



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 341 926**

51 Int. Cl.:
C12P 21/06 (2006.01)
C07K 14/00 (2006.01)
C07H 17/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **99909701 .7**
96 Fecha de presentación : **01.03.1999**
97 Número de publicación de la solicitud: **1060261**
97 Fecha de publicación de la solicitud: **20.12.2000**

54 Título: **Poliproteínas con dedos de cinc que tienen enlazadores mejorados.**

30 Prioridad: **02.03.1998 US 76454 P**

45 Fecha de publicación de la mención BOPI:
29.06.2010

45 Fecha de la publicación del folleto de la patente:
29.06.2010

73 Titular/es:
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
5 Cambridge Center
Kendall Square, Room Ne25-230
Cambridge, Massachusetts 02142-1324, US

72 Inventor/es: **Kim, Jin-Soo**
Pabo, Carl, O.

74 Agente: **Lehmann Novo, María Isabel**

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

Poliproteínas con dedos de cinc que tienen enlazadores mejorados.

5 Antecedentes de la invención

Los dedos de cinc pertenecientes a la familia Cys₂-His₂ constituyen uno de los motivos de fijación de DNA más comunes encontrados en los eucariotas, y estos dedos de cinc han proporcionado un marco muy atractivo para el diseño y la selección de proteínas de fijación de DNA con nuevas especificidades de secuencia. Numerosos estudios han utilizado métodos de presentación de fago o ideas de diseño para explorar y alterar sistemáticamente la especificidad de las interacciones dedo de cinc-DNA (Desjarlais & Berg, *Proteins Struct. Funct. Genet.* 12:101-104 (1992); Desjarlais & Berg, *Proc. Natl. Acad. Sci. USA* 90:2256-2260 (1993); Rebar & Pabo, *Science* 263:671-673 (1994); Jamieson *et al.*, *Biochemistry* 33:5689-5695 (1994); Choo & Klug, *Proc. Natl. Acad. Sci. USA* 91:11163-11167 (1994); Wu *et al.*, *Proc. Natl. Acad. Sci. USA* 92:344-348 (1995); y Greisman & Pabo, *Science* 275:657-661 (1997)).

El diseño por computadora basado en estructura ha sido utilizado para enlazar dedos de cinc Cys₂-His₂ con otros dominios de fijación de DNA, que incluyen otras proteínas con dedos de cinc, para generar proteínas híbridas que reconocen sitios extendidos (Pomerantz *et al.*, *Science* 267: 93-96 (1995); Kim *et al.*, *Proc. Natl. Acad. Sci. USA* 94: 3616-3620 (1997)). Por ejemplo, proteínas con dedos de cinc se han enlazado a un dominio de dimerización GAL4 para desarrollar nuevos homo- y hetero-dímeros (Pomerantz *et al.*, *Biochemistry* 4: 965-970 (1997)), y a un dominio de nucleasa para generar nuevas enzimas de restricción (Kim *et al.*, *Proc. Natl. Acad. Sci. USA* 93: 1156-1160 (1996)). La fusión dedo de cinc/homeodominio está siendo testada respecto a aplicaciones potenciales en terapia génica (Rivera *et al.*, *Nature Med.* 2: 1028-1032 (1996)).

Ha habido también varios intentos de aumentar la afinidad y especificidad de las proteínas con dedos de cinc por adición con dedos adicionales a una proteína de tres dedos (Rebar, (Tesis Ph.D.), *Selection Studies of Zinc Finger-NA Recognition*, Massachusetts Institute of Technology (1997); Shi, Y. (Tesis Ph.D.) *Molecular Mechanisms of Zinc Finger Protein-Nucleic Acid Interactions*, Johns Hopkins University (1995)) o por enlace en tándem de dos proteínas de tres dedos (Liu *et al.*, *Proc. Natl. Acad. Sci. USA* 94:5525-5530 (1997)). Sin embargo, estas estrategias de diseño previas para proteínas con dedos múltiples, todas las cuales utilizaban enlazadores "TGEKP" canónicos (enlazadores que tienen la secuencia de aminoácidos treonina-glicina-glutamato-lisina-prolina) para conectar los dedos adicionales, daban como resultado aumentos de afinidad relativamente moderados. Existe por consiguiente necesidad de desarrollar enlazadores que proporcionen afinidad y especificidad incrementadas a proteínas con dedos de cinc quiméricas.

35 Sumario de la invención

La presente invención proporciona por tanto un método de utilización de diseño basado en la estructura para seleccionar enlazadores flexibles y producir proteínas con dedos de cinc con afinidad y especificidad mejoradas. La presente invención proporciona también un método para producir proteínas con dedos de cinc de acuerdo con las reivindicaciones con afinidad y especificidad mejoradas. Las proteínas con dedos de cinc producidas utilizando estos métodos tienen afinidades de fijación en el intervalo femtomolar y proporcionan, v.g., niveles altos (mayores que aproximadamente 70 veces) de represión de la transcripción en un solo sitio diana. Tales proteínas con dedos de cinc pueden utilizarse para regulación de la expresión génica, v.g., como agentes terapéuticos, diagnósticos, y para aplicaciones de investigación tales como genómica funcional.

En un aspecto, la presente invención proporciona un método de producción de una proteína con dedos de cinc que se fija a sitios diana adyacentes de acuerdo con la reivindicación 1.

En otro aspecto, la presente invención proporciona una proteína con dedos de cinc de acuerdo con la reivindicación 9.

En una realización, la presente invención proporciona ácidos nucleicos que codifican las proteínas con dedos de cinc.

En una realización, los dominios primero y segundo son polipéptidos con dedos de cinc. En otra realización, el polipéptido con dedos de cinc se selecciona del grupo constituido por Zif268 y NRE. En otra realización, los polipéptidos con dedos de cinc son heterólogos.

En otra realización, la proteína con dedos de cinc comprende adicionalmente un polipéptido de dominio regulador.

En una realización, la proteína con dedos de cinc tiene afinidad femtomolar para los sitios diana adyacentes. En otra realización, la proteína con dedos de cinc tiene aproximadamente una afinidad 2-4 femtomolar para los sitios diana adyacentes.

En una realización, el enlazador flexible tiene una longitud de 8 u 11 aminoácidos. En otra realización, el enlazador flexible tiene la secuencia RQKDGERP o RQKDGGGSERP.

En una realización, los sitios diana están separados por uno o dos nucleótidos.

En una realización, los sitios diana adyacentes están separados por cero nucleótidos y el enlazador flexible tiene una longitud de 6 aminoácidos. En otra realización, los sitios diana adyacentes están separados por un solo nucleótido y el enlazador flexible tiene una longitud de siete, ocho o nueve aminoácidos. En otra realización, los sitios diana adyacentes están separados por dos nucleótidos y el enlazador flexible tiene una longitud de diez, once o doce aminoácidos. En otra realización, los sitios diana adyacentes están separados por tres nucleótidos y el enlazador flexible tiene una longitud de 12 aminoácidos.

Breve descripción de los dibujos

La Figura 1 representa el diseño basado en la estructura de un péptido con 6 dedos, 268//NRE. La estructura cocristalina del complejo Zif268-DNA y el molde B-DNA (utilizado en la unión) se alinearon mediante fosfatos superpuestos (Pavletich & Pabo, *Science* 252:809817 (1991); Elrod-Erickson *et al.*, *Structure* 4:1171-1180 (1996)). En este modelo, dos péptidos de 3 dedos se fijan a sitios correspondientes de 9 pares de bases (las bases se representan en color blanco) separados por una laguna de 2 pb (las bases se representan en color gris). Obsérvese que la orientación de un péptido de 3 dedos coincide casi exactamente con la del otro péptido de 3 dedos debido a que una vuelta helicoidal de este DNA subenrollado contiene 11 pares de bases.

La Figura 2 muestra representaciones esquemáticas de péptidos con dedos de cinc y de constructos informadores utilizados en los estudios de transfección descritos en esta memoria. La Figura 2A muestra péptidos con dedos de cinc. Cada dedo está representado con un círculo. Se muestra la secuencia de aminoácidos de un enlazador en el péptido Zif268 (que tiene un enlazador canónico "TGEKP", y los enlazadores más largos utilizados para conectar los péptidos de 3 dedos se indican al pie. En cada caso, el recuadro de la izquierda denota la región helicoidal e incluye el segundo de los residuos His conservados del dedo; la línea en zigzag denota la primera hoja β del dedo siguiente, que incluye el primero de los residuos Cys conservados. La Figura 2B ilustra promotores de genes informadores de luciferasa. Las posiciones de los nucleótidos de la secuencia TATA, el codón inicial, y los sitios de fijación de los dedos de cinc se enumeran con respecto al sitio de comienzo de la transcripción (+1).

La figura 3 representa un ensayo de desplazamiento de gel. Se incubaron diversas cantidades (0, 0,01, 0,1, y 1 nM) del péptido NRE durante 1 hora con sitios de fijación libres (pistas 1-4) o sitios de fijación preincubados con 0,1 nM del péptido Zif268 durante 0,5 horas (pistas 5-8). Se indican las posiciones del DNA libre y de los complejos proteína-DNA.

La Figura 4 muestra estudios de fijación de competición. La Figura 4A, el péptido 268/NRE (5 pM) se preincubó con diversas cantidades (0,05, 0,5, 5 y 50 nM) de DNAs competidores fríos (pistas 3-14) durante 1 hora, y se añadió luego un ligero exceso molar (sobre la concentración de péptido) del sitio N/Z marcado (608 pM) a la mezcla de reacción. Se analizaron partes alícuotas por electroforesis en gel en diversos momentos, y este gel muestra los resultados al cabo de 600 horas de tiempo de incubación a la temperatura ambiente. En la Figura 4B, el péptido 268//NRE (pistas 2-6) o el péptido Zif268 (pistas 7-11) se mezcló con el sitio N/Z marcado, se añadió un ligero exceso molar (sobre la concentración de péptido) de sitio N/Z sin marcar (de tal modo que el 70% del sitio marcado se desplazaría en ausencia de DNA de esperma de salmón), y se incluyeron diversas cantidades de DNA de esperma de salmón (0, 0,1, 1, 10, y 100 μ g/ml). Las muestras se analizaron por electroforesis en gel después de 24 horas de incubación.

La Figura 5 muestra gráficos (Figuras 5A, 5B, 5C y 5D) que ilustran la represión de la transcripción *in vivo* por péptidos con dedos de cinc. Se transfectaron células 293 humanas como ha sido descrito (Cepek *et al.*, *Genes Dev.* 10: 2079-2088 (1996)) utilizando el método de precipitación con fosfato de calcio. Se midieron las actividades de luciferasa y β -galactosidasa 48 horas más tarde. Las actividades de luciferasa se dividieron por las actividades de β -galactosidasa correspondientes para proporcionar las actividades de luciferasa relativas. Los niveles de represión (múltiplos de represión) se obtuvieron dividiendo 1) las actividades relativas de luciferasa de las células transfectadas con el plásmido de expresión vacío por 2) las de las células transfectadas con los plásmidos de expresión con dedos de cinc. Se utilizan escalas diferentes en los gráficos para los diferentes informadores. Los péptidos 68/NR, 68/NRE, 68//NR, y 68//NRE son variantes de proteínas de fusión con 6 dedos a las que les faltan uno o dos de los dedos terminales. Así, el péptido 68/NR contiene los dedos 2 y 3 del péptido Zif268 fusionados (por el más corto de los dos enlazadores) a los dedos 1 y 2 del péptido NRE. Los datos representan un promedio de tres experimentos independientes, y se muestra el error estándar del valor medio.

Descripción detallada de la invención

I. Introducción

La presente invención proporciona una estrategia de diseño para enlazadores que fusionan dos dominios de fijación de DNA de una proteína con dedos de cinc. Estos enlazadores son flexibles y más largos que los enlazadores canónicos utilizados previamente, permitiendo la fijación de la proteína con dedos de cinc a su sitio diana sin introducción de tensión alguna. El sitio diana es típicamente un sitio diana "compuesto", que se compone de dos sitios diana adyacentes que están separados por cero a tres nucleótidos. Cada uno de los sitios diana adyacentes es reconocido por un solo dominio de fijación de DNA de la proteína con dedos de cinc. La estrategia de diseño de enlazadores implica diseño basado en estructura para determinar una longitud mínima de un enlazador entre dos dominios de fijación de DNA,

y adición subsiguiente de aminoácidos adicionales al enlazador para proporcionar al menos aproximadamente 1-2 angstroms adicionales de flexibilidad al enlazador. La presente invención proporciona por tanto proteínas con dedos de cinc que tienen afinidad femtomolar para su sitio diana, y que reprimen eficazmente la expresión génica, v.g., más de aproximadamente 70 veces, cuando están direccionadas a un solo sitio.

Los análisis estructurales y bioquímicos demuestran que el DNA está en muchos casos ligeramente desenrollado cuando se fija a péptidos con dedos de cinc (Pavletich & Pabo, *Science* 252:809-817 (1991); Shi & Berg, *Biochemistry* 35:3845-3848 (1996); Nekludova & Pabo, *Proc. Natl. Acad. Sci. USA* 91:6948-6952 (1994)). Estudios de modelización han demostrado que en el DNA B ideal, el enlazador canónico es demasiado corto para permitir un acoplamiento favorable de Zif268 (Elrod-Erickson *et al.*, *Structure* 4: 1171-1180 (1996)); el DNA debe estar ligeramente desenrollado para interactuar con los dedos de cinc en el modo observado en el complejo Zif268. Esencialmente, parece ser que la periodicidad helicoidal de los dedos de cinc no coincide exactamente con la periodicidad helicoidal del B-DNA. Dado que la tensión de desenrollado puede llegar a ser un problema más importante cuando existen más dedos (las periodicidades helicoidales del péptido y el DNA pueden llegar a desfasarse progresivamente), se testaron enlazadores más largos y más flexibles en el diseño de las proteínas con dedos múltiples (véase Kim & Pabo, *Proc. Natl. Acad. Sci. U.S.A.* 95: 2812-2817 (1998)).

La presente invención demuestra que pueden utilizarse enlazadores de seis aminoácidos a doce aminoácidos para producir proteínas con dedos de cinc que tienen afinidad incrementada. Por ejemplo, se utilizó un enlazador de ocho aminoácidos para una proteína con dedos de cinc que reconocía sitios diana adyacentes separados por un solo par de bases. Se utilizó un enlazador de once aminoácidos para una proteína con dedos de cinc que reconocía sitios diana adyacentes separados por dos pares de bases. Los enlazadores de la invención pueden diseñarse también utilizando modelización basada en estructura. En la modelización basada en estructura, se fabrica un modelo que muestra la fijación de cada polipéptido de dominio de fijación de DNA a su sitio diana de DNA. El modelo se utiliza luego para determinar la separación física de los dominios a medida que los mismos se fijan a sitios diana adyacentes. La separación física entre los dominios se utiliza para determinar la longitud mínima del enlazador utilizado para conectar el aminoácido C-terminal del primer dominio con el aminoácido N-terminal del segundo dominio, sin impedimento estérico para el enlazador o los dominios de fijación de DNA. Esta longitud se incrementa luego en 1-2 Å, a fin de crear un enlazador ligeramente más largo y flexible que evita la introducción de la tensión en la proteína con dedos de cinc quimérica.

A menudo se utilizan programas de computadora para modelización basada en estructura, aunque los modelos pueden construirse también físicamente. Ejemplos de programas de computadora utilizados para modelización basada en estructura incluyen Insight II (Biosym Technologies, San Diego) y Quanta 4.0 (Molecular Simulations (Burlington, MA)). Los programas utilizan a menudo información derivada de estudios obtenidos por cristalografía de rayos X de proteínas de fijación de DNA a fin de proporcionar las coordenadas apropiadas para las proteínas. Esta información puede obtenerse también de bases de datos disponibles públicamente tales como el Brookhaven Protein Data Bank. Esta información puede utilizarse también para extrapolar distancias y coordenadas para proteínas de fijación de DNA cuya estructura cristalina se desconoce. Los modelos de B-DNA son bien conocidos en la técnica. Las coordenadas relevantes (v.g., distancias y tamaños) se utilizan con programas de modelización por computadora de elección, utilizando las instrucciones del fabricante y los parámetros por defecto. Alternativamente, pueden utilizarse parámetros individualizados. La modelización basada en estructura puede realizarse como se describe, v.g. en Kim & Pabo, *Proc. Nat'l. Acad. Sci. U.S.A.* 95:2812-2817 (1998); Pavletich & Pabo, *Science* 252:809-817 (1991); Rebar, Tesis Ph.D. (Massachusetts Institute of Technology, Cambridge MA) (1997); Liu *et al.*, *Proc. Nat'l. Acad. Sci. U.S.A.* 94:5525-5530 (1997); Pomerantz *et al.*, *Science* 267:93-96 (1995); Pomerantz *et al.*, *Proc. Nat'l. Acad. Sci. U.S.A.* 92:9752-9756 (1995); Li *et al.*, *Nature Biotechnology* 16:190-195 (1998); Kim *et al.*, *Proc. Natl. Acad. Sci. USA* 94:3616-3620 (1997); y Pomerantz *et al.*, *Biochemistry* 4:965-970 (1997).

Dos criterios básicos sugieren que las alineaciones de dominios de fijación de DNA tienen potencial para combinación en una proteína quimérica que fija DNA: (1) la ausencia de colisiones entre los dominios, y (2) posicionamiento coherente de las regiones carboxil- y amino-terminales de los dominios, es decir, los dominios están orientados de tal manera que la región carboxil-terminal de un polipéptido puede estar unida a la región amino-terminal del polipéptido siguiente.

El enlazador utilizado para unir los dos dominios de fijación de DNA puede comprender cualquier secuencia de aminoácidos que no impida sustancialmente la interacción de los dominios de fijación del DNA con sus sitios diana respectivos. Residuos de aminoácido preferidos para los enlazadores de la presente invención incluyen, pero sin carácter limitante, glicina, alanina, leucina, serina, valina y treonina. Una vez que se ha seleccionado la longitud de la secuencia de aminoácidos, puede seleccionarse la secuencia del enlazador, v.g., por tecnología de bibliotecas de presentación de fago (véase, v.g. la patente U.S. No. 5260203), o por utilización de secuencias de enlazadores existentes naturalmente o sintéticas, como un entramado (v.g., GTGQKP y GEKP, véase Liu *et al.*, *Proc. Nat'l. Acad. Sci. U.S.A.* 94:5525-5530 (1997); véase también Whitlow *et al.*, *Methods: A Companion to Methods in Enzymology* 2:97-105 (1991)). Típicamente, los enlazadores de la invención se construyen preparando ácidos nucleicos recombinantes que codifican el enlazador y los dominios de fijación de DNA, que se condensan por la vía de la secuencia de aminoácidos enlazadora. Opcionalmente, los enlazadores pueden producirse también utilizando síntesis de péptidos, y enlazarse luego a los dominios polipeptídicos de fijación de DNA.

Las proteínas con dedos de cinc de la invención se componen de dos o más dominios de fijación de DNA, donde cada dominio de fijación de DNA comprende un polipéptido con dedos de cinc. El segundo dominio de fijación de DNA comprende un dominio de fijación con dedos de cinc, sea el mismo dominio o un dominio heterólogo. Proteínas con dedos de cinc adecuadas incluyen cualquier proteína de la familia Cys₂-His₂, v.g., SP-1, SP-1C, ZIF268, NRE, Tramtrack, GLI, YY1, o TFIIIA (véase, v.g., Jacobs, *EMBO J.* 11:4507 (1992); Desjarlais & Berg, *PNAS* 90:2256-2260 (1993); Christy *et al.*, *PNAS* 85:7857-7861 (1988); Greisman & Pabo, *Science* 275:657-661 (1997); Fairall *et al.*, *Nature* 366:483 (1993); Paveltich *et al.*, *Science* 261:1701 (1993)).

En otra realización, la proteína con dedos de cinc está enlazada a al menos uno o más dominios reguladores, descritos más adelante. Dominios reguladores preferidos incluyen dominios represores o activadores de factores de transcripción tales como KRAB y VP16, dominios co-represores y co-activadores, DNA-metil-transferasas, histona-acetiltransferasas, histona-desacetilasas, y endonucleasas tales como FokI. Las secuencias de aminoácidos de los dominios de fijación de DNA pueden ser existentes naturalmente o no existentes naturalmente (o modificadas).

La expresión de proteínas con dedos de cinc puede controlarse también por sistemas tipificados por los sistemas regulados por tet y el sistema RU-486 (véase, v.g., Gossen & Bujard, *PNAS* 89:5547 (1992); Oligino *et al.*, *Gene Ther.* 5:491-496 (1998); Wang *et al.*, *Gene Ther.* 4:432-441 (1997); Neering *et al.*, *Blood* 88:1147-1155 (1996); y Rendahl *et al.*, *Nat. Biotechnol.* 16:757-761 (1998)). Éstos imparten control de moléculas pequeñas en la expresión de la proteína con dedos de cinc y por tanto imparten control de moléculas pequeñas en el o los sitios diana de interés. Esta beneficiosa característica podría utilizarse en modelos de cultivo de células, en terapia génica, y en animales y plantas transgénicos.

La especificidad de fijación de las proteínas de fijación de DNA hace que las mismas sean particularmente útiles debido a que poseen propiedades de fijación de DNA distintas de las de proteínas conocidas. Las proteínas prefieren fijarse a los sitios diana adyacentes y, por tanto, pueden utilizarse para modular la expresión de genes que tengan los sitios diana adyacentes. Estas proteínas con dedos de cinc tienen una afinidad para los sitios diana adyacentes que se encuentra en el intervalo femtomolar, v.g., 100 femtomoles, 10 femtomoles, o menos, en algunos casos tan baja como 2-4 femtomoles, y en algunos casos 1 femtomolar o menor.

Las proteínas con dedos de cinc producidas utilizando el método de la invención tienen numerosas aplicaciones, que incluyen aplicaciones terapéuticas, diagnósticas, y de investigación tales como en modelos de células o animales y en genómica funcional. Por ejemplo, las proteínas con dedos de cinc pueden utilizarse para regular la expresión génica, permitiendo nuevas aplicaciones terapéuticas en humanos y otros mamíferos, v.g., tratamiento de enfermedades genéticas, cáncer, infecciones fúngicas, protozoarias, bacterianas y virales, isquemia, enfermedad vascular, artritis, trastornos inmunológicos, etc., proporcionando además medios para desarrollar plantas con fenotipos alterados, con inclusión de resistencia a las enfermedades, maduración de los frutos, composición de azúcares y aceites, rendimiento, y color. Adicionalmente, las proteínas con dedos de cinc de la presente invención pueden utilizarse para ensayos de diagnóstico y para ensayos de genómica funcional.

Como se describe en esta memoria, las proteínas con dedos de cinc pueden diseñarse para reconocer cualquier sitio diana adecuado para cualquiera de los usos descritos en esta memoria, v.g., genes eucariotas y procariotas, genes celulares, genes virales, genes de protozoos, genes fúngicos, y genes bacterianos. En general, los genes adecuados a regular incluyen citoquinas, linfoquinas, factores de crecimiento, factores mitógenos, factores quimiotácticos, factores onco-activos, receptores, canales de potasio, proteínas G, moléculas de transducción de señales, y otros genes relacionados con enfermedades.

Un tema general en la función de los factores de transcripción es que la fijación simple y la proximidad suficiente al promotor son todo lo que se requiere por regla general. El posicionamiento exacto con relación al promotor, la orientación y, dentro de ciertos límites, la distancia no son muy importantes. Esta característica permite una flexibilidad considerable en la elección de los sitios para construcción de proteínas con dedos de cinc. El sitio diana reconocido por la proteína con dedos de cinc puede ser por consiguiente cualquier sitio adecuado del gen diana que permita la activación o represión de la expresión génica por una proteína con dedos de cinc, enlazada opcionalmente a un dominio regulador.

Sitios diana preferidos que incluyen regiones adyacentes a, situadas aguas abajo, o aguas arriba del sitio de comienzo de la transcripción. Adicionalmente, sitios diana que están localizados en regiones intensificadoras, sitios represores, sitios de pausa de RNA-polimerasa, y sitios reguladores específicos (v.g., sitios SP-1, elementos de respuesta a la hipoxia, elementos de reconocimiento de receptores nucleares, sitios de fijación de p53), sitios en la región codificante del cDNA o en una región codificante del marcador de secuencia expresado (EST). Como se describe más adelante, típicamente cada dedo reconoce 2-4 pares de bases, fijándose una proteína con dedos de cinc de 2 dedos a un sitio diana de 4 a 7 pares de bases, fijándose una proteína con dedos de cinc de 3 dedos a un sitio de 6 a 10 pares de bases, y fijándose una proteína con dedos de cinc de 6 dedos a dos sitios diana adyacentes, teniendo cada sitio diana de 6 a 10 pares de bases.

Las proteínas con dedos de cinc de la invención pueden testarse respecto a actividad *in vivo* utilizando un ensayo simple (*Current Protocols in Molecular Biology* (Ausubel *et al.*, ed., 1994)). El ensayo *in vivo* utiliza un plásmido que codifica la proteína con dedos de cinc quimérica, que se co-expresa con un plásmido informador que contiene un gen de test, v.g., el gen de luciferasa, el gen de cloranfenicol-acetil-transferasa (CAT) o el gen de la hormona

del crecimiento humana (hGH), con un sitio diana para la proteína con dedos de cinc quimérica. Los dos plásmidos se introducen juntos en células hospedadoras. Un segundo grupo de células sirve como grupo de control y aloja un plásmido que codifica el factor de transcripción y un plásmido que contiene el gen informador sin el sitio de fijación para el factor de transcripción.

Se mide la producción de transcritos de genes informadores o la cantidad de actividad de la proteína relevante; si la síntesis de mRNA por el gen informador o la cantidad de actividad de la proteína relevante es mayor que la del gen de control, el factor de transcripción es un regulador positivo de la transcripción. Si la síntesis del mRNA del gen informador o la cantidad de actividad de la proteína relevante es menor que la del control, el factor de transcripción es un regulador de la transcripción negativo.

Opcionalmente, el ensayo puede incluir un plásmido de control de la eficiencia de transfección. Este plásmido expresa un producto génico independiente del gen informador, y la cantidad de este producto génico indica de modo aproximado cuántas células están capturando los plásmidos y con qué grado de eficiencia está siendo introducido el DNA en las células. La proteína con dedos de cinc puede testarse también con respecto a modulación de un gen endógeno *in vivo*, utilizando métodos conocidos por los expertos en la técnica.

En una realización, la presente invención proporciona una fusión en la cual el péptido de 3 dedos Zif268 se enlazó a un péptido con 3 dedos de diseño (designado "NRE") que reconoce específicamente un elemento de respuesta a una hormona nuclear (Greisman & Pabo, *Science* 275: 657 (1997)). Ensayos de desplazamiento en gel indican que este péptido de 6 dedos, 268//NRE se fija a un sitio de DNA compuesto de 18 pares de bases con una constante de disociación en el campo femtomolar. Los enlazadores ligeramente más largos utilizados en esta proteína de fusión proporcionan una mejora espectacular en la afinidad de fijación de DNA, actuando mucho mejor que los enlazadores canónicos "TGEKP" que se han utilizado en estudios previos. Experimentos de transfección de cultivo de tejidos muestran también que el péptido 268//NRE es un represor extremadamente eficaz, proporcionando una represión de 72 veces cuando se direcciona a un sitio de fijación próximo al sitio de comienzo de la transcripción. La utilización de esta estrategia y péptidos enlazadores seleccionados por presentación de fago hace posible el diseño de nuevas proteínas de fijación de DNA con afinidad y especificidad de fijación extraordinarias para uso en aplicaciones biológicas y terapia génica.

Los nuevos péptidos de 6 dedos se fijan mucho más fuertemente que las proteínas con dedos múltiples consignadas previamente que utilizaban un enlazador "DGEKP" convencional para conectar dos módulos de 3 dedos o para añadir dedos adicionales a una proteína de 3 dedos. Las proteínas con dedos múltiples con enlazadores canónicos habían sido testadas por Rebar (Rebar (Tesis Ph. D.), *Selection Studies of Zinc Finger-DNA Recognition*, Massachusetts Institute of Technology (1997)), por Shi (Shi, (Tesis Ph.D.), *Molecular Mechanisms of Zinc Finger Protein-Nucleic Acid Interactions*, Johns Hopkins University (1995)), y por Liu *et al.* (Liu *et al.*, *Proc. Natl. Acad. Sci. USA* 94:5525-5530 (1997)). Cada estudio comparaba la fijación de la nueva proteína con dedos múltiples (en el sitio extendido apropiado) con la fijación del péptido de 3 dedos original. Utilizando enlazadores canónicos, un péptido de 4 dedos se fijaba 6,3 veces más fuertemente que el péptido de 3 dedos correspondiente (Rebar (Ph. D. Thesis), *Selection Studies of Zinc Finger-DNA Recognition*, Massachusetts Institute of Technology (1997)), un constructo con 5 dedos no exhibía mejora alguna en K_d con respecto al péptido de 3 dedos original (Shi, (Tesis Ph.D.), *Molecular Mechanisms of Zinc Finger Protein-Nucleic Acid Interactions*, Johns Hopkins University (1995)), y los péptidos de 6 dedos se fijaban de 58 a 74 veces más fuertemente que los péptidos de 3 dedos correspondientes (Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 94:5525-5530 (1997)).

En contraste, los péptidos descritos en esta memoria (véase la sección de ejemplos) se fijan 6.000-90.000 veces más fuertemente que los péptidos de 3 dedos originales. Parece verosímil que los enlazadores más largos utilizados en los constructos 268//NRE y 268//NRE tienen que disipar cierta tensión que se acumula cuando una serie mayor de dedos están conectados todos ellos con enlazadores canónicos. Presumiblemente, esto implica un ligero desapareamiento en la periodicidad helicoidal del DNA y la periodicidad helicoidal preferida de los dedos de cinc, haciendo que los mismos queden ligeramente desalineados, en particular cuando están conectados 4 o más dedos por enlazadores canónicos.

II. Definiciones

Como se utilizan en esta memoria, los términos siguientes tienen los significados adscritos a ellos a no ser que se especifique otra cosa.

El término "proteína con dedos de cinc" o "ZFP" o "polipéptido con dedos de cinc" hace referencia a una proteína que tiene dominios de fijación de DNA que están estabilizados por cinc. Los dominios de fijación de DNA individuales se conocen típicamente como "dedos". Una proteína con dedos de cinc tiene al menos un dedo, típicamente 2 dedos, 3 dedos, o 6 dedos. Cada dedo fija de dos a cuatro pares de bases de DNA, típicamente tres o cuatro pares de bases de DNA (el "subsitio"). Una proteína con dedos de cinc se fija a una secuencia de ácido nucleico denominada un sitio diana o segmento diana. Cada dedo comprende típicamente un subdominio de fijación de DNA de aproximadamente 30 aminoácidos, formador de quelatos de cinc. Un motivo ilustrativo que caracteriza una clase de estas proteínas (clase C_2H_2) es -Cys-(X)₂₋₄-Cys-(X)₁₂-His-(X)₃₋₅-His (donde X es cualquier aminoácido). Estudios realizados han demostrado que un solo dedo de cinc de esta clase está constituido por una hélice alfa que contiene los dos residuos histidina invariantes coordinados con cinc junto con los dos residuos cisteína de una sola vuelta beta (véase, v.g., Berg & Shi, *Science* 271:1081-1085 (1996)).

Un “sitio diana” es la secuencia de ácido nucleico reconocida por una proteína con dedos de cinc o un polipéptido heterólogo de fijación de DNA. Para una proteína con dedos de cinc, un solo sitio diana tiene por regla general aproximadamente 4 a aproximadamente 10 pares de bases. Típicamente, una proteína con dedos de cinc de 2 dedos reconoce un sitio diana de siete pares de bases, una proteína con dedos de cinc de 3 dedos reconoce un sitio diana de seis a diez pares de bases, y una proteína con dedos de cinc de 6 dedos reconoce dos sitios diana adyacentes de nueve a diez pares de bases.

Un “subsitio” es una subsecuencia del sitio diana, y corresponde a una porción del sitio diana reconocida por un solo dedo, v.g., un subsitio de 2-4 bases, típicamente un subsitio de 3 bases. Un sitio diana comprende al menos dos, típicamente tres, cuatro, cinco, seis o más subsitios, uno para cada dedo de la proteína.

El término “sitios diana adyacentes” hace referencia a sitios diana no solapantes que están separados por 0 a 3 pares de bases.

La “separación física” entre dos dominios de fijación de DNA hace referencia a la distancia entre dos dominios cuando los mismos están fijados a sus sitios diana respectivos. Esta distancia se utiliza para determinar una longitud mínima de un enlazador. Una longitud mínima de un enlazador es la longitud que permitiría que los dos dominios se conectaran sin proporcionar impedimento estérico a los dominios o al enlazador (un enlazador mínimo). Un enlazador que proporciona más de la longitud mínima es un “enlazador flexible”.

“Diseño basado en estructura” hace referencia a métodos de determinación de la longitud de enlazadores mínimos y enlazadores flexibles, utilizando modelos físicos o de computadora de proteínas de fijación de DNA unidas a sus sitios diana respectivos.

“ K_d ” hace referencia a la constante de disociación para el compuesto, es decir, la concentración de un compuesto (v.g., una proteína con dedos de cinc) que proporciona la fijación semimáxima del compuesto a su diana (es decir, la mitad de las moléculas del compuesto están fijadas a la diana) en condiciones dadas (v.g. cuando $[diana] \ll K_d$), como se mide utilizando un sistema de ensayo dado (véase, v.g., la patente U.S. No. 5789538). El sistema de ensayo utilizado para medir la K_d debe seleccionarse de modo que proporcione la medida más exacta del valor K_d real de la proteína con dedos de cinc. Puede utilizarse cualquier sistema de ensayo, con tal que el mismo proporcione una medida exacta del valor K_d real de la proteína con dedos de cinc. En una realización, el K_d para las proteínas con dedos de cinc de la invención se mide utilizando un ensayo de cambio de la movilidad electroforética (“EMSA”), como se describe en esta memoria. A no ser que se haga un ajuste para la pureza o actividad de las proteínas con dedos de cinc, los cálculos de K_d pueden dar como resultado una subestimación del valor K_d verdadero de una proteína con dedos de cinc dada.

La expresión “adyacente a un sitio de iniciación de la transcripción” se refiere a un sitio diana que se encuentra dentro de aproximadamente 50 bases aguas arriba o aguas abajo de un sitio de iniciación de la transcripción. “Aguas arriba” de un sitio de iniciación de la transcripción se refiere a un sitio diana que se encuentra a más de aproximadamente 50 bases en dirección 5' del sitio de iniciación de la transcripción (es decir, en la región no transcrita del gen).

La expresión “sitio de pausa de RNA-polimerasa” se describe en Uptain *et al.*, *Annu. Rev. Biochem.* 66: 117-172 (1997).

El término “heterólogo” es un término relativo que, cuando se utiliza con referencia a porciones de un ácido nucleico indica que el ácido nucleico comprende dos o más subsecuencias que no se encuentran en la misma relación una a otra en la naturaleza. Por ejemplo, un ácido nucleico que se produce recombinantemente tiene típicamente dos o más secuencias de genes no afines dispuestos sintéticamente para producir un nuevo ácido nucleico funcional, v.g., un promotor de una fuente y una región codificante de otra fuente. Los dos ácidos nucleicos son por tanto heterólogos uno a otro en este contexto. Cuando se añaden a una célula, los ácidos nucleicos recombinantes serían también heterólogos respecto a los genes endógenos de la célula. Así, en un cromosoma, un ácido nucleico heterólogo podría incluir un ácido nucleico no nativo (no existente naturalmente) que se ha integrado en el cromosoma, o un ácido nucleico extracromosómico no nativo (no existente naturalmente). En contraste, un fragmento de cromosoma que ha sufrido translocación naturalmente no sería considerado como heterólogo en el contexto de esta solicitud de patente, dado que comprende una secuencia de ácido nucleico endógena que es nativa para la célula mutada.

Un “dominio regulador” se refiere a una proteína o un dominio proteínico que tiene una actividad tal como actividad de modulación de la transcripción, actividad de modificación del DNA, actividad modificadora de proteína y análoga cuando está fijado a un dominio de fijación de DNA, es decir, una proteína con dedos de cinc. Ejemplos de dominios reguladores incluyen proteínas o dominios efectores de proteínas, v.g., factores y co-factores de transcripción (v.g., KRAB, MAD, ERD, SID, la subunidad p65 del factor nuclear kappa B, el factor 1 de respuesta precoz al crecimiento, y receptores de hormonas nucleares, VP16, VP64), endonucleasas, integrasas, recombinasas, metiltransferasas, histona-acetiltransferasas, histona-desacetilasas, etc. Activadores y represores incluyen co-activadores y co-represores (véase, v.g., Utley *et al.*, *Nature* 394: 498-502 (1998)).

Un “dominio de fijación de DNA heterólogo” se refiere a un dominio de fijación de DNA de una proteína que no es una proteína con dedos de cinc, tal como una enzima de restricción, un receptor de hormona nuclear, una proteína de homeodominio tal como dentada o antenopodia, una proteína bacteriana con motivo hélice-vuelta-hélice tal como

el represor lambda y el represor tet, Gal4, la proteína de fijación TATA, proteínas con motivo hélice-bucle-hélice tales como myc y myo D, proteínas del tipo de cremallera de leucina tales como fos y jun, y proteínas con motivo de hoja beta tales como los represores met, arc, y arc, y mnt.

5 “Humanizada” hace referencia a una secuencia polipeptídica no humana que se ha modificado para minimizar la inmunorreactividad en humanos, típicamente por alteración de la secuencia de aminoácidos para mimetizar secuencias humanas existentes, sin alterar sustancialmente la función de la secuencia polipeptídica (*véase, v.g., Jones et al., Nature* 321: 522-525 (1986), y la solicitud de patente publicada del Reino Unido No. 8707252). Las secuencias de cadena principal para las proteínas con dedos de cinc se seleccionan preferiblemente a partir de proteínas con dedos de cinc
10 C₂H₂ humanas existentes (*v.g., SP-1*). Los dominios funcionales se seleccionan preferiblemente de genes humanos existentes (*v.g., el dominio de activación de la subunidad p65 de NF-kappa B*). En caso de ser posible, las secuencias de hélice de reconocimiento se seleccionarán de los miles de dominios de reconocimiento de DNA de proteínas con dedos de cinc existentes proporcionados por la secuenciación del genoma humano. En la medida de lo posible, los dominios se combinarán como unidades a partir de las mismas proteínas existentes. La totalidad de estos pasos
15 minimizarán la introducción de nuevos epítopes de juntura en las proteínas con dedos de cinc quiméricas y harán que las proteínas con dedos de cinc manipuladas genéticamente sean menos inmunógenas.

“Ácido nucleico” hace referencia a desoxirribonucleótidos o ribonucleótidos y polímeros de los mismos en forma mono- o bicatenaria. El término abarca ácidos nucleicos que contienen análogos de nucleótidos conocidos o residuos o
20 eslabones de cadena principal modificados, que son sintéticos, existentes naturalmente, y no existentes naturalmente, que tienen propiedades de fijación similares al ácido nucleico de referencia, y que son metabolizados de manera similar a los nucleótidos de referencia. Ejemplos de tales análogos incluyen, sin limitación, fosforotioatos, fosforamidatos, metilfosfonatos, metilfosfonatos quirales, 2-O-metil-ribonucleótidos, y ácidos nucleicos peptídicos (PNAs). A no ser que se indique otra cosa, una secuencia de ácido nucleico particular abarca también implícitamente variantes de la
25 misma modificadas de manera conservadora (*v.g., sustituciones de codones degenerados*) y secuencias complementarias, así como la secuencia indicada explícitamente. El término ácido nucleico se utiliza intercambiamente con gen, cDNA, mRNA, oligonucleótido, y polinucleótido. Las secuencias nucleotídicas se presentan en esta memoria en la orientación convencional 5'-3'.

30 Los términos “polipéptido”, “péptido” y “proteína” se utilizan intercambiamente en esta memoria haciendo referencia a un polímero de residuos de aminoácidos. Los términos se aplican a polímeros de aminoácidos en los cuales uno o más residuos de aminoácido es un análogo o mimético de un aminoácido existente naturalmente correspondiente, así como a polímeros de aminoácidos existentes naturalmente. Los polipéptidos pueden modificarse, *v.g., por la adición de residuos carbohidrato para formar glicoproteínas*. Los términos “polipéptido”, “péptido” y “proteína”
35 incluyen glicoproteínas, así como no-glicoproteínas. Las secuencias polipeptídicas se presentan en esta memoria en la orientación convencional N-terminal a C-terminal.

El término “aminoácido” se refiere a aminoácidos existentes naturalmente y sintéticos, así como a análogos de aminoácidos y miméticos de aminoácidos que se comportan de una manera similar a los aminoácidos existentes naturalmente. Los aminoácidos existentes naturalmente son los codificados por el código genético, así como aquellos aminoácidos que se modifican posteriormente, *v.g. hidroxiprolina, carboxiglutamato, y O-fosfoserina*. La expresión
40 análogos de aminoácidos hace referencia a compuestos que tienen la misma estructura química básica que un aminoácido existente naturalmente, es decir, un carbono alfa que está unido a un hidrógeno, un grupo carboxilo, un grupo amino, y un grupo R, *v.g., homoserina, norleucina, metionina-sulfóxido, metionina, y metil-sulfonio*. Tales análogos tienen grupos R modificados (*v.g. norleucina*) o cadenas principales peptídicas modificadas, pero retienen la misma estructura química básica que un aminoácido existente naturalmente. La expresión miméticos de aminoácidos hace
45 referencia a compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que se comporta de una manera similar a un aminoácido existente naturalmente.

50 “Variantes modificadas de manera conservadora” se aplica tanto a secuencias de aminoácidos como secuencias de ácido nucleico. Con respecto a secuencias de ácido nucleico particulares, la expresión variante modificada de manera conservadora se refiere a aquellos ácidos nucleicos que codifican secuencias de aminoácidos idénticas o esencialmente idénticas, o en el caso en que el ácido nucleico no codifica una secuencia de aminoácidos, a secuencias esencialmente idénticas. Específicamente, pueden conseguirse sustituciones de codones degenerados por generación de secuencias
55 en las cuales la tercera posición de uno o más codones seleccionados (o de todos) está sustituida con residuos de base mixta y/o residuos desoxiinosina (*Batzer et al., Nucleic Acid Res.* 19:5081 (1991); *Ohtsuka et al., J. Biol. Chem.* 260:26052608 (1985); *Rossolini et al., Mol. Cell. Probes* 8:91-98 (1994)). Debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier proteína dada. Por ejemplo, los codones GCA, GCC, GCG y GCU codifican todos ellos el aminoácido alanina. Así, en cualquier posición en la que una alanina está especificada por un codón en un aminoácido de esta memoria, el codón puede alterarse por cualquiera de los codones correspondientes descritos sin alterar el polipéptido codificado. Tales variaciones de ácidos nucleicos son “variaciones silenciosas”, que son una especie de variaciones modificadas de manera conservadora. Cada secuencia de ácido nucleico de esta memoria que codifica un polipéptido describe cualquier posible variación silenciosa del
60 ácido nucleico. Una persona con experiencia reconocerá que cada codón en un ácido nucleico (excepto AUG, que es ordinariamente el único codón para metionina, y TGG, que es ordinariamente el único codón para triptófano) puede modificarse para producir una molécula funcionalmente idéntica. De acuerdo con ello, cada variación silenciosa de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita.

En cuanto a las secuencias de aminoácidos y ácidos nucleicos, sustituciones, deleciones o adiciones individuales que alteran, añaden o delecionan un solo aminoácido o nucleótido o un pequeño porcentaje de aminoácidos o nucleótidos en la secuencia crean una “variante modificada de manera conservadora”, donde la alteración da como resultado la sustitución de un aminoácido con un aminoácido químicamente similar. Tablas de sustituciones conservadoras que proporcionan aminoácidos funcionalmente similares son bien conocidos en la técnica. Tales variantes modificadas de manera conservadora se suman a y no excluyen las variantes polimórficas y alelos de la invención.

Los grupos siguientes contienen cada uno aminoácidos que son sustituciones conservadoras de uno por otro:

- 1) alanina (A), glicina (G);
- 2) serina (S), treonina (T);
- 3) ácido aspártico (D), ácido glutámico (E);
- 4) asparagina (N), glutamina (Q);
- 5) cisteína (C), metionina (M);
- 6) arginina (R), lisina (K), histidina (H);
- 7) isoleucina (I), leucina (L), valina (V); y
- 8) fenilalanina (F), tirosina (Y), triptófano (W).

(Véase, v.g., Creighton, *Proteins* (1984) para una discusión de las propiedades de los aminoácidos).

III. Diseño de Proteínas con Dedos de cinc

Las proteínas con dedos de cinc de la invención comprenden al menos un polipéptido con dedos de cinc enlazado por un enlazador flexible a al menos un segundo dominio de fijación de DNA, que opcionalmente es un segundo polipéptido con dedos de cinc. La proteína con dedos de cinc puede contener más de dos dominios de fijación de DNA, así como uno o más dominios reguladores. Los polipéptidos con dedos de cinc de la invención pueden manipularse genéticamente para reconocer un sitio diana seleccionado en el gen de elección. Típicamente, una cadena principal de cualquier C_2H_2 ZFP adecuado, tal como SP-1, SP-1C, o Zif268, se utiliza como el entramado para los polipéptidos con dedos de cinc manipulados genéticamente (véase, v.g., Jacobs, *EMBO J* 11:4507 (1992); Desjarlais & Berg, *PNAS* 90:2256-2260 (1993)). Pueden utilizarse luego diversos métodos para diseñar y seleccionar un polipéptido con dedos de cinc con afinidad alta para su diana. Un polipéptido con dedos de cinc puede estar diseñado o seleccionado para fijarse a cualquier sitio diana adecuado en el gen diana, con afinidad alta.

Puede utilizarse cualquier método adecuado conocido en la técnica para diseñar y construir ácidos nucleicos que codifican polipéptidos con dedos de cinc, v.g., presentación de fago, mutagénesis aleatoria, bibliotecas combinatorias, diseño racional por computadora, selección por afinidad, PCR, clonación a partir de cDNA o bibliotecas genómicas, construcción sintética y análogos (véase, v.g., patente U.S. No. 5786538; Wu *et al.*, *PNAS* 92:344-348 (1995); Jamieson *et al.*, *Biochemistry* 33:5689-5695 (1994); Rebar & Pabo, *Science* 263:671-673 (1994); Choo & Klug, *PNAS* 91:11163-11167 (1994); Choo & Klug, *PNAS* 91: 11168-11172 (1994); Desjarlais & Berg, *PNAS* 90:2256-2260 (1993); Desjarlais & Berg, *PNAS* 89:7345-7349 (1992); Pomerantz *et al.*, *Science* 267:93-96 (1995); Pomerantz *et al.*, *PNAS* 92:9752-9756 (1995); y Liu *et al.*, *PNAS* 94:5525-5530 (1997); Griesman & Pabo, *Science* 275:657-661 (1997); Desjarlais & Berg, *PNAS* 91:11-99-11103 (1994)).

Utilizando estos métodos, puede sintetizarse luego un polipéptido con dedos de cinc que se fija al sitio preseleccionado. Estos métodos de selección del sitio diana están basados, en parte, en el reconocimiento de que la presencia de uno o más sitios D-able en un segmento diana confiere(n) el potencial para mayor afinidad de fijación en un polipéptido con dedos de cinc seleccionado o diseñado para fijarse a dicho sitio con relación a polipéptidos con dedos de cinc que se fijan a segmentos diana que carecen de sitios D-able.

Un sitio o subsitio D-able es una región de un sitio diana que permite que un solo dedo de cinc diseñado adecuadamente se fije a cuatro bases en lugar de tres del sitio diana.

La fijación de un solo dedo de cinc a un segmento diana de cuatro bases impone restricciones tanto en cuanto a la secuencia de la cadena diana como en la secuencia de aminoácidos del dedo de cinc. El sitio diana dentro de la cadena diana debería incluir el motivo del sitio “D-able” 5’ NNGK 3’, en el cual N y K son códigos convencionales de ambigüedad IUPAC-IUB. Un dedo de cinc para fijación a un sitio de este tipo debería incluir un residuo arginina en la posición -1 y un ácido aspártico (o menos preferiblemente un ácido glutámico) en la posición +2. Los residuos arginina en la posición -1 interaccionan con el residuo G en el sitio D-able. El residuo de ácido aspártico (o ácido glutámico) en la posición +2 del dedo de cinc interacciona con la base de la cadena opuesta complementaria a la base K en el sitio D-able. Es la interacción entre ácido aspártico (símbolo D) y la base de la cadena opuesta (cuarta base)

la que confiere el nombre de sitio D-able. Como resulta evidente a partir de la fórmula del sitio D-able, existen dos subtipos de sitio D-able: 5' NNGG 3' y 5' NNGT 3'. Para el primer sitio, el ácido aspártico o ácido glutámico en la posición +2 de un dedo de cinc interacciona con un C en la cadena opuesta al sitio D-able. En el último sitio, el ácido aspártico o ácido glutámico en la posición +2 del dedo de cinc interacciona con un A en la cadena opuesta al sitio D-able. En general, se prefiere NNGG a NNGT.

En el diseño de un polipéptido con dedos de cinc de 3 dedos, debe seleccionarse un sitio diana en el cual al menos un dedo de la proteína, y opcionalmente, 2 o los 3 dedos tengan el potencial para fijarse a un sitio D-able. Esto puede lograrse seleccionando un sitio diana del interior de un gen diana mayor que tenga la fórmula 5'-NNx aNy bNcz-3', en donde

cada uno de los juegos (x, a), (y, b) y (z, c) es (N, N), o (G, K);

al menos uno de (x, a), (y, b), y (z, c) es (G, K) y

N y K son códigos de ambigüedad IUPAC-IUB.

Dicho de otro modo, al menos uno de los tres juegos (x, a), (y, b) y (z, c) es el juego (G, K), lo que significa que la primera posición del juego es G y la segunda posición es G o T. Aquellos de los tres juegos (en su caso) que no son (G, K) son (N, N), lo que indica que la primera posición del juego puede estar ocupada por cualquier nucleótido y la segunda posición del juego puede estar ocupada por cualquier nucleótido. Como ejemplo, el juego (x, a) puede ser (G, K) y los juegos (y, b), y (z, c) pueden ser ambos (N, N).

En la fórmula 5'-NNx aNy bNcz-3', los tripletes de NNx aNy y bNcz representan los tripletes de bases en la cadena diana fijados por los 3 dedos en un polipéptido con dedos de cinc. Si solamente uno de x, y y z es un G, y este G va seguido por un K, el sitio diana incluye un solo subsitio D-able. Por ejemplo, si solamente x es G y a es K, el sitio se lee 5'-NNG KNy bNcz-3' con el subsitio D-able resaltado. Si tanto x como y pero no z son G, y a y b son K, entonces el sitio diana tiene dos subsitios D-able solapantes como sigue: 5'-NNG **KNG** KNz c-3', representándose uno solo de tales sitios en negrita y el otro en cursiva. Si los tres de x, y, y z son G y a, b y c son K, entonces el segmento diana incluye 3 subsitios D-able, como sigue 5' **NNG** **KNG** *KNG* K 3', representándose los subsitios D-able en negrita, cursiva y subrayado.

Estos métodos operan por tanto por selección de un gen diana, y búsqueda sistemática dentro de las subsecuencias posibles del gen de sitios diana de acuerdo con la fórmula 5'-NNx aNy bNcz-3', como se ha descrito arriba. En algunos de dichos métodos, cada una de las posibles subsecuencias de 10 bases contiguas en cualquiera de las cadenas de un gen diana potencial se evalúa para determinar si la misma se ajusta a la fórmula anterior, y, en caso afirmativo, cuántos sitios D-able están presentes. Típicamente, una comparación de este tipo se realiza mediante computadora, y se imprime como salida una lista de sitios diana de acuerdo con la fórmula. Opcionalmente, tales sitios diana pueden imprimirse en diferentes subconjuntos de acuerdo con el número de sitios D-able que estén presentes.

En una variación, los métodos de la invención identifican segmentos primero y segundo diana, cada uno de los cuales está independientemente de acuerdo con la fórmula anterior. Los dos segmentos diana en tales métodos se ven limitados a ser adyacentes o próximos (es decir, dentro de aproximadamente 0-5 bases) uno de otro en el gen diana. La estrategia que subyace en la selección de segmentos diana próximos consiste en permitir el diseño de un polipéptido con dedos de cinc formado por el enlace de dos polipéptidos componentes con dedos de cinc específicos para los segmentos diana primero y segundo respectivamente. Estos principios pueden extenderse para seleccionar sitios diana que estén fijados por polipéptidos con dedos de cinc con cualquier número de dedos componentes. Por ejemplo, un sitio diana adecuado para una proteína de 9 dedos tendría 3 segmentos componentes, cada uno de ellos conforme a la fórmula anterior.

Los sitios diana identificados por los métodos anteriores pueden someterse a evaluación ulterior por otros criterios o pueden utilizarse directamente para el diseño o selección (en caso necesario) y producción de un polipéptido con dedos de cinc específico para dicho sitio. Un criterio adicional para evaluación de los sitios diana potenciales es su proximidad a regiones particulares dentro de un gen. Si un polipéptido con dedos de cinc debe utilizarse para reprimir un gen celular por sí mismo (es decir, sin enlazar el polipéptido con dedos de cinc a un resto represor), entonces la localización óptima parece encontrarse a o dentro de 50 pb aguas arriba o aguas abajo del sitio de comienzo de la transcripción, para interferir con la formación del complejo de transcripción (Kim & Pabo, *J. Biol. Chem.* 272:29795-29800 (1977)) o competir por una proteína esencial de fijación de intensificadores. EN cambio, si un polipéptido con dedos de cinc se fusiona a un dominio funcional tal como el dominio represor KRAB o el dominio activador VP16, la localización del sitio de fijación es considerablemente más flexible y puede estar situada fuera de las regiones reguladoras conocidas. Por ejemplo, un dominio KRAB puede reprimir la transcripción en un promotor hasta al menos 3 kpb del punto en el que está fijado KRAB (Margolin *et al.*, *PNAS* 91: 4509-4513 (1994)). Así, pueden seleccionarse sitios diana que no incluyen o solapan necesariamente segmentos de significado biológico demostrable con genes diana, tales como secuencias reguladoras. Otros criterios para evaluación ulterior de segmentos diana incluyen la disponibilidad previa de polipéptidos con dedos de cinc que se fijen a tales segmentos o segmentos afines, y/o la facilidad de diseño de nuevos polipéptidos con dedos de cinc que se fijen a un segmento diana dado.

Después que se ha seleccionado un segmento diana, puede proporcionarse un polipéptido con dedos de cinc que se fija al segmento por una diversidad de métodos. El más sencillo de los métodos consiste en proporcionar un polipéptido con dedos de cinc previamente caracterizado procedente de una colección existente conocida que se fija al sitio diana. Sin embargo, en muchos casos, no existen tales polipéptidos con dedos de cinc. Un enfoque alternativo puede utilizarse también para diseñar nuevos polipéptidos con dedos de cinc, que utiliza información en una base de datos de polipéptidos con dedos de cinc existentes y sus afinidades de fijación respectivas. Un enfoque adicional consiste en diseñar un polipéptido con dedos de cinc basado en reglas de sustitución como se ha expuesto anteriormente. Otra alternativa adicional consiste en seleccionar un polipéptido con dedos de cinc que tiene especificidad para una diana dada por un proceso empírico tal como presentación de fago. En algunos de estos métodos, cada dedo componente de un polipéptido con dedos de cinc está diseñado o seleccionado independientemente de otros dedos componentes. Por ejemplo, cada dedo puede obtenerse a partir de un polipéptido con dedos de cinc preexistente diferente, o cada dedo puede someterse a aleatorización y selección separadas.

Una vez que un polipéptido con dedos de cinc ha sido seleccionado, diseñado, o proporcionado de algún otro modo a un segmento diana dado, el polipéptido con dedos de cinc o el DNA que lo codifica se sintetizan. Métodos ilustrativos para síntesis y expresión de DNA codificante de proteínas con dedos de cinc se describen más adelante. El polipéptido con dedos de cinc o un polinucleótido que codifica el mismo puede utilizarse luego para modulación de la expresión, o análisis del gen diana que contiene el sitio diana al cual se fija el polipéptido con dedos de cinc.

IV. Expresión y purificación de proteínas con dedos de cinc producidas utilizando los métodos de la invención

Proteínas con dedos de cinc que comprenden un enlazador flexible y ácidos nucleicos que codifican dichas proteínas quiméricas con dedos de cinc pueden producirse utilizando técnicas de rutina en el campo de la genética recombinante. Textos básicos que describen los métodos generales de uso en esta invención incluyen Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2ª edición 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); y *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 994)). Además, esencialmente cualquier ácido nucleico puede solicitarse bajo petición de cualquiera de una diversidad de fuentes comerciales. Análogamente, los péptidos y anticuerpos pueden solicitarse bajo petición de cualquiera de una diversidad de fuentes comerciales.

Se utilizan típicamente dos métodos alternativos para crear las secuencias codificantes requeridas para expresar polipéptidos de fijación de DNA de nuevo diseño y el enlazador flexible. Un protocolo es un procedimiento de ensamblaje basado en PCR que utiliza 6 oligonucleótidos solapantes (para producir un solo polipéptido con dedos de cinc de 3 dedos). Tres oligonucleótidos corresponden a secuencias “universales” que codifican porciones del dominio de fijación de DNA entre las hélices de reconocimiento. Estos oligonucleótidos se mantienen constantes para todos los constructos con dedos de cinc. Los otros tres oligonucleótidos “específicos” están diseñados para codificar las hélices de reconocimiento. Estos oligonucleótidos contienen sustituciones fundamentalmente en las posiciones -1, 2, 3 y 6 de las hélices de reconocimiento que los hacen específicos para cada uno de los diferentes dedos de cinc.

Para producir un polipéptido con dedos de cinc de 3 dedos, la síntesis PCR se lleva a cabo en dos pasos. En primer lugar, se crea un molde de DNA bicatenario por combinación de los seis oligonucleótidos (3 universales, 3 específicos) en una reacción PCR de 4 ciclos con un paso de reasociación a baja temperatura, reasociándose con ello los oligonucleótidos para formar un “entramado” de DNA. Las lagunas en el entramado se rellenan por medio de polimerasa termoestable de alta fidelidad, siendo suficiente también la combinación de las polimerasas Taq y Pfu. En la segunda fase de la construcción, el molde con dedos de cinc se amplifica mediante iniciadores externos diseñados para incorporar sitios de restricción en cada extremo para clonación en un vector lanzadera o directamente en un vector de expresión.

Un método alternativo de clonación de las proteínas de fijación de DNA de nuevo diseño está basado en la reasociación de oligonucleótidos complementarios que codifican las regiones específicas de la proteína quimérica con dedos de cinc deseada. Esta aplicación particular requiere que los oligonucleótidos estén fosforilados antes del paso de fijación final. Esto se realiza usualmente antes del comienzo de las reacciones de reasociación, pero el quinasado puede tener lugar también después de la reasociación. Resumidamente, los oligonucleótidos “universales” que codifican las regiones constantes de las proteínas (oligos 1, 2 y 3 anteriores) se reasocian con sus oligonucleótidos complementarios. Adicionalmente, los oligonucleótidos “específicos” que codifican las hélices de reconocimiento de los dedos están reasociados con sus oligonucleótidos complementarios respectivos. Estos oligos complementarios están diseñados para rellena la región que se había rellena previamente con polimerasa en el protocolo arriba descrito. Los oligos complementarios a los oligos 1 y dedo 3 comunes se manipulan genéticamente para dejar secuencias colgantes específicas para los sitios de restricción utilizados en la clonación al vector de elección. El segundo protocolo de ensamblaje difiere del protocolo inicial en los aspectos siguientes: el “entramado” que codifica el ZFP de nuevo diseño está compuesto enteramente por DNA sintético, eliminando con ello el paso de rellena con polimerasa, y adicionalmente el fragmento a clonar en el vector no requiere amplificación. Por último, el diseño de dejar colgantes específicos de secuencia elimina la necesidad de digestiones con enzimas de restricción del fragmento de inserción.

El fragmento resultante que codifica el polipéptido con dedos de cinc de nuevo diseño se fija a un vector de expresión. Las secuencias que codifican el enlazador flexible y el segundo dominio de fijación de DNA (opcionalmente un polipéptido con dedos de cinc) se fijan también al vector para crear una proteína con dedos de cinc quimérica.

Típicamente, el enlazador flexible está codificado por un oligonucleótido que se fija al vector de expresión entre los dos dominios de fijación de DNA. El segundo dominio de fijación de DNA puede construirse como se ha descrito arriba, o puede clonarse u obtenerse a partir de una fuente alternativa utilizando métodos bien conocidos en la técnica, v.g., PCR y análogos. Vectores de expresión que se utilizan comúnmente incluyen, pero sin carácter limitante, un vector de expresión bacteriano modificado pMAL-c2 (New England BioLabs, "NEB") o un vector de expresión eucariota, pcDNA (Promega).

El ácido nucleico que codifica la proteína con dedos de cinc de elección se somete típicamente a clonación en vectores intermedios para transformación en células procariotas o eucariotas para replicación y/o expresión, v.g., para determinación de K_d . Los vectores intermedios son típicamente vectores procariotas, v.g., plásmidos, o vectores lanzadera, o vectores de insecto, para almacenamiento o manipulación del ácido nucleico que codifica la proteína con dedos de cinc o producción de proteína. El ácido nucleico que codifica una proteína con dedos de cinc se clona también típicamente en un vector de expresión, para administración a una célula de planta, célula animal, preferiblemente una célula de mamífero o una célula humana, célula fúngica, célula bacteriana, o célula de protozoo.

Para obtener la expresión de un gen o ácido nucleico clonado, una proteína con dedos de cinc quimérica se subclona típicamente en un vector de expresión que contiene un promotor para dirigir la transcripción. Promotores bacterianos y eucariotas adecuados son bien conocidos en la técnica y se describen, v.g., en Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2ª edición 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); y *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 1994). Sistemas de expresión bacterianos para la expresión de la proteína con dedos de cinc están disponibles, v.g., en *E. coli*, *Bacillus sp.*, y *Salmonella* (Palva *et al.*, *Gene* 22:229-235 (1983)). Kits para tales sistemas de expresión están disponibles comercialmente. Sistemas de expresión eucariotas para células de mamífero, levadura, y células de insecto son bien conocidos en la técnica y están también disponibles comercialmente.

El promotor utilizado para dirigir la expresión de un ácido nucleico de proteínas con dedos de cinc depende de la aplicación particular. Por ejemplo, se utiliza típicamente un promotor constitutivo fuerte para expresión y purificación de proteínas con dedos de cinc. En contraste, cuando una proteína con dedos de cinc se administra *in vivo* para regulación génica, se utiliza un promotor constitutivo o un promotor inducible, dependiendo del uso particular de la proteína con dedos de cinc. El promotor puede incluir también típicamente elementos que son sensibles a la transactivación, v.g., elementos de respuesta hipóxica, elementos de respuesta Gal4, elementos de respuesta al represor lac, y sistemas de control de moléculas pequeñas tales como sistemas reguladas por tet y el sistema RU-486 (véase, v.g., Gossen & Bujard, *Proc. Natl. Acad. Sci. U.S.A.* 89:5547 (1992); Oligino *et al.*, *Gene Ther.* 5:491-496 (1998); Wang *et al.*, *Gene Ther.* 4:432-441 (1997); Neering *et al.*, *Blood* 88:1147-1155 (1996); y Rendahl *et al.*, *Nat. Biotechnol.* 16:757-761 (1998)).

Además del promotor, el vector de expresión contiene típicamente una unidad de transcripción o casete de expresión que contiene todos los elementos adicionales requeridos para la expresión del ácido nucleico en células hospedadoras, sean procariotas o eucariotas. Una casete de expresión típica contiene por tanto un promotor enlazado operativamente, v.g., a la secuencia de ácido nucleico que codifica la proteína con dedos de cinc, y señales requeridas, v.g., para poliadenilación eficiente del transcrito, terminación de la transcripción, sitios de fijación de ribosoma, o terminación de la traducción. Elementos adicionales de la casete pueden incluir, v.g., intensificadores, y señales intrónicas remodeladas heterólogas.

El vector de expresión particular utilizado para transportar la información genética al interior de la células se selecciona con relación al uso propuesto de la proteína con dedos de cinc, v.g., expresión en plantas, animales, bacterias, hongos, protozoos, etc. Vectores de expresión bacterianos estándar incluyen plásmidos tales como plásmidos basados en pBR322, pSKF, pET23D, y sistemas de expresión por fusión disponibles comercialmente, tales como GST y LacZ. Una proteína de fusión preferida es la proteína de fijación de maltosa, "MBP". Tales proteínas de fusión se utilizan para purificación de la proteína con dedos de cinc. Pueden añadirse también marcadores epitópicos a proteínas recombinantes a fin de proporcionar métodos de aislamiento convenientes, para monitorización de la expresión, y para monitorización de la localización celular y subcelular, v.g., c-myc o FLAG.

Vectores de expresión que contienen elementos reguladores de virus eucariotas se utilizan a menudo en vectores de expresión eucariotas, v.g., vectores SV40, vectores del virus del papiloma, y vectores derivados del virus Epstein-Barr. Otros vectores eucariotas ilustrativos incluyen pMSG, pAV009/A+, pMTO10/A+, pMAMneo-5, baculovirus pDSVE, y cualquier otro vector que permita la expresión de proteínas bajo la dirección del promotor SV40 precoz, el promotor SV40 tardío, el promotor metalotioneína, el promotor del virus del tumor mamario murino, el promotor del virus del sarcoma de Rous, el promotor polihedrina, u otros promotores que hayan demostrado eficacia para expresión en células eucariotas.

Algunos sistemas de expresión tienen marcadores para selección de líneas de células transfectadas de manera estable tales como timidina-quinasa, higromicina-B-fosfotransferasa, y dihidrofolato-reductasa. También son adecuados sistemas de expresión de alto rendimiento, tales como la utilización de un vector de baculovirus en células de insecto, con una secuencia codificante de proteína con dedos de cinc bajo la dirección del promotor polihedrina u otros promotores fuertes de baculovirus.

Los elementos que se incluyen típicamente en vectores de expresión incluyen también un replicón que funciona en *E. coli*, un gen que codifica resistencia a antibióticos a fin de permitir la selección de bacterias que alojan plásmidos recombinantes, y sitios de restricción singulares en regiones no esenciales del plásmido que permiten la inserción de secuencias recombinantes.

Se utilizan métodos estándar de transfección para producir líneas de células bacterianas, de mamífero, de levadura o insecto que expresan grandes cantidades de proteína, las cuales se purifican luego utilizando técnicas estándar (véase, v.g., Colley *et al.*, *J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, en *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). La transformación de células eucariotas y procariotas se realiza de acuerdo con técnicas estándar (véase, v.g., Morrison, *J. Bact.* 132:349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* 101:347-362 (Wu *et al.*, eds, 1983).

Puede utilizarse cualquiera de los procedimientos bien conocidos para introducción de secuencias extrañas de nucleótidos en células hospedadoras. Éstos incluyen el uso de transfección con fosfato de calcio, polibreno, fusión de protoplastos, electroporación, liposomas, microinyección, DNA desnudo, vectores plasmídicos, vectores virales, a la vez episómicos e integradores, y cualquiera de los otros métodos bien conocidos para introducción de DNA genómico clonado, cDNA, DNA sintético u otro material genético extraño en una célula hospedadora (Véase, v.g., Sambrook *et al.*, *supra*). Únicamente es necesario que el procedimiento particular de manipulación genética utilizado sea capaz de introducir con éxito al menos un gen en la célula hospedadora capaz de expresar la proteína de elección.

Cualquier método adecuado de purificación de proteínas conocido por los expertos en la técnica puede utilizarse para purificar las proteínas con dedos de cinc de la invención (véase Ausubel, *supra*, Sambrook, *supra*). Adicionalmente, puede utilizarse cualquier hospedador adecuado, v.g., células bacterianas, células de insecto, células de levadura, células de mamífero, y análogos.

En una realización, la expresión de la proteína con dedos de cinc fusionada a una proteína de fijación de maltosa (MBP-proteína con dedos de cinc) en la tensión bacteriana JM109 permite la purificación directa mediante una columna de amilosa (NEB). Niveles altos de expresión de la proteína con dedos de cinc química pueden obtenerse por inducción con IPTG, dado que la fusión MBP-proteína con dedos de cinc en el plásmido de expresión pMal-c2 se halla bajo el control del promotor tac inducible por IPTG (NEB). Bacterias que contienen los plásmidos de fusión MBP-proteína con dedos de cinc se inoculan en medio 2 x YT que contiene ZnCl_2 10 μM , 0,02% glucosa, más 50 $\mu\text{g/ml}$ de ampicilina y se agitan mediante sacudidas a 37°C. En la fase de crecimiento exponencial-medio se añade IPTG hasta 0,3 mM y se mantienen los cultivos agitando mediante sacudidas. Al cabo de 3 horas, las bacterias se cosechan por centrifugación, se destruyen por sonicación, y a continuación se separa el material insoluble por centrifugación. Las proteínas MBP-proteína con dedos de cinc se capturan en una resina fijada a amilosa, se lavan a fondo con tampón que contiene Tris-HCl 20 mM (pH 7,5), NaCl 200 mM, DTT 5 mM y ZnCl_2 50 μM , y se eluyen luego con maltosa esencialmente en el mismo tampón (la purificación está basada en un protocolo estándar de NEB). Las proteínas purificadas se cuantifican y se guardan para análisis bioquímico.

Las propiedades bioquímicas de las proteínas purificadas, v.g., K_d , pueden caracterizarse por cualquier ensayo adecuado. En una realización, se caracteriza K_d por ensayos de cambio de la movilidad electroforética ("EMSA") (Buratowski & Chodosh, en *Current Protocols in Molecular Biology* pp. 12.2.1-12.2.7 (Ausubel ed., 1996)).

V. Dominios reguladores

Las proteínas con dedos de cinc producidas utilizando los métodos de la invención pueden estar asociados opcionalmente con dominios reguladores para modulación de la expresión génica. La proteína con dedos de cinc puede asociarse covalente o no covalentemente con uno o más dominios reguladores, alternativamente dos o más dominios reguladores, siendo los dos o más dominios dos copias del mismo dominio, o dos dominios diferentes. Los dominios reguladores pueden enlazarse covalentemente a la proteína con dedos de cinc, v.g., por un enlazador de aminoácidos, como parte de una proteína de fusión. Las proteínas con dedos de cinc pueden asociarse también con un dominio regulador por la vía de un dominio de dimerización no covalente, v.g., una cremallera de leucina, un dominio N-terminal de la proteína STAT, o una proteína de fijación FK506 (véase, v.g., O'Shea, *Science* 254: 539 (1991), Barahmand-Pour *et al.*, *Curr. Top. Microbiol. Immunol.* 211:121-128 (1996); Klemm *et al.*, *Annu. Rev. Immunol.* 16:569-592 (1998); Klemm *et al.*, *Annu. Rev. Immunol.* 16:569-592 (1998); Ho *et al.*, *Nature* 382:822-826 (1996); y Pomeranz *et al.*, *Biochem.* 37:965 (1998)). El dominio regulador puede estar asociado con la proteína con dedos de cinc en cualquier posición adecuada, con inclusión del término C o N de la proteína con dedos de cinc.

Dominios reguladores comunes para adición a la proteína con dedos de cinc producida utilizando los métodos de la invención incluyen, v.g., dominios de fijación de DNA heterólogos de factores de transcripción, dominios efectores de factores de transcripción (activadores, represores, co-activadores, co-represores), silenciadores, receptores de hormonas nucleares, factores de transcripción de oncogenes (v.g., miembros de las familias myc, jun, fos, myb, max, mad, rel, ets, bcl, myb, mos etc.); y proteínas asociadas a cromatina y sus modificadores (v.g., quinasas, acetilasas y desacetilasas).

Polipéptidos de factores de transcripción a partir de los cuales puede obtenerse un dominio regulador incluyen aquéllos que están implicados en transcripción regulada y basal. Tales polipéptidos incluyen factores de transcripción, sus dominios efectores, coactivadores, silenciadores, receptores de hormonas nucleares (véase, v.g., Goodrich *et al.*, *Cell* 84:825-30 (1996)) para una revisión de proteínas y elementos de ácidos nucleico implicados en transcripción; factores de transcripción en general se revisan en Barnes & Adcock, *Clin. Exp. Allergy* 25 Suppl. 2:46-9 (1995) y Roeder, *Methods Enzymol.* 273:165-71 (1996)). Se conocen también bases de datos dedicadas a factores de transcripción (véase, v.g., *Science* 269:630 (19995)). Factores de transcripción de receptores de hormonas nucleares se describen, por ejemplo, en Rosen *et al.*, *J. Med. Chem.* 38:4855-74 (1995). La familia de factores de transcripción C/EBP ha sido revisada en Wedel *et al.*, *Immunobiology* 193:171-85 (1995). Coactivadores y co-represores que median la regulación de la transcripción por receptores de hormonas nucleares han sido revisados, por ejemplo, en Meier, *Eur. J. Endocrinol.* 134(2):158-9 (1996); Kaiser *et al.*, *Trends Biochem. Sci.* 21:342-5 (1996); y Utley *et al.*, *Nature* 394:498-502 (1998)). Factores de transcripción GATA, que están implicados en la regulación de la hematopoyesis, se describen, por ejemplo, en Simon, *Nat. Genet.* 11:9-11 (1995); Weiss *et al.*, *Exp. Hematol.* 23:99-107. La proteína de fijación de la secuencia TATA (TBP) y sus polipéptidos TAF asociados (que incluyen TAF30, TAF55, TAF80, TAF110, TAF150 y TAF250) se describen en Goodrich & Tjian, *Curr. Opin. Cell Biol.* 6:403-9 (1994) y Hurley, *Curr. Opin. Struct. Biol.* 6:69-75 (1996). La familia STAT de factores de transcripción se revisa, por ejemplo, en Barahmand-Pour *et al.*, *Curr. Top. Microbiol. Immunol.* 211:121-8 (1996). Factores de transcripción implicados en enfermedades se revisan en Aso *et al.*, *J. Clin. Invest.* 97:1561-9 (1996).

En una realización, el dominio de represión KRAB de la proteína humana KOX-1 se utiliza como represor de la transcripción (Thiesen *et al.*, *New Biologist* 2:363-374 (1990); Margolin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 91:4509-4513 (1994); Pengue *et al.*, *Nucl. Acids Res.* 22:2908-2914 (1994); Witzgall *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 91:4514-4518 (1994)). En otra realización, KAP-1, un co-represor de KRAB, se utiliza con KRAB (Friedman *et al.*, *Genes Dev.* 10:2067-2078 (1996)). Alternativamente, puede utilizarse KAP-1 solo con una proteína con dedos de cinc. Otros factores de transcripción y dominios de factores de transcripción preferidos que actúan como represores de la transcripción incluyen MAD (véase, v.g., Sommer *et al.*, *J. Biol. Chem.* 273:6632-6642 (1998); Gupta *et al.*, *Oncogene* 16:1149-1159 (1998); Queva *et al.*, *Oncogene* 16:967-977 (1998); Larsson *et al.*, *Oncogene* 15:737-748 (1997); Laherty *et al.*, *Cell* 89:349-356 (1997); y Cultraro *et al.*, *Mol. Cell. Biol.* 17:2353-2359 (1997)); FKHR (cabeza de horquilla en el gen del rabdosarcoma; Ginsberg *et al.*, *Cancer Res.* 15:3542-3546 (1998); Epstein *et al.*, *Mol. Biol.* 18:4118-4130 (1998)); EGR-1 (producto 1 del gen de respuesta temprana al crecimiento; Yan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 95:8298-8303 (1998); y Liu *et al.*, *Cancer Gene Ther.* 5:3-28 (1998)); el dominio represor del factor represor ets2 (ERD; Sgouras *et al.*, *EMBO J.* 14:4781-4793 (1995)); y el dominio de interacción MAD smSIN3 (SID; Ayer *et al.*, *Mol. Cell. Biol.* 16:5772-5781 (1996)).

En una realización, el dominio de activación HSV VP16 se utiliza como un activador de la transcripción (véase, v.g., Hagmann *et al.*, *J. Virol.* 71:5952-5962 (1997)). Otros factores de transcripción preferidos que podrían suministrar dominios de activación incluyen el dominio de activación VP64 (Seipel *et al.*, *EMBO J.* 11:4961-4968 (1996)); receptores de hormonas nucleares (véase, v.g., Torchia *et al.*, *Curr. Opin. Cell Biol.* 10:373-383 (1998)); la subunidad p65 del factor nuclear kappa B (Bitko & Barik, *J. Virol.* 72:5610-5618 (1998) y Doyle & Hunt, *Neuroreport* 8:2937-2942 (1997)); y EGR-1 (producto 1 del gen de respuesta temprana al crecimiento; Yan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 95:8298-8303 (1998); y Liu *et al.*, *Cancer Gene Ther.* 5:3-28 (1998)).

Quinasas, fosfatasa, y otras proteínas que modifican polipéptidos implicados en la regulación génica son también útiles como dominios reguladores para las proteínas quiméricas con dedos de cinc. Tales modificadores están implicados en muchos casos en la conmutación encendido/apagado de la transcripción mediada, por ejemplo, por hormonas. Quinasas implicadas en la regulación de la transcripción se revisan en Davis, *Mol. Reprod. Dev.* 42:459-67 (1995), Jackson *et al.*, *Adv. Second Messenger Phosphoprotein Res.* 28:279-86 (1993), y Boulikas, *Crit. Rev. Eukaryot. Gene Expr.* 5:1-77 (1995), mientras que las fosfatasa se revisan, por ejemplo, en Schonthal & Semin, *Cancer Biol.* 6:239-48 (1995). Tirosina-quinasas nucleares se describen en Wang, *Trends Biochem. Sci.* 19:373-6 (1994).

Como se ha descrito, pueden obtenerse también dominios útiles a partir de los productos génicos de oncogenes (v.g., miembros de las familias myc, jun, fos, myb, max, mad, rel, ets, bcl, myb, mos) y sus factores y modificadores asociados. Oncogenes se describen, por ejemplo, en Cooper, *Oncogenes*, 2ª edición, The Jones and Bartlett Series in Biology, Boston, MA, Jones and Bartlett Publishers, 1995. Los factores de transcripción ets se revisan en Waslylk *et al.*, *Eur. J. Biochem.* 211:7-18 (1993) y Crepieux *et al.*, *Crit. Rev. Oncog.* 5:615-38 (1994). Los oncogenes myc se revisan, por ejemplo, en Ryan *et al.*, *Biochem. J.* 314:713-21 (1996). Los factores de transcripción jun y fos se describen, por ejemplo, en *The Fos and Jun Families of Transcription Factors*, Angel & Herrlich, eds. (1994). El oncogén max se revisa en Hurlin *et al.*, *Cold Spring Harb. Symp. Quant. Biol.* 59:109-16. La familia del gen myb se revisa en Kanei-Ishii *et al.*, *Curr. Top. Microbiol. Immunol.* 211:89-98 (1996). La familia del gen mos se revisa en Yew *et al.*, *Curr. Opin. Genet. Dev.* 3:19-25 (1993).

En otra realización, se utiliza histona-acetiltransferasa como activador de la transcripción (véase v.g., Jin & Scotto, *Mol. Cell. Biol.* 18:4377-4384 (1998); Wolffe, *Science* 272:371-372 (1996); Taunton *et al.*, *Science* 272:408-411 (1996); y Hassig *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 95:3519-3524 (1998)). En otra realización, se utiliza histona-desacetilasa como represor de la transcripción (véase v.g., Jin & Scotto, *Mol. Cell. Biol.* 18:4377-4384 (1998); Syntichaki & Thireos, *J. Biol. Chem.* 273:24414-24419 (1998); Sakaguchi *et al.*, *Genes Dev.* 12:2831-2841 (1998); y Martinez *et al.*, *J. Biol. Chem.* 273:23781-23785 (1998)).

Además de dominios reguladores, a menudo la proteína con dedos de cinc se expresa como una proteína de fusión tal como la proteína de fijación de maltosa ("MBP"), glutatión-S-transferasa (GST), hexahistidina, c-myc, y el epítipo FLAG, para facilidad de purificación, monitorización de la expresión, o monitorización de la localización celular y subcelular.

Ejemplos

Los ejemplos siguientes se proporcionan únicamente a modo de ilustración y no como limitación. Los expertos en la técnica reconocerán fácilmente una diversidad de parámetros no críticos que podrían cambiarse o modificarse conduciendo a resultados esencialmente similares.

Métodos

Construcción de plásmidos. Los plásmidos de expresión con dedos de cinc utilizados en estudios de transfección se construyeron por amplificación mediante PCR de segmentos de DNA que codificaban los dedos deseados del péptido Zif268 y/o el péptido NRE. Estos segmentos de DNA se insertaron en los sitios HindIII y BamHI de pCS, que se habían construido por subclonación de un dúplex de oligonucleótidos

5'-AGCTACCATGGCCAAGGAAACCGCAGCTGCCAAAT

TCGAAAGACAGCATATGGATTCTAAGCTTCGCGGATCCT-3' (SEQ ID NO: 1)

5'-CTAGAGGATCCGCGAAGCTTAGAATCCATATGCTGTCT

TTCGAATTTGGCAGCTGCGGTTTCCTTGGCCATGGT-3' (SEQ ID NO: 2)

en los sitios HinIII y XbaI de pcDNA3 (Invitrogen). Estos plásmidos de expresión se diseñaron para producir péptidos con dedos de cinc que tenían a la vez un marcador S-péptido (Kim & Raines, *Protein Sci.* 2:348-356 (1993); Kim & Raines i. 219:165-166 (1995)) y una señal de localización nuclear del antígeno T grande de SV40 (Kalderón *et al.*, *Cell* 39:499-509 (1984)) en su término N. Se construyeron plásmidos informadores por mutagénesis orientada utilizando el kit QuikChange (Stratagene). La construcción del plásmido molde (pGL3-TATA/Inr) para la mutagénesis fue descrita previamente (Kim & Pabo, *J. Biol. Chem.* 272:29795-29800 (1997)). Las secuencias de DNA de todos los constructos se confirmaron por secuenciación didesoxi.

Producción y purificación de las proteínas. Los segmentos de DNA que codificaban los péptidos Zif268, NRE, y 268//NRE se amplificaron por PCR y se subclonaron en pGEX-6P-3 (Pharmacia). Las proteínas con dedos de cinc se expresaron en *E. coli* como fusiones con glutatión-S-transferasa (GST) y se purificaron utilizando cromatografía de afinidad de acuerdo con el protocolo del fabricante. Estos constructos no tenían un marcador S-péptido o una señal de localización nuclear SV40. GST se retiró subsiguientemente por digestión con PreScissionTM Protease (Pharmacia). Las concentraciones de proteínas se estimaron utilizando electroforesis en gel SDS-poliacrilamida con seroalbúmina bovina como estándar (Pomerantz *et al.*, *Science* 267:93-96 (1995)). Las concentraciones de proteínas con dedos de cinc activos se determinaron esencialmente como se ha descrito (Rebar & Pabo, *Science* 263:671-673 (1994)). Estos dos métodos dieron resultados comparables, lo que indicaba que prácticamente la totalidad de la proteína era activa.

Ensayo de desplazamiento en gel. Las reacciones de fijación de DBNA contenían el péptido con dedos de cinc apropiado y uno o más sitios de fijación en una solución de bis-tris-propano 20 mM de pH 7,0, NaCl 100 mM, MgCl₂ 5 mM, ZnSO₄ 20 mM, 10% de glicerol, 0,1% de Nonidet P40, DTT 5 mM, y 0,10 mg/ml de seroalbúmina bovina en un volumen total de 10 ml. Todos los experimentos de fijación se realizaron a la temperatura ambiente. Las secuencias de DNA de los sitios de fijación se dan a continuación: sitio N, 5'-TCTGC AAGGGTT CA GCGACACCAACCAA-3' (SEQ ID NO: 3); sitio Z, 5'-GTGTGTGTGTGATCT GCGTGGGCG GTAAG-3' (SEQ ID NO: 4); sitio NZ, 5'-TCTGC AAGGGTTCA GCGTGGGCG GTAAG-3' (SEQ ID NO: 5); sitio N/Z, 5'-TCTGC AAGGGTTCA G GCGTGGGCG GTAAG-3' (SEQ ID NO: 6); y sitio N//Z, 5'-TCTGC AAGGGTTCA GT GCGTGGGCG GTAAG-3' (SEQ ID NO: 7). En cada caso, las secuencias de reconocimiento de 9 pb están subrayadas. Los DNAs marcados utilizados en los ensayos de desplazamiento en gel se prepararon por extensión Klenow o reacción de quinasas.

Para determinar las constantes de disociación, se incubaron diluciones seriadas tres veces del péptido Zif268 o NRE con un DNA sonda marcado (0,4-1,4 pM) a la temperatura ambiente durante 1 hora, y las mezclas de reacción se sometieron luego a electroforesis en gel. Las señales radiactivas se cuantificaron por análisis Phosphorimager; las constantes de disociación aparentes se determinaron como se ha descrito (Rebar & Pabo, *Science* 263:671-673 (1994)).

Las tasas de asociación y disociación se determinaron también por ensayo de desplazamiento en gel. Para iniciar la reacción de fijación cuando se determinaban las constantes de asociación, se añadió un DNA sonda mar-

ES 2 341 926 T3

cado (concentración final, ~ 0,4 pM) al péptido con dedos de cinc (concentración final, 5-10 pM) a la temperatura ambiente, y se analizaron en diversos momentos (0-20 min) partes alícuotas por electroforesis en gel. La fracción fijada en el tiempo t se determinó por análisis Phosphorimager de los geles. Los datos se ajustaron luego (programa KaleidaGraph™ (Synergy Software)) a la ecuación:

$$F = F_{\text{final}} [1 - \exp(-k_{\text{obs}} \times t)]$$

donde F es la fracción fijada en el tiempo t; F_{final} es la fracción calculada fijada a la terminación de la reacción; y K_{obs} es la constante de velocidad (Hoopes *et al.*, *J. Biol. Chem.* 267:11539-11547 (1992)). La constante de tasa de asociación se calculó a partir de la ecuación:

$$k_{\text{on}} = (F_{\text{final}} \times k_{\text{obs}})/[P]$$

donde [P] es la concentración de la proteína con dedos de cinc. Las constantes de tasa de disociación se determinaron esencialmente como se ha descrito (Kim *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 94:3616-3620 (1997)). Las proteínas (concentración final, 100 pM) se preincubaron con un DNA sonda marcado durante 1 hora y se añadió luego un gran exceso de DNA sonda sin marcar (concentración final, 20 nM). Se retiraron partes alícuotas en diversos momentos y se analizaron por electroforesis en gel. La fracción de sitio marcado se normalizó a la fracción encontrada al final del periodo de preincubación de 1 hora. Se representó el logaritmo natural de la fracción normalizada fijada en función del tiempo, y se determinó la tasa de disociación a partir de la pendiente. Todos los puntos de datos para medidas rápidas de la tasa de asociación y la tasa de disociación se corrigieron por el tiempo muerto de electroforesis.

Estudios de fijación de competición. El péptido 268//NRE (concentración final, 5 pM) se incubó primeramente durante 1 hora con diversas cantidades de un DNA competidor frío (0, 0,05, 0,5, 5, y 50 nM), y a continuación se añadió el sitio marcado N/Z (6-8 pM). Se analizaron muestras por electroforesis en gel después de 2, 24, 48, 96, 190 y 600 horas. Las relaciones de especificidad ($K_{\text{dc}}/K_{\text{d}}$) se calcularon a partir de la ecuación:

$$K_{\text{dc}}/K_{\text{d}} = \{[C]/[P]_{\text{t}}\} \times (F_0 \times F)/(F_0 - F)(1 - F)$$

donde K_{dc} es la constante de disociación para la fijación al DNA competidor; K_{d} es la constante de disociación para la fijación al sitio quimérico intacto; [C] es la concentración de DNA competidor; $[P]_{\text{t}}$ es la concentración total de la proteína; F_0 es la fracción fijada en ausencia del DNA competidor; y F es la fracción fijada en presencia del DNA competidor. Esta ecuación supone que la concentración de proteína libre es significativamente menor que la de proteína fijada a DNA. Este criterio debe satisfacerse fácilmente, dado que el valor K_{d} del péptido 268//NRE en el sitio N/Z es 3,8 fM, y se utilizaron 5 pM del péptido de fusión en estos experimentos de competición.

Experimentos de competición con DNA de esperma de salmón contenían el péptido 268//NRE o Zif268 (200 pM), el sitio marcado N/Z, y un ligero exceso molar de sitio N/Z sin marcar. Se añadieron cantidades variables de DNA de esperma de salmón, y las muestras se analizaron por electroforesis en gel al cabo de 2, 24, y 48 horas de incubación. Cuando se calcularon relaciones de especificidad, se supuso que cada base en el DNA de esperma de salmón representa el comienzo de un sitio de fijación potencial (inespecífico).

Ensayo de cotransfección transitoria. Las células 293 se transfectaron por precipitación con fosfato de calcio con un choque de glicerol como se ha descrito (Cepek *et al.*, *Genes Dev.* 10:2079-2088 (1996)). Los experimentos de transfección utilizaron típicamente células a 10-30% de confluencia en cultivos monocapa (placas de 6 pocillos), y se añadieron los plásmidos siguientes: 0,2 mg del plásmido de expresión vacío (pCS) o de plásmidos de expresión que codificaban proteínas con dedos de cinc; 0,2 mg de un plásmido informador; 1 mg de plásmido activador (GAL4-VP16); 0,1 mg de plásmido de expresión de β -galactosidasa (pCMVb; Clontech); y 2,5 mg de plásmido portador (pUC19). Las actividades de luciferasa y β -galactosidasa en las células transfectadas se midieron como se ha descrito (Kim *et al.*, *Proc. Natl. Acad. Sci. USA* 94:3616-3620 (1997); Kim & Pabo, *J. Biol. Chem.* 272:29795-29800 (1997)). Todos los péptidos con dedos de cinc expresados en las células 293 se cuantificaron utilizando el kit S. Tag™ Rapid Assay (Novagen) (Kim & Raines, *Protein Sci.* 2:348-356 (1993); Kim & Raines, *Anal. Biochem.* 219:165-166 (1995)).

Resultados

Diseño de péptidos con dedos de cinc múltiples basado en estructura. La estrategia de diseño implicaba enlazar dos péptidos de 3 dedos de cinc, utilizando enlazadores más largos (no canónicos) en la unión para evitar la introducción de tensión alguna. Para reducir ulteriormente cualquier riesgo de interferencia o colisión entre los dedos, los enlazadores se diseñaron de modo que pudieran acomodar sitios de fijación compuestos con uno o dos pares de bases adicionales insertados entre los sitios de fijación de nueve pares de bases individuales. Los estudios consignados en este documento utilizaron el péptido de 3 dedos Zif268 (que reconoce el sitio 5'-GCG TGG GCG-3'; SEQ ID NO: 8) y un péptido con 3 dedos "NRE" (una variante de Zif268 seleccionada previamente por presentación de fago) que se fija fuerte

y específicamente a parte de un elemento de respuesta de hormona nuclear (5'-AAG GGT TCA-3'; SEQ ID NO: 9) (Greisman & Pabo, *Science* 275:657-661 (1997)). El sitio diana compuesto con un par de bases adicionales en el centro tiene la secuencia 5'-AAG GGT TCA G GCG TGG GCG-3' (SEQ ID NO: 10) y se conoce como el sitio N/Z (N denota el sitio de fijación para el péptido NRE, y Z el sitio de fijación para Zif268). El sitio con dos pares de bases adicionales en el centro tiene la secuencia 5'-AAG GGT TCA GT GCG TGG GCG-3' (SEQ ID NO: 11) y se conoce como el sitio N//Z.

El diseño basado en estructura, con el complejo Zif268 (Pavletich & Pabo, *Science* 252:809-817 (1991); Elrod-Erickson *et al.*, *Structure* 4:1171-1180 (1996)) como modelo, se utilizó para determinar la longitud apropiada de enlazadores para la construcción de proteínas con dedos múltiples que podrían reconocer cada sitio de fijación (véanse las Figuras 1 y 2). En el sitio N/Z, se encontró que un contenido de ocho residuos entre el residuo Leu en el extremo de la hélice a del primer péptido y el residuo Tyr en la primera hoja b del péptido siguiente podría permitir una flexibilidad suficiente. Un enlazador canónico "TGEKP" tiene 4 residuos (a saber, Gly-Glu-Lys-Pro) en esta región. En el sitio N//Z, parecía razonable utilizar 11 residuos entre el Leu y el Tyr (Fig. 1A). Cada enlazador (Fig. 1A) contenía secuencias que flanquean naturalmente el término N y el término C del péptido de 3 dedos Zif268. Para permitir una flexibilidad adicional, se incluyó una glicina en el enlazador más corto (que es todavía 4 residuos más largo que un enlazador canónico), y se incluyó una secuencia Gly-Gly-Gly-Ser en el enlazador más largo (que es 7 residuos más largo que un enlazador canónico). Utilizando una notación análoga a la empleada para los sitios de fijación, la proteína de fusión con el enlazador más corto se designa como 268/NRE y la proteína de fusión con el enlazador más largo se designa como 268//NRE.

Ensayos de desplazamiento en gel para determinar las constantes de disociación y semividas de los complejos proteínas-DNA. Los péptidos con dedos de cinc Zif268, NRE, y 268//NRE se expresaron y purificaron a partir de *E. coli*, y se utilizaron en varias series de experimentos de desplazamiento en gel. Una serie preliminar de experimentos se diseñó simplemente para determinar si dos proteínas de 3 dedos podrían fijarse en sitios adyacentes de 9 pb (cualquier interferencia en la fijación de los péptidos no enlazados podría reducir la afinidad de una proteína con dedos múltiples para los sitios compuestos). Los primeros experimentos utilizaron un fragmento de DNA (al que se hace referencia como el sitio NZ) con los sitios de fijación NRE y Zif268 directamente yuxtapuestos (5'-AAG GGT TCA GCG TGG GCG-3'; SEQ ID NO: 12). Cantidades variables del péptido NRE se incubaron con el sitio marcado NZ en presencia o ausencia de Zif268 (Figura 3). Se determinó que el péptido NRE de 3 dedos se fija en realidad con una fuerza ligeramente mayor al sitio NZ con Zif268 prefijado que al sitio libre. La constante de disociación aparente (K_d) del péptido NRE es 180 pM cuando el mismo se fija solo, pero 60 pM cuando Zif268 está prefijado al sitio cercano. Se obtuvieron resultados similares en el sitio N/Z. Estos experimentos demuestran que no existe colisión alguna entre los péptidos fijados en sitios adyacentes y sugieren que puede existir incluso cierto efecto cooperativo moderado. Parece ser que los límites previos en la afinidad de las proteínas con dedos múltiples (Rebar (Tesis Ph.D.), *Selection Studies of Zinc Finger-DNA Recognition*, Massachusetts Institute of Technology (1997); Shi, (Tesis Ph.D.), *Molecular Mechanisms of Zinc Finger Protein-Nucleic Acid Interactions*, Johns Hopkins University (1995); Liu *et al.*, *Proc. Natl. Acad. Sci. USA* 94:5525-5530 (1997)) se debían a problemas con el diseño de los enlazadores.

Una segunda serie de estudios de fijación confirma la eficacia del nuevo diseño de enlazadores. Titulaciones de equilibrio demuestran que el péptido 268//NRE tiene una afinidad significativamente mayor para los sitios compuestos que para los sitios de 9 pb individuales (Tabla 1). La proteína de fusión se fija a los sitios de 9 pb aislados con valores K_d s similares a los del péptido NRE (180 pM) y el péptido Zif268 (14 pM) para sus sitios de fijación. En contraste, la proteína de fusión 268//NRE fija los sitios compuestos tan fuertemente que las constantes de disociación son demasiado pequeñas para poder ser determinadas fácilmente por titulación de la proteína. Al menos 0,4 pM de DNA sonda marcado eran necesarios en estos experimentos de desplazamiento en gel, haciendo difícil la determinación exacta de valores K_d de < 1 pM. Dadas estas dificultades técnicas, se decidió medir la tasa de asociación y la tasa de disociación para la fijación del péptido 268//NRE y utilizar estas tasas para estimar la constante de fijación de equilibrio (Tabla 1). Estudios paralelos con los péptidos de 3 dedos proporcionaron controles útiles. Las tasas de asociación para los péptidos 268//NRE, NRE, y Zif268 eran rápidas y estaban próximas al límite controlado por difusión (108 a 109 M⁻¹s⁻¹) (Von Hippel & Berg, *I.* 264:675-678 (1989)). Las tasas de disociación exhibían diferencias asombrosas: los péptidos de 3 dedos tienen semividas de < 39 segundos, mientras que el péptido 268//NRE tiene una semivida de 370 horas en el sitio NZ. Estudios de control demuestran que el péptido 268//NRE forma un complejo mucho menos estable con un solo sitio de 9 pb simple (así, la semivida = 150 segundos en el sitio N). Tanto los dedos NRE como los dedos Zif268 tienen que fijarse a sus subsitios respectivos de 9 pb para formar el complejo extraordinariamente estable observado con el péptido 268//NRE en el sitio NZ.

En todos los casos en que pudieron realizarse medidas en paralelo, los valores K_d calculados a partir de la relación de las constantes cinéticas (K_{off}/K_{on}) estaban en concordancia satisfactoria con los determinados a partir de estudios de equilibrio (Tabla 1). Esto proporcionó confianza en la utilización de los datos cinéticos para determinar K_d s en los casos en que la titulación directa era inviable. Los cálculos demuestran que el péptido 268//NRE tiene afinidad femtomolar para los sitios de fijación compuestos, con un K_d de $2,1 \times 10^{-15}$ M (2,1 fM) en el sitio NZ, 3,7 fM en el sitio N/Z, y 3,0 fM en el sitio N//Z (la consistencia de estos 3 valores K_d s es también estimulante dado que podía esperarse que el enlazador más largo, flexible, debería acomodar fácilmente cualquiera de estas separaciones). Los datos demuestran que el nuevo diseño de enlazador es muy eficaz: el péptido de fusión 268//NRE se fija mucho más fuertemente (5.000-95.000 veces) al sitio compuesto que a los sitios individuales de 9 pb, y se fija mucho más fuertemente (6.000-90.000 veces) que cualquiera de los péptidos de 3 dedos originales.

TABLA 1

Constantes de Disociación y datos de Tasa

Proteína	Sitio de fijación	K_d , pM	$k_{on} \cdot M^{-1} s^{-1}$	$k_{off} \cdot s^{-1}$
268//NRE	N	190 ± 50	$2.5 \pm 0.4 \times 10^8$	$4.7 \pm 2.9 \times 10^{-7}$
268//NRE	Z	10*		
268//NRE	NZ	$<1.0^\dagger$	$2.5 \pm 0.2 \times 10^8$	$5.2 \pm 0.9 \times 10^{-7}$
268//NRE	N/Z	$<1.0^\dagger$	$2.5 \pm 0.2 \times 10^8$	$9.2 \pm 0.7 \times 10^{-7}$
268//NRE	N//Z	$<1.0^\dagger$	$2.6 \pm 0.6 \times 10^8$	$7.7 \pm 1.3 \times 10^{-7}$
NRE	N/Z	180 ± 43	$>7.3 \times 10^7$	$>5.9 \times 10^{-7}$
Zif268	NZ	12 ± 3		
Zif268	N/Z	14 ± 4	$>7.0 \times 10^8$	$1.4 \pm 0.4 \times 10^{-7}$
Zif268	N//Z	14 ± 1		

Todas las constantes se determinaron en al menos dos experimentos separados, y se indica el SEM.

* No pudo determinarse un valor K_d exacto debido a que este complejo daba una banda sucia en los geles.

† Como se explica en el texto, estos valores K_d no pudieron medirse directamente. La estimación de K_d a partir de la relación k_{off}/K_{on} da valores de 2,1 fM en el sitio NZ, 3,7 fM en el sitio N/Z y 3,0 fM en el sitio N//Z.

Se utilizaron también experimentos de competición para estudiar adicionalmente la afinidad y especificidad del péptido de 6 dedos 268//NRE (Figura 4A). Una serie de experimentos testaron directamente con qué bondad podrían competir los sitios N y Z de 9 pb con el sitio compuesto N/Z en cuanto a fijación al péptido de fusión. En estos experimentos, se mezclaron cantidades variables de sitio N o Z frío con una cantidad limitante del péptido 268//NRE. Después de 1 hora de incubación, se añadió un ligero exceso molar (con relación a la cantidad total de proteína de fusión) de sitio N/Z marcado. En estas condiciones, aproximadamente el 70% del DNA marcado es desplazado en ausencia de DNA competidor. Muestras tomadas en diversos momentos se analizaron por electroforesis en gel. Dado que la concentración del péptido 268//NRE en este experimento (5 pM) es varios órdenes de magnitud mayor que la constante de disociación del péptido para el sitio N/Z, prácticamente la totalidad del péptido se fija al sitio N/Z cuando no se añade DNA competidor alguno. Cualquier disminución en la cantidad de sitio N/Z desplazado en presencia de DNA competidor refleja la fijación del péptido 268//NRE al sitio de competición.

La equilibración en estos experimentos requiere cientos de horas, y la estabilidad de la proteína purificada resulta realmente un problema importante (el sitio compuesto se añade en último lugar, y la equilibración requiere largo tiempo dado que la proteína de fusión puede encontrar sitios Z fríos centenares o millares de veces antes que encuentre

por primera vez un sitio N/Z marcado). Después de pre-equilibración con concentraciones altas del sitio N o Z frío, se determinó que la fracción de marcador N/Z desplazada aumenta continuamente con tiempos de incubación crecientes de hasta aproximadamente 600 horas. Después de 600 horas de incubación, una fracción significativa del sitio N/Z marcado es desplazada incluso en presencia de un exceso molar de 10.000 veces del sitio N o Z frío. Las relaciones de especificidad (calculadas como se ha descrito arriba) indican que el péptido 268//NRE prefiere el sitio compuesto sobre el sitio N por un factor de al menos $3.800 + 1.600$ y que el péptido de fusión prefiere el sitio compuesto sobre el sitio Z por un factor de al menos $320 + 44$. Estos experimentos confirman directamente la notable especificidad del péptido de 6 dedos, pero estos valores son sólo límites inferiores en las relaciones de especificidad. La muestra de proteína pierde algo de su actividad durante el largo tiempo de incubación requerido por estos experimentos (la actividad de la proteína libre tiene una semivida de aproximadamente 2 días en estas condiciones), y la proteína desnaturalizada no tendrá nunca oportunidad de desplazar el sitio marcado N/Z.

Se utilizaron experimentos de competición con DNA de esperma de salmón para estimar la relación de las constantes de fijación específica/inespecífica para el péptido 268//NRE (Figura 4B). Estos experimentos demostraron que el péptido 268//NRE discrimina muy eficazmente contra el DNA inespecífico e indican una relación de especificidad (K_{ds}/K_d) de $8,8 + 1,5 \times 10^6$. Experimentos paralelos con el péptido de 3 dedos Zif268 dan una relación de especificidad de $1,2 + 0,1 \times 10^5$. Estudios previos, utilizando DNA de timo de ternero como competidor y condiciones ligeramente diferentes, habían dado una relación de especificidad de $0,31 \times 10^5$ para el péptido Zif268 (Greisman & Pabo, *Science* 275:657-661 (1997)). Tomados en su conjunto, los datos respecto a afinidad y especificidad del péptido de fusión de 6 dedos 268//NRE sugerían que el mismo podría servir como represor muy eficaz e indicaban ciertamente que podría ser un excelente candidato para análisis ulterior *in vivo*.

Se utilizaron estudios de cotransfección transitoria en la línea de células humana 293 para ver si los nuevos péptidos con dedos múltiples podrían reprimir eficazmente la transcripción de genes informadores. En un estudio previo, se había demostrado que el péptido Zif268 podría reprimir eficazmente la transcripción tanto basal como activada por VP16 cuando el péptido Zif268 se fijaba a un sitio próximo a la secuencia TATA o al elemento iniciador (Kim & Pabo, *J. Biol. Chem.* 272: 29795-29800 (1997)). En este estudio actual, se utilizaron un informador de luciferasa y constructos promotores similares en los cuales se incorporaban sitios de fijación apropiados (Z, N, N/Z, y N//Z) en posiciones comparables próximas al elemento iniciador (Fig. 1B).

Se determinó que el péptido 268//NRE proporciona una represión 72 veces mayor de la transcripción activada por VP16 en un promotor que contiene el sitio N/Z y una represión 47 veces mayor en un promotor que contiene el sitio N//Z (Figs. 5A-5C). El péptido 268//NRE proporciona una represión 68 veces mayor en el sitio N/Z. Claramente, estos péptidos de fusión son represores muy eficaces en sitios con las separaciones apropiadas. Experimentos paralelos con los péptidos de 3 dedos indican represión, pero indican que los mismos son considerablemente menos eficaces que los péptidos de fusión. Así, el péptido NRE proporciona una represión 1,9 veces mayor con un sitio N en el promotor; represión 1,8 veces mayor con un sitio N/Z; represión 2,7 veces mayor con un sitio N//Z; y no indica represión alguna con un sitio Z aislado. El péptido Zif268 proporciona una represión 13 veces mayor respecto al promotor Z; represión 8,9 veces mayor respecto al promotor N/Z; represión 15 veces mayor respecto al promotor N//Z; y no proporciona represión alguna con un sitio N aislado. Experimentos adicionales demuestran que es necesario acoplamiento covalente para alcanzar los niveles de represión mucho mayores obtenidos con las proteínas de fusión en el sitio N/Z.

Así pues, la co-expresión de los péptidos Zif268 y NRE como cadenas polipeptídicas separadas (por inclusión de ambos plásmidos de expresión en los ensayos de cotransfección) proporciona solamente una represión 8,5 veces mayor en el sitio N/Z, un nivel comparable (dentro del valor experimental) a la represión 8,9 veces mayor obtenida en este sitio con el péptido Zif268 aislado. Esto es mucho menor que la represión 68 veces y 72 veces mayor que proporcionan las proteínas de fusión 268//NRE y 268//NRE en el sitio N/Z, y está claro que estos efectos "sinérgicos" requieren enlace covalente.

Debe indicarse que los dedos adicionales en los péptidos de fusión pueden tener ciertos efectos represivos moderados incluso en casos en los que solamente 3 de los dedos pueden fijarse específicamente. Así, los péptidos de 6 dedos (268//NRE y 268//NRE) proporcionan una represión de 21 a 23 veces mayor del promotor Z. Un nivel de represión similar (22 veces mayor) se obtiene con el péptido 268//NRE en el sitio N//Z. La modelización sugiere que el enlazador es demasiado corto para permitir la fijación específica de la totalidad de los 6 dedos en este sitio. Estos niveles de represión son coherentemente algo mayores que el nivel observado con el péptido Zif268 aislado en el sitio Z (represión 13 veces mayor). Parece posible (cuando el péptido 268//NRE se fija al sitio Z) que 1) los dedos NRE estén libres e interfieran todavía estéricamente con el ensamblaje del complejo de transición, o que 2) los dedos NRE establezcan contactos inespecíficos débiles con el DNA y aumenten por tanto ligeramente la estabilidad del complejo. Estudios ulteriores indican que todos los péptidos se expresan a niveles comparables.

Los péptidos con dedos de cinc expresados en las células 293 tenían un marcador S-péptido, y la cantidad de péptido se cuantificó utilizando un ensayo de ribonucleasas después de activación con proteína S (Kim & Raines, *Protein Sci.* 2:348-356 (1993); Kim & Raines *Anal. Biochem.* 219:156-166 (1995)). Una estimación conservadora indica que los niveles de expresión de los péptidos en las células son significativamente mayores (al menos 100 veces) que las constantes de disociación de los péptidos de 3 dedos. Se construyeron también plásmidos que podrían codificar variantes de 4 y 5 dedos de los péptidos 268//NRE y 268//NRE. Estos se testaron en estudios de transfección de cultivos de tejido, y proporcionaron típicamente niveles de represión intermedios entre los obtenidos con los péptidos de 3 dedos y los obtenidos con los péptidos de 6 dedos (Figs. 5A-5D).

REIVINDICACIONES

1. Un método de construcción de una proteína con dedos de cinc que se fija a sitios diana adyacentes, comprendiendo el método los pasos de
 - (i) seleccionar un primer y un segundo polipéptido de dominio de fijación de DNA de la proteína con dedos de cinc, en donde cada uno de los dominios comprende un polipéptido con dedos de cinc, y en donde el primer dominio se fija a un primer sitio diana y el segundo dominio se fija a un segundo sitio diana, sitios diana que son sitios diana adyacentes y no solapantes que están separados por 0 a 3 pares de bases;
 - (ii) seleccionar un enlazador flexible en el cual, si los sitios diana adyacentes están separados por cero nucleótidos, el enlazador flexible tiene una longitud de seis aminoácidos o en donde, si los sitios diana adyacentes están separados por un solo nucleótido, el enlazador flexible tiene una longitud de siete, ocho o nueve aminoácidos o en donde, si los sitios diana adyacentes están separados por dos nucleótidos, el enlazador flexible tiene una longitud de diez, once o doce aminoácidos, o en donde, si los sitios diana adyacentes están separados por 3 nucleótidos, el enlazador flexible tiene una longitud de 12 aminoácidos; y
 - (iii) fusionar los dominios primero y segundo con el enlazador flexible, produciendo de este modo una proteína quimérica con dedos de cinc que se fija a sitios diana adyacentes.
2. El método de la reivindicación 1, en donde el enlazador flexible tiene la secuencia de aminoácidos RQKDGGERP.
3. El método de la reivindicación 1, en donde el enlazador flexible tiene la secuencia de aminoácidos RQKDGGGSERP.
4. El método de la reivindicación 1, en donde el polipéptido con dedos de cinc se selecciona del grupo constituido por Zif268 y NRE.
5. El método de la reivindicación 1, en donde los polipéptidos con dedos de cinc son heterólogos.
6. El método de la reivindicación 1, en donde la proteína con dedos de cinc tiene afinidad femtomolar para los sitios diana adyacentes.
7. El método de la reivindicación 6, en donde la proteína con dedos de cinc tiene aproximadamente una afinidad 2-4 femtomolar para los sitios diana adyacentes.
8. El método de la reivindicación 1, en donde la proteína con dedos de cinc comprende adicionalmente un polipéptido de dominio regulador.
9. Una proteína con dedos de cinc que se fija a sitios diana adyacentes, comprendiendo la proteína con dedos de cinc:
 - (i) un primer y un segundo polipéptido de dominio de fijación de DNA de la proteína con dedos de cinc, en donde cada uno de los dominios comprende un polipéptido con dedos de cinc, y en donde el primer dominio se fija a un primer sitio diana y el segundo dominio se fija a un segundo sitio diana, sitios diana que son sitios diana adyacentes y no solapantes que están separados por 0 a 3 pares de bases; y
 - (ii) un enlazador flexible; en donde el primer y el segundo dominios están fusionados con el enlazador flexible, en donde, si los sitios diana adyacentes están separados por cero nucleótidos, el enlazador flexible tiene una longitud de seis aminoácidos, o en donde, si los sitios diana adyacentes están separados por un solo nucleótido, el enlazador flexible tiene una longitud de siete, ocho o nueve aminoácidos o en donde, si los sitios diana adyacentes están separados por dos nucleótidos, el enlazador flexible tiene una longitud de diez, once o doce aminoácidos de longitud o en donde, si los sitios diana adyacentes están separados por tres nucleótidos, el enlazador flexible tiene una longitud de doce aminoácidos.
10. La proteína con dedos de cinc de la reivindicación 9, en donde el enlazador flexible tiene la secuencia de aminoácidos RQKDGGERP.
11. La proteína con dedos de cinc de la reivindicación 9, en donde el enlazador flexible tiene la secuencia de aminoácidos RQKDGGGSERP.
12. La proteína con dedos de cinc de la reivindicación 9, en donde el polipéptido con dedos de cinc se selecciona del grupo constituido por Zif268 y NRE.

ES 2 341 926 T3

13. La proteína con dedos de cinc de la reivindicación 9, en donde los polipéptidos con dedos de cinc son heterólogos.

5 14. La proteína con dedos de cinc de la reivindicación 9, en donde la proteína con dedos de cinc tiene afinidad femtomolar para los sitios diana adyacentes.

15. La proteína con dedos de cinc de la reivindicación 14, en donde la proteína con dedos de cinc tiene aproximadamente una afinidad 2,4-femtomolar para los sitios diana adyacentes.

10 16. La proteína con dedos de cinc de la reivindicación 9, en donde la proteína con dedos de cinc comprende adicionalmente un polipéptido de dominio regulador.

17. Un ácido nucleico aislado que codifica la proteína con dedos de cinc de la reivindicación 9.

15

20

25

30

35

40

45

50

55

60

65

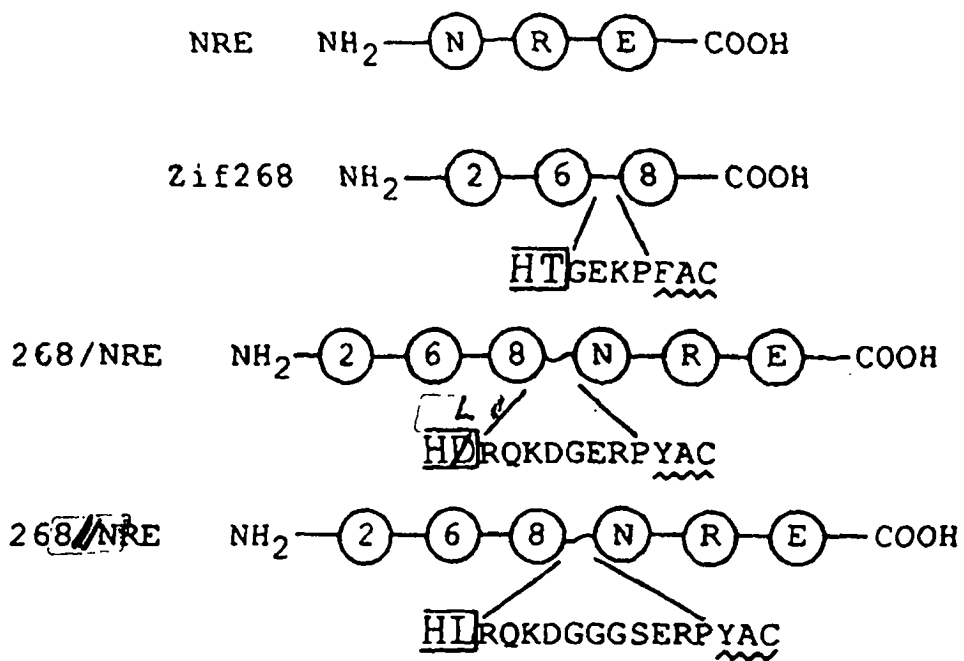


FIG. 1A.

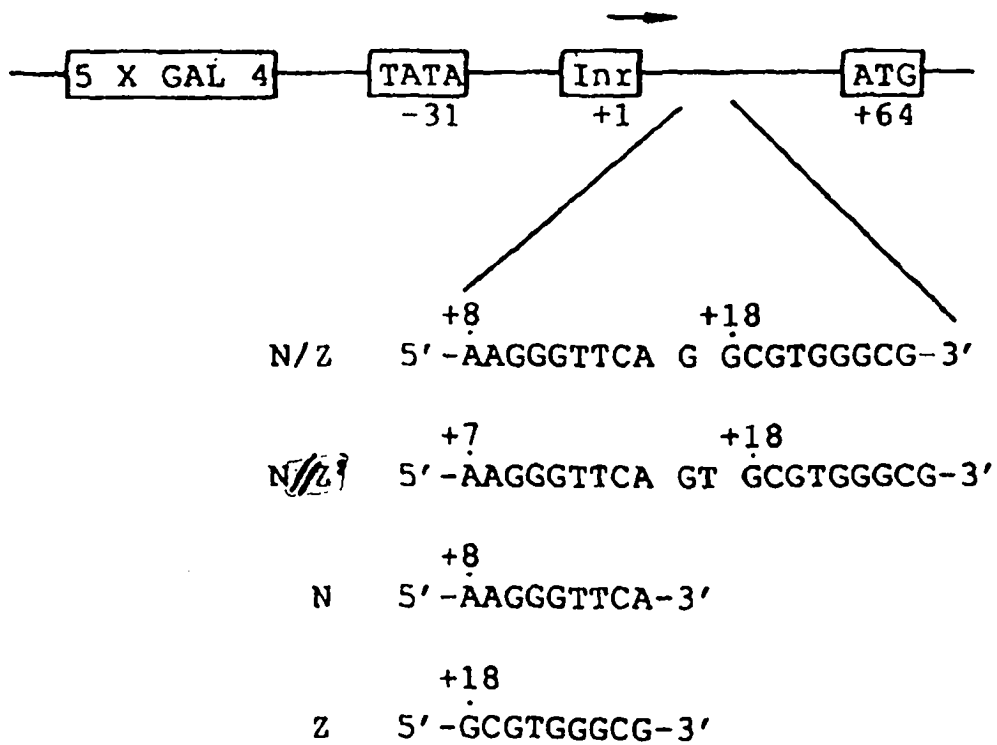


FIG. 1B.

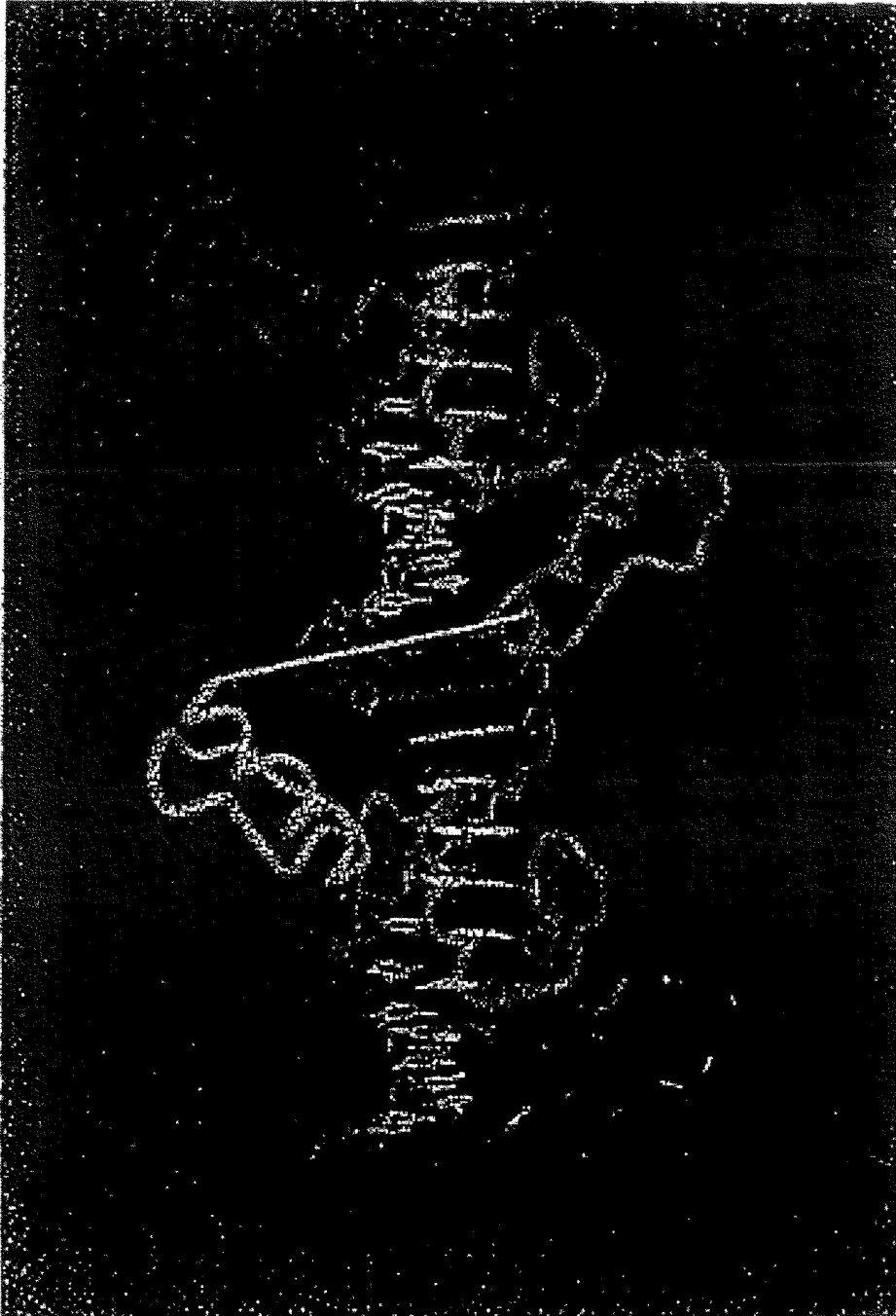


FIG. 2.

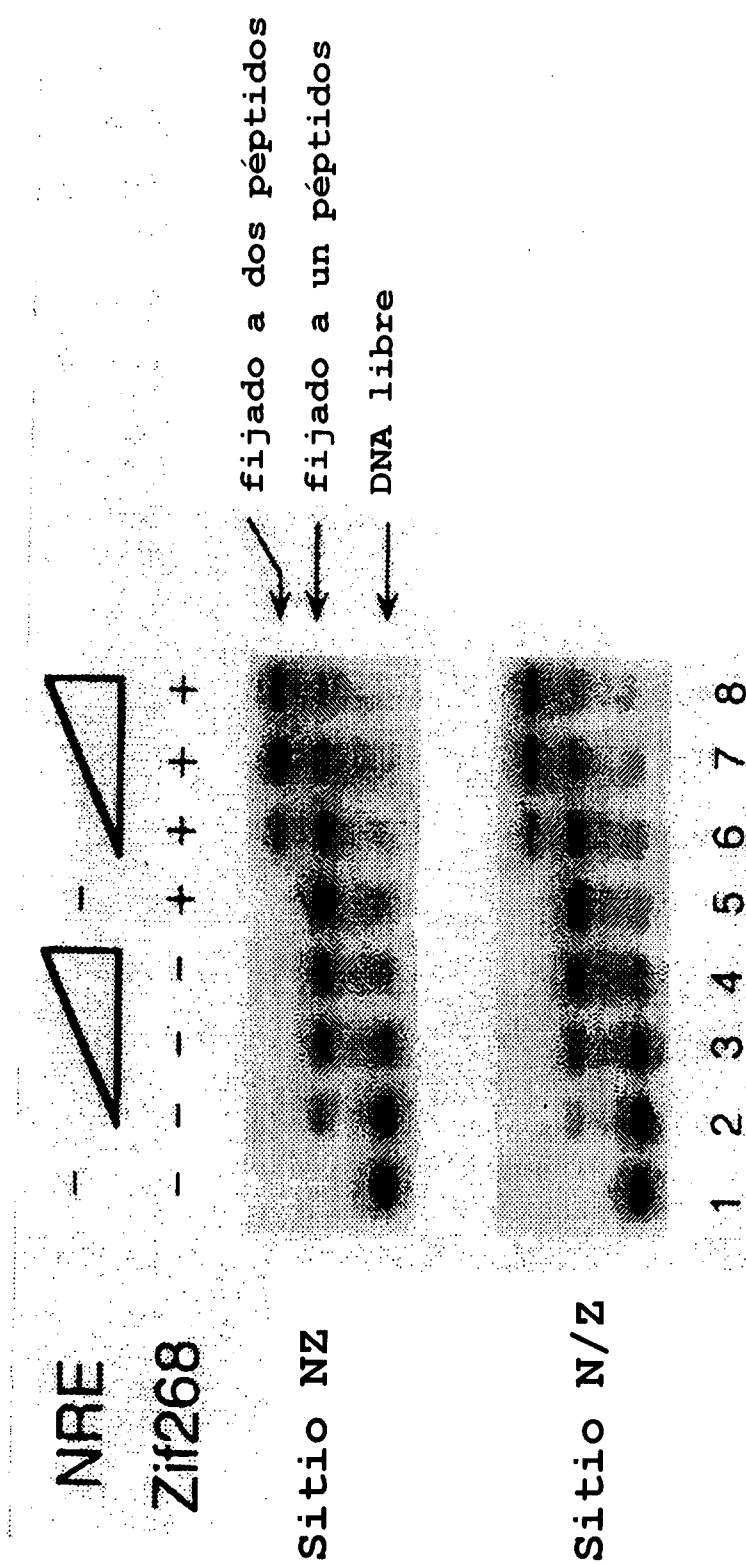


FIG. 3.

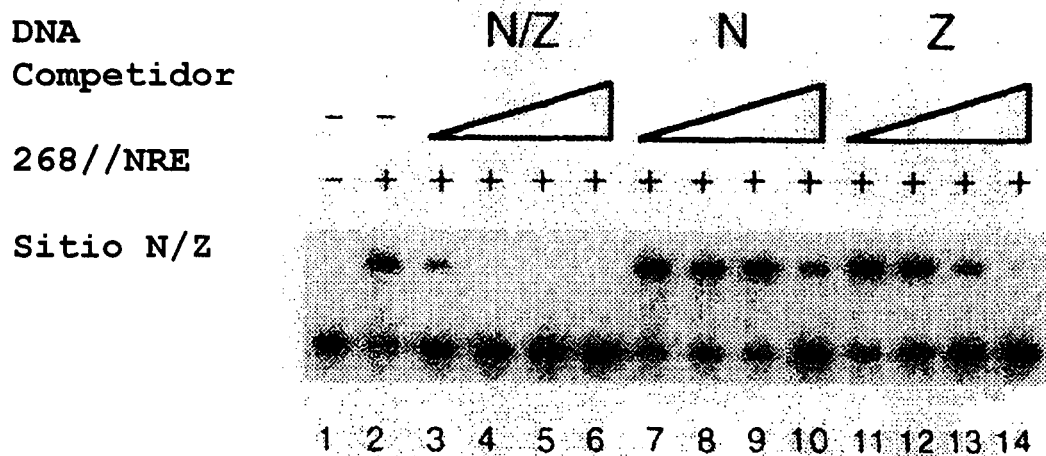


FIG. 4A.

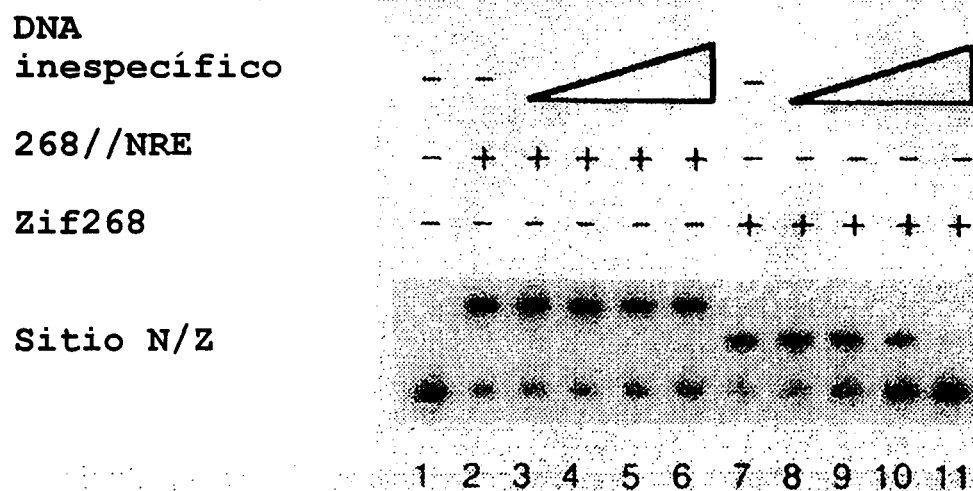
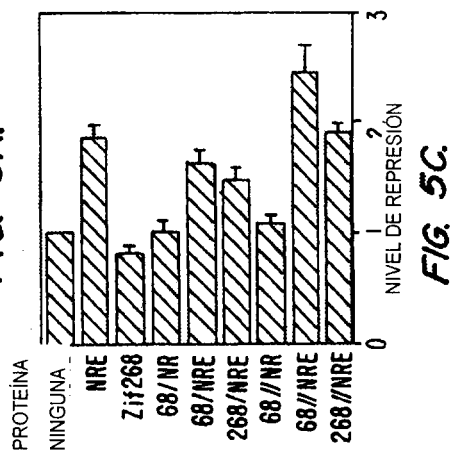
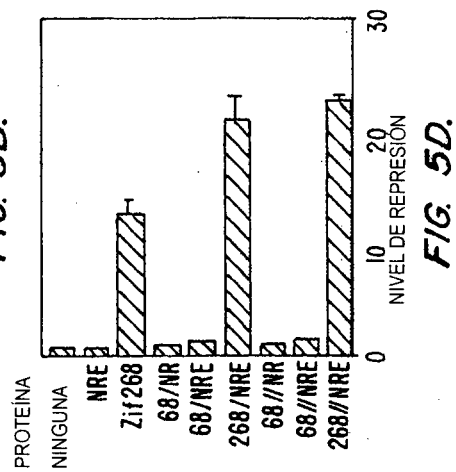
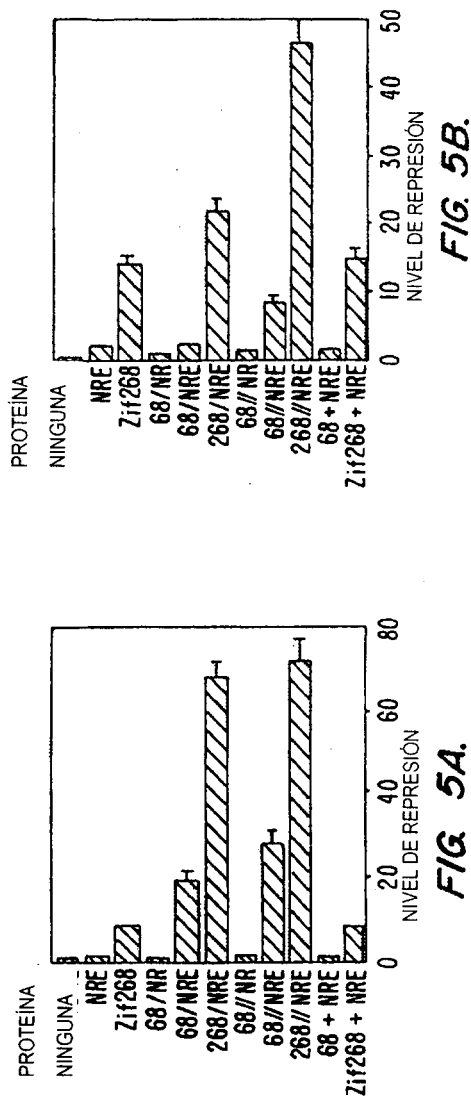


FIG. 4B.



ES 2 341 926 T3

LISTA DE SECUENCIAS

<110> Massachusetts Institute of Technology

⁵ <120> Poliproteínas con dedos de cinc que tienen enlazadores mejorados

<130> SMK/FP5882253

10 <140> EP 99909701.7

<141> 1999-03-01

<150> PCT/US99/04441

15 <151> 1999-03-01

<150> US 60/076,454

<151> 1998-03-02

<160> 24

<170> PatentIn Ver. 2.0

25

 $\langle 210 \rangle$ 1

<211> 74

<212> DNA

30

<213> Secuencia Artificial

 $\langle 220 \rangle$

35 <223> Descripción de la Secuencia Artificial: dúplex de oligonucleótidos

<400> 1

40

agctaccatg gccaaaggaaa ccgcagctgc caaattcgaa agacagcata tggattctaa 60
gcttcgcgga tcct 74

<210> 2

<211> 74

45

<212> DNA

<213> Secuencia Artificial

 $\langle 220 \rangle$

50

<223> Descripción de la Secuencia Artificial: dúplex de oligonucleótidos

<400> 2

55

ctagaggatc cgcgaagctt agaatccata tgcctgtcttt cgaatttggc agctgcgggt 60
tccttggcca tggt 74

<210> 3

60

<211> 29

<212> DNA

<213> Secuencia Artificial

65

 $\langle 220 \rangle$

<223> Descripción de la Secuencia Artificial: Sitio de fijación de dedo de cinc de sitio N

ES 2 341 926 T3

	<220>	
	<221> Proteína_fijación	
	<222> (6)..(14)	
5	<223> Secuencia de reconocimiento de dedo de cinc de sitio N	
	<400> 3	
10	tctgcaagggttcaggcgac accaaccaa	29
	<210> 4	
	<211> 29	
15	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
20	<223> Descripción de la Secuencia Artificial: Sitio de fijación de dedo de cinc de sitio Z	
	<220>	
	<221> Proteína_fijación	
25	<222> (16)..(24)	
	<223> Secuencia de reconocimiento de dedo de cinc de sitio Z de 9 pb	
	<400> 4	
30	gtgtgtgtgtgatctgcgtggcggttaag	29
	<210> 5	
35	<211> 28	
	<212> DNA	
	<213> Secuencia Artificial	
40	<220>	
	<223> Descripción de la Secuencia Artificial: Sitio de fijación de dedo de cinc de sitio NZ	
45	<220>	
	<221> Proteína_fijación	
	<222> (6)..(14)	
	<223> Secuencia de reconocimiento de dedo de cinc de sitio N de 9 pb	
50	<220>	
	<221> Proteína_fijación	
	<222> (15)..(23)	
55	<223> Secuencia de reconocimiento de dedo de cinc de sitio Z de 9 pb	
	<400> 5	
60	tctgcaagggttcagcgtggcggttaag	28
	<210> 6	
	<211> 29	
65	<212> DNA	
	<213> Secuencia Artificial	

ES 2 341 926 T3

	<220>	
	<223> Descripción de la Secuencia Artificial: Sitio de fijación de dedo de cinc de sitio N/Z	
5	<220>	
	<221> Proteína_fijación	
	<222> (6)..(14)	
	<223> Secuencia de reconocimiento de dedo de cinc de sitio N de 9 pb	
10	<220>	
	<221> Proteína_fijación	
	<222> (16)..(24)	
15	<223> Secuencia de reconocimiento de dedo de cinc de sitio Z de 9 pb	
	<400> 6	
20	tctgcaaggg ttcaggcgtg ggcggttaag	29
	<210> 7	
	<211> 30	
25	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
30	<223> Descripción de la Secuencia Artificial: Sitio de fijación de dedo de cinc de sitio N//Z	
	<220>	
35	<221> Proteína_fijación	
	<222> (6)..(14)	
	<223> Secuencia de reconocimiento de dedo de cinc de sitio N de 9 pb	
40	<220>	
	<221> Proteína_fijación	
	<222> (17)..(25)	
	<223> Secuencia de reconocimiento de dedo de cinc de sitio Z de 9 pb	
45	<400> 7	
	tctgcaaggg ttcagtcgt ggcggttaag	30
50	<210> 8	
	<211> 9	
	<212> DNA	
55	<213> Secuencia Artificial	
	<220>	
60	<223> Descripción de la Secuencia Artificial: Sitio de reconocimiento del péptido Zif268 de tres dedos	
	<400> 8	
	gcgtgggcg	9
65	<210> 9	
	<211> 9	

ES 2 341 926 T3

	<212> DNA	
	<213> secuencia Artificial	
5	<220>	
	<223> Descripción de la Secuencia Artificial: Sitio de fijación del péptido NRE (elemento de respuesta nuclear a las hormonas) de tres dedos	
10	<400> 9	
	aagggttca	9
15	<210> 10	
	<211> 19	
	<212> DNA	
	<213> Secuencia Artificial	
20	<220>	
	<223> Descripción de la Secuencia Artificial: Sitio diana compuesto de sitio N/Z con un solo par de bases adicional en el centro	
25	<400> 10	
	aagggttcag gcgtgggcg	19
30	<210> 11	
	<211> 20	
	<212> DNA	
35	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: Sitio diana compuesto de sitio N//Z con dos pares de bases adicionales en el centro	
40	<400> 11	
	aagggttcag tgcgtgggcg	20
45	<210> 12	
	<211> 18	
50	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
55	<223> Descripción de la Secuencia Artificial: Sitio NZ con sitios de fijación de NRE y Zif268 directamente yuxtapuestos	
	<400> 12	
60	aagggttcag cgtgggcg	18
	<210> 13	
65	<211> 5	
	<212> PRT	
	<213> Secuencia Artificial	

ES 2 341 926 T3

<220>

<223> Descripción de la Secuencia Artificial: enlazador canónico "TGEKP"

5 <400> 13

Thr Gly Glu Lys Pro

1 5

10

<210> 14

<211> 8

15 <212> PRT

<213> Secuencia Artificial

<220>

20 <223> Descripción de la Secuencia Artificial: enlazador flexible más corto para la proteína de fusión 268/NRE que contiene una glicina

<400> 14

25

Arg Gln Lys Asp Gly Glu Arg Pro

1 5

30 <210> 15

<211> 11

<212> PRT

35 <213> Secuencia Artificial

<220>

40 <223> Descripción de la Secuencia Artificial: enlazador flexible más largo para la proteína de fusión 268//NRE que contiene una secuencia Gly-Gly-Gly-Ser

<400> 15

45 Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro
1 5 10

<210> 16

50 <211> 9

<212> PRT

<213> Secuencia Artificial

55 <220>

<223> Descripción de la Secuencia Artificial: región enlazadora en Zif268

<400> 16

60

His Thr Gly Glu Lys Pro Phe Ala Cys

1 5

65 <210> 17

<211> 13

ES 2 341 926 T3

<212> PRT
 <213> Secuencia Artificial

5 <220>
 <223> Descripción de la Secuencia Artificial: región enlazadora en 268/NRE

<400> 17

10 His Leu Arg Gln Lys Asp Gly Glu Arg Pro Tyr Ala Cys
 1 5 10

15 <210> 18
 <211> 16
 <212> PRT
 <213> Secuencia Artificial

20 <220>
 <223> Descripción de la Secuencia Artificial: región enlazadora en 268//NRE

25 <400> 18

 His Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro Tyr Ala Cys
 1 5 10 15

30 <210> 19
 <211> 6
 <212> PRT

35 <213> Secuencia Artificial

<220>
 40 <223> Descripción de la Secuencia Artificial: secuencia enlazadora sintética

<400> 19

45 Gly Thr Gly Gln Lys Pro
 1 5

50 <210> 20
 <211> 4
 <212> PRT
 <213> Secuencia Artificial

55 <220>
 <223> Descripción de la Secuencia Artificial: secuencia enlazadora sintética

60 <400> 20

 Gly Glu Lys Pro
 1

65 <210> 21
 <211> 25

ES 2 341 926 T3

<212> PRT

<213> Secuencia Artificial

5 <220>

<223> Descripción de la Secuencia Artificial:motivo ilustrativo para la proteína con dedos de cinc C2H2

<220>

10 <221> MOD_RES

<222> (2)..(5)

<223> Xaa = cualquier aminoácido

15 <220>

<221> MOD_RES

<222> (4)..(5)

20 <223> Xaa puede estar presente o ausente

<220>

<221> MOD_RES

25 <222> (7)..(18)

<223> Xaa = cualquier aminoácido

<220>

30 <221> MOD_RES

<222> (20)..(24)

<223> Xaa = cualquier aminoácido

35 <220>

<221> MOD_RES

<222> (23)..(24)

40 <223> Xaa puede estar presente o ausente

<400> 21

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

45 1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His

50 20 25

<210> 22

<211> 10

55 <212> DNA

<213> Secuencia artificial

<220>

60 <223> Descripción de la Secuencia artificial: sitio diana de dedos de cinc con dos subsitios "D-able" solapantes

<220>

<221> base_modificada

65 <222> (1)..(2)

<223> n = g, a, c o t

ES 2 341 926 T3

<220>
 <221> base_modificada
 <222> (5)
 5 <223> n = g, a, c o t

 <220>
 <221> base_modificada
 10 <222> (8)
 <223> n = g, a, c o t

 <220>
 15 <221> base_modificada
 <222> (9)
 <223> n = a, c o t; si g, entonces la posición 10 es g o t
 20
 <220>
 <221> base_modificada
 <222> (10)
 25 <223> n = a o c; si la posición 9 es g, entonces n = g o t

 <400> 22
 30 nngkngkrmn 10

 <210> 23
 <211> 10
 35 <212> DNA
 <213> Secuencia Artificial

 <220>
 40 <223> Descripción de la Secuencia Artificial'. sitio diana con dedos de cinc que tiene tres subsitios "D-able" solapantes

 <220>
 <221> base_modificada
 45 <222> (1)..(2)
 <223> n = g, a, c o t

 <220>
 50 <221> base_modificada
 <222> (5)
 <223> n = g, a, c o t
 55
 <220>
 <221> base_modificada
 <222> (8)
 60 <223> n = g, a, c o t

 <400> 23
 65 nngkngkngk 10

ES 2 341 926 T3

<210> 24

<211> 4

<212> PRT

5 <213> Secuencia Artificial

<220>

10 <223> Descripción de la Secuencia Artificial:secuencia incluida en el enlazador más largo para permitir flexibilidad adicional

<400> 24

15 Gly Gly Gly Ser
1

20

25

30

35

40

45

50

55

60

65