<220>

<221> SITIO

<222> (31)..(34)

<223> donde Xaa es cualquier aminoácido

40 <400> 119

Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
20 25 30

Xaa Xaa

<210> 120

<211> 34
 <212> PRT
 <213> Secuencia Artificial

5 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

10 <220>
 <221> SITIO
 <222> (5) .. (8)
 <223> donde Xaa es cualquier aminoácido

15 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

20 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

25 <220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina

30 <220>
 <221> SITIO
 <222> (24) .. (29)
 <223> donde Xaa puede estar presente o ausente

35 <220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 120

Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5				10					15		

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	His	Xaa	Xaa
			20					25					30		

Xaa	Xaa
-----	-----

40 <210> 121
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

45 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO

<222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 5 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 10 <222> (7) .. (8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 <222> (10) .. (21)
 <223> donde Xaa es cualquier aminoácido

 15 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 20 <220>
 <221> SITIO
 <222> (24) .. (29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 25 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 <220>
 <221> SITIO
 <222> (31)..(34)
 30 <223> donde Xaa es cualquier aminoácido

 <400> 121
 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

 Xaa Xaa
 <210> 122
 <211> 34
 35 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 <220>
 40 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 45 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 <220>
 <221> SITIO

<222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 5 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 10 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 <220>
 <221> SITIO
 15 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 <220>
 20 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 <220>
 25 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 25 <221> SITIO
 <222> (31) .. (34)
 <223> donde Xaa es cualquier aminoácido

 <400> 122

 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
 20 25 30

 Xaa Xaa
 30 <210> 123
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

 <220>
 35 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 <220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 40 <220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 <220>
 45 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO

<222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 5 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 10 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 <220>
 <221> SITIO
 <222> (23) .. (29)
 <223> donde Xaa es cualquier aminoácido

 15 <220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

 20 <220>
 <221> SITIO
 <222> (31) .. (34)
 <223> donde Xaa es cualquier aminoácido

 <400> 123

 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
 20 25 30

 Xaa Xaa
 25 <210> 124
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

 <220>
 30 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 <220>
 <221> SITIO
 <222> (1) .. (3)
 <223> donde Xaa es cualquier aminoácido

 35 <220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 <220>
 40 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 45 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO

<222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 5 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 <222> (24) .. (29)
 10 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 15 <220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

 <400> 124

 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

 Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

 20 Xaa Xaa
 <210> 125
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

 25 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 <220>
 <221> SITIO
 <222> (1)..(3)
 30 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 35 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 40 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 <220>
 <221> SITIO
 45 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO

<222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 <220>
 <221> SITIO
 5 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 10 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

 15 <400> 125

 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
 20 25 30

 Xaa Xaa

 <210> 126
 <211> 34
 <212> PRT
 20 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 <220>
 <221> SITIO
 25 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 <222> (5)..(8)
 30 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 35 <220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 <220>
 40 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 45 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO

<222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 5 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 <220>
 <221> SITIO
 10 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

 <400> 126

 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

 Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

 Xaa Xaa

 <210> 127
 <211> 34
 15 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 <220>
 20 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 <220>
 25 <221> SITIO
 <222> (5) .. (8)
 <223> donde Xaa es cualquier aminoácido

 <220>
 30 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 35 <220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 <220>
 40 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquiera aminoácido

 <220>
 45 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO

	<222>	(30)
	<223>	donde Xaa es cualquier amoniácido excepto histidina
	<220>	
5	<221>	SITIO
	<222>	(31)..(34)
	<223>	donde Xaa es cualquier aminoácido
	<400>	127
	Xaa	Xaa Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
	1	5 10 15
	Xaa	Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
		20 25 30
	Xaa	Xaa
10	<210>	128
	<211>	34
	<212>	PRT
	<213>	Secuencia Artificial
	<220>	
	<223>	Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico
15	<220>	
	<221>	SITIO
	<222>	(1)..(3)
	<223>	donde Xaa es cualquier aminoácido
	<220>	
20	<221>	SITIO
	<222>	(5)..(8)
	<223>	donde Xaa es cualquier amoniácido
	<220>	
	<221>	SITIO
25	<222>	(9)
	<223>	donde Xaa es cualquier aminoácido excepto cisteína
	<220>	
	<221>	SITIO
	<222>	(7)..(8)
30	<223>	donde Xaa puede estar presente o ausente
	<220>	
	<221>	SITIO
	<222>	(10)..(21)
	<223>	donde Xaa es cualquier aminoácido
35	<220>	
	<221>	SITIO
	<222>	(22)
	<223>	donde Xaa es cualquier aminoácido excepto histidina
	<220>	
40	<221>	SITIO
	<222>	(23) .. (29)
	<223>	donde Xaa es cualquier aminoácido
	<220>	
	<221>	SITIO
45	<222>	(24)..(29)
	<223>	donde Xaa puede estar presente o ausente
	<220>	
	<221>	SITIO

<222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 <220>
 <221> SITIO
 5 <222> (31) .. (34)
 <223> donde Xaa es cualquier aminoácido

 <400> 128

 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

 Xaa Xaa

 10 <210> 129
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 15 <220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 20 <220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 25 <220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 30 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 35 <220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 40 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 45 <220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO

<222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina

<220>
 <221> SITIO
 5 <222> (31) .. (34)
 <223> donde Xaa es cualquier aminoácido

<400> 129

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10					15		

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			20					25						30		

Xaa Xaa

10 <210> 130
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

15 <220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

20 <220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

25 <220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

30 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

35 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

40 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

45 <220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO

<222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 <220>
 <221> SITIO
 5 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

 <400> 130

 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

 Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

 Xaa Xaa

 10 <210> 131
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

 15 <220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 20 <220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 25 <220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 30 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 35 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 40 <220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 45 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO

<222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 5 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

 <400> 131

 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
 20 25 30

 Xaa Xaa
 10 <210> 132
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico
 15 <220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

 20 <220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto cisteína

 25 <220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

 30 <220>
 <221> SITIO
 <222> (7) .. (8)
 <223> donde Xaa puede estar presente o ausente

 <220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto cisteína
 35 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

 40 <220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina

 45 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

 <220>
 <221> SITIO
 <222> (24) .. (29)

<223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (30)
 5 <223> donde Xaa es cualquier aminoácido excepto histidina

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

10 <400> 132

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10					15	

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			20					25					30		

Xaa Xaa

<210> 133
 <211> 34
 <212> PRT
 15 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 20 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (4)
 25 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

30 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (10)..(21)
 35 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (23)..(29)
 40 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (29) .. (29)
 45 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (31)..(34)

<223> donde Xaa es cualquier aminoácido

<400> 133

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
20 25 30

Xaa Xaa

<210> 134

5 <211> 34

<212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

10 <220>

<221> SITIO

<222> (1)..(3)

<223> donde Xaa es cualquier aminoácido

<220>

15 <221> SITIO

<222> (5)..(8)

<223> donde Xaa es cualquier aminoácido

<220>

20 <221> SITIO

<222> (7)..(8)

<223> donde Xaa puede estar presente o ausente

<220>

25 <221> SITIO

<222> (9)

<223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

<220>

<221> SITIO

<222> (10)..(21)

<223> donde Xaa es cualquier aminoácido

30 <220>

<221> SITIO

<222> (23)..(29)

<223> donde Xaa es cualquier aminoácido

<220>

35 <221> SITIO

<222> (24)..(29)

<223> donde Xaa puede estar presente o ausente

<220>

40 <221> SITIO

<222> (31)..(34)

<223> donde Xaa es cualquier aminoácido

<400> 134

Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
20 25 30

Xaa Xaa

<210> 135
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

5 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 10 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

15 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

20 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

25 <220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

30 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

35 <220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 135
 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 40 1 5 10 15
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
 20 25 30
 Xaa Xaa

<210> 136
 <211> 34
 <212> PRT
 45 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO

	<222>	(1)..(3)
	<223>	donde Xaa es cualquier aminoácido
	<220>	
5	<221>	SITIO
	<222>	(5)..(8)
	<223>	donde Xaa es cualquier aminoácido
	<220>	
	<221>	SITIO
10	<222>	(7)..(8)
	<223>	donde Xaa puede estar presente o ausente
	<220>	
	<221>	SITIO
	<222>	(10)..(21)
	<223>	donde Xaa es cualquier aminoácido
15	<220>	
	<221>	SITIO
	<222>	(23)..(29)
	<223>	donde Xaa es cualquier aminoácido
20	<220>	
	<221>	SITIO
	<222>	(24)..(29)
	<223>	donde Xaa puede estar presente o ausente
	<220>	
25	<221>	SITIO
	<222>	(30)
	<223>	donde Xaa es cualquier aminoácido excepto histidina o cisteína
	<220>	
	<221>	SITIO
30	<222>	(31) .. (34)
	<223>	donde Xaa es cualquier aminoácido
	<400>	136
	Xaa	Xaa Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
	1	5 10 15
	Xaa	Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
	20	25 30
	Xaa	Xaa
	.	
35	<210>	137
	<211>	34
	<212>	PRT
	<213>	Secuencia Artificial
	<220>	
	<223>	Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico
	<220>	
40	<221>	SITIO
	<222>	(1)..(3)
	<223>	donde Xaa es cualquier aminoácido
	<220>	
	<221>	SITIO
45	<222>	(4)
	<223>	donde Xaa es cualquier aminoácido excepto histidina o cisteína

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

5 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

10 <220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

15 <220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

20 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

25 <220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 137

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10					15	

Xaa	Xaa	Xaa	Xaa	Xaa	His	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	His	Xaa	Xaa
			20					25					30		

30 Xaa Xaa

<210> 138
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

35 <220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

40 <220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

45 <220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

5

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

10

<220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

15

<220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

20

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

25

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 138
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
 20 25 30
 Xaa Xaa
 <210> 139
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

30

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

35

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

40

<220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

45

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

5

<220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

10

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

15

<220>
 <221> SITIO
 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

20

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 139

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15

Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

Xaa Xaa

25

<210> 140
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

30

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

35

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

40

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

45

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

5 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

10 <220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

15 <220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 140

Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5				10							15	

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	His	Xaa	Xaa
			20					25					30		

Xaa Xaa

20 <210> 141
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

25 <220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

30 <220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

35 <220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

40 <220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

45 <220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

5

<220>
 <221> SITIO
 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

10

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 141
 Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa Xaa Xaa Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30
 Xaa Xaa

15

<210> 142
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

20

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

25

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

30

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

35

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

40

<220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

<220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

45

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

5

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 142

Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10					15	

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			20					25						30	

10

Xaa Xaa
 <210> 143
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

15

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

20

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

25

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

30

<220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

35

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

40

<220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

45

<220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

5

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 143

	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
10	1				5					10					15	
	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
				20				25					30			

Xaa Xaa

<210> 144
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

15

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

20

<220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

25

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

30

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

35

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

40

<220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

45

<220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

5

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 144

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
.1				5				10					15		

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			20					25					30		

10

Xaa Xaa
 <210> 145
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

15

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)

20

<223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

25

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

30

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

35

<220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

<220>
 <221> SITIO
 <222> (10)..(21)

40

<223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

45

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

5

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 145

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10						15	

Xaa	Xaa	Xaa	Xaa	Xaa	His	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
			20					25						30		

10

Xaa Xaa
 <210> 146
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

15

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

20

<220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

25

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

30

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

35

<220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

40

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

45

<220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

5

<220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 146

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10						15	

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	His	Xaa	Xaa
			20					25						30		

10

Xaa Xaa
 <210> 147
 <211> 34
 <212> PRT
 <213> Secuencia Artificial

15

<220>
 <223> Descripción de la Secuencia Artificial: componente de dedos de zinc no-canónico

<220>
 <221> SITIO
 <222> (1)..(3)
 <223> donde Xaa es cualquier aminoácido

20

<220>
 <221> SITIO
 <222> (4)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

25

<220>
 <221> SITIO
 <222> (5)..(8)
 <223> donde Xaa es cualquier aminoácido

30

<220>
 <221> SITIO
 <222> (7)..(8)
 <223> donde Xaa puede estar presente o ausente

<220>
 <221> SITIO
 <222> (9)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

35

<220>
 <221> SITIO
 <222> (10)..(21)
 <223> donde Xaa es cualquier aminoácido

40

<220>
 <221> SITIO
 <222> (22)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

45

<220>
 <221> SITIO
 <222> (23)..(29)
 <223> donde Xaa es cualquier aminoácido

<220>
 <221> SITIO
 <222> (24)..(29)
 <223> donde Xaa puede estar presente o ausente

5 <220>
 <221> SITIO
 <222> (30)
 <223> donde Xaa es cualquier aminoácido excepto histidina o cisteína

10 <220>
 <221> SITIO
 <222> (31)..(34)
 <223> donde Xaa es cualquier aminoácido

<400> 147
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1 5 10 15
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

15 Xaa Xaa

REIVINDICACIONES

1. Una proteína de unión a ADN con dedos de zinc aislada que ha sido modificada genéticamente para unirse a una secuencia de ADN diana, donde
 - (i) la proteína de unión a ADN con dedos de zinc comprende 3 dedos de zinc,
 - 5 (ii) los dedos de zinc primero y segundo son de la clase C2H2 y el tercer dedo de zinc es un dedo de zinc no canónico que comprende la secuencia $X_3\text{-Cys-X}_{2-4}\text{-Cys-X}_{12}\text{-His-X}_{1-7}\text{-Cys-X}_4$ donde X es cualquier aminoácido y donde X_{12} comprende una región de la hélice de reconocimiento que se une al sitio de diana; y
 - (iii) la región de la hélice de reconocimiento del dedo de zinc no canónico está modificada genéticamente para unirse al sitio diana.
- 10 2. La proteína de unión a ADN con dedos de zinc aislada de la reivindicación 1, donde la secuencia diana está en una célula vegetal.
3. La proteína de unión a ADN con dedos de zinc aislada de la reivindicación 1, donde la secuencia diana está en una célula animal.
- 15 4. La proteína de unión a ADN con dedos de zinc aislada de la reivindicación 1, donde la secuencia diana está en una célula humana.
5. La proteína de unión a ADN con dedos de zinc aislada de la reivindicación 1, donde la secuencia diana es una secuencia promotora.
6. La proteína de unión a ADN con dedos de zinc aislada de la reivindicación 1, que consiste en tres dedos de zinc.
7. La proteína de unión a ADN con dedos de zinc aislada de la reivindicación 1, donde la secuencia diana comprende aproximadamente de 9 a aproximadamente 14 pares de bases contiguos.
- 20 8. Un polipéptido aislado que codifica proteína de unión a ADN con dedos de zinc según la reivindicación 1.
9. Un vector de expresión que comprende el polinucleótido de la reivindicación 8.
10. Una célula hospedadora que comprende el polinucleótido de la reivindicación 8.
11. Un polipéptido de fusión que comprende:
 - 25 (a) Una proteína de unión a ADN con dedos de zinc aislada que ha sido modificada para unirse a una secuencia diana de ADN, donde
 - (i) la proteína de unión a ADN con dedos de zinc comprende 3 dedos de zinc,
 - (ii) los dedos de zinc primero y segundo son de la clase C2H2 y el tercer dedo de zinc es un dedo de zinc no canónico que comprende la secuencia $X_3\text{-Cys-X}_{2-4}\text{-Cys-X}_{12}\text{-His-X}_{1-7}\text{-Cys-X}_4$, donde X es cualquier aminoácido y donde X_{12} comprende una región de la hélice de reconocimiento que se une al sitio de diana; y
 - 30 (iii) la región de la hélice de reconocimiento del dedo de zinc no canónico está modificada para unirse al sitio diana; y
 - (b) por lo menos un dominio funcional
- 35 12. El polipéptido de fusión de la reivindicación 11, donde el dominio funcional es un dominio represivo.
13. El polipéptido de fusión de la reivindicación 12, donde el dominio funcional es ROM2 o AtD2A.
14. El polipéptido de fusión de la reivindicación 11, donde el dominio funcional es un dominio de activación.
15. El polipéptido de fusión de la reivindicación 14, donde el dominio de activación se selecciona del grupo que consiste en PVALF, ERF-2, OsGA1, HALF-1, C1, AP1, ARF-5, ARF-6, ARF-7, ARF-8, CPRF-1, CPRF4, MYC-RP/GP y TRAB1.
- 40 16. El polipéptido de fusión de la reivindicación 11, donde el dominio funcional se selecciona del grupo que consiste en un dominio aislante, una proteína remodeladora de la cromatina o un dominio de unión a metilo.
17. Una célula que comprende el polipéptido de fusión de la reivindicación 11.
18. Un polinucleótido aislado que codifica el polipéptido de fusión de la reivindicación 11.

19. Un vector de expresión que comprende el polinucleótido de la reivindicación 18.
20. Una célula hospedadora que comprende el polinucleótido de la reivindicación 18.
21. Un método para modular la expresión de un gen dónde dicho método comprende la etapa de contactar una región de ADN con una molécula de fusión según la reivindicación 11.
- 5 22. El método de la reivindicación 21, donde el gene está en una célula vegetal.
23. El método de la reivindicación 22, donde la proteína de unión a ADN con dedos de zinc de la molécula de fusión se une a un sitio diana en un gen que codifica un producto seleccionado del grupo que consiste en FAD2-1, desaturasa delta-9, desaturasa delta-12, desaturasa delta-15, acetil-CoA carboxilasa, acil-ACP tioesterasa, ADP glucosa pirofosforilasa, almidón sintasa, celulosa sintasa, sacarosa sintasa, liasa de hidroperóxidos de ácidos grasos, poligalacturonasa, y EPSP sintasa.
- 10 24. El método de la reivindicación 21, donde el gen está en una célula animal.
25. El método de la reivindicación 24, donde el gen está en una célula humana.
26. Una composición farmacéutica que comprende una proteína con dedos de zinc no canónica según cualquiera de las reivindicaciones 1 o 11 y un excipiente aceptable farmacéuticamente.
- 15 27. Una composición que comprende un polinucleótido según cualquiera de las reivindicaciones 8 o 18 y un excipiente aceptable farmacéuticamente.
28. Un método para producir una proteína de unión a ADN con dedos de zinc de la reivindicación 1 o un polipéptido de fusión de la reivindicación 11, que comprende las etapas
- 20 a) introducir una ácido nucléico que codifica una proteína de unión a ADN con dedos de zinc de la reivindicación 1 o el polipéptido de fusión de la reivindicación 11 dentro de la célula hospedadora, y
- b) expresar la proteína de unión a ADN con dedos de zinc de la reivindicación 1 o el polipéptido de fusión de la reivindicación 11 en la célula hospedadora.
29. El método de la reivindicación 28 comprende además purificar la proteína de unión a ADN con dedos de zinc de la reivindicación 1 o el polipéptido de fusión de la reivindicación 11.

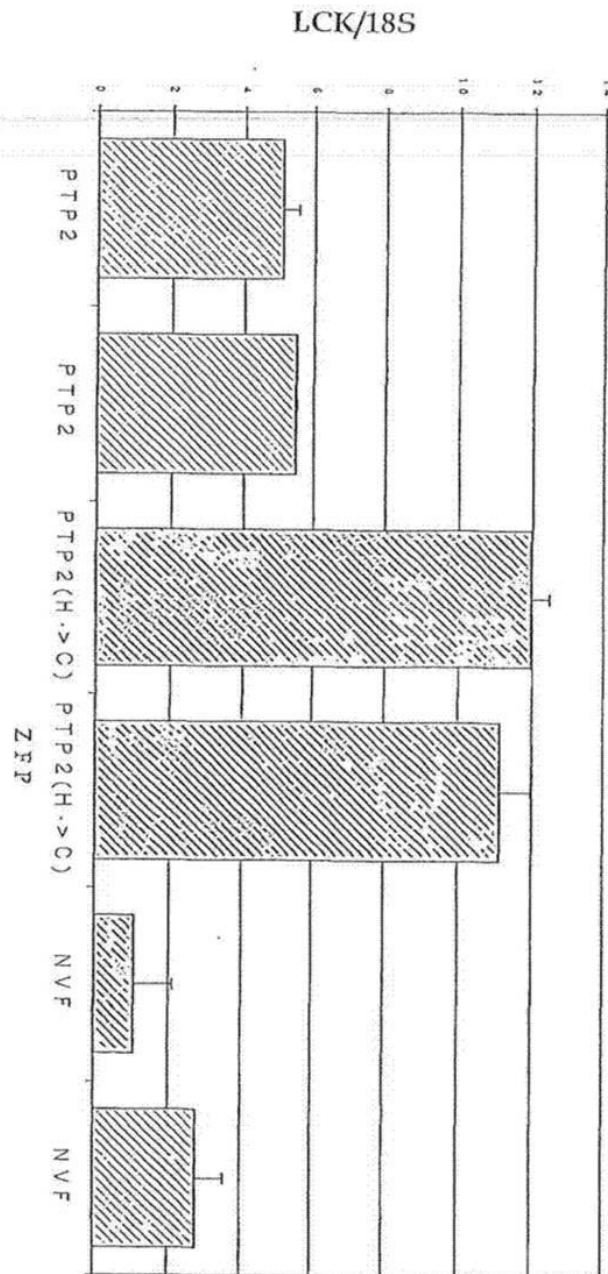


FIGURA 1

FIGURA 2

