

NUCLEASA EFECTORA TAL PARA LA INACTIVACIÓN ESPECÍFICA DEL
CORRECEPTOR DEL VIH CCR5

La invención se refiere a una nueva nucleasa efectora TAL (TALEN) para la inactivación específica del correceptor del VIH CCR5.

Aparte de su propia función en la célula, el receptor de quimiocinas CCR5 desempeña un papel importante en la infección del VIH. Aquí se presenta en las cepas denominadas CCR5-trópicas del VIH como un correceptor que interviene en la infección inicial de VIH. Si no hay ningún CCR5 en la superficie de un linfocito T auxiliar, los VIH no pueden fusionarse con la célula huésped y no se produce la infección. Así, una delección homocigota en el gen CCR5 (CCR5 Δ 32), que se presenta aproximadamente en el 1 % de los europeos occidentales y de los americanos "blancos" ("caucásicos") protege casi por completo de una infección con cepas CCR-trópicas del VIH. En consecuencia, CCR5 es una diana muy interesante para el tratamiento del VIH.

Las estrategias farmacológicas previas dirigidas a bloquear CCR5 requieren un tratamiento de por vida en el marco de un tratamiento combinado antirretrovírico (TAR). A largo plazo, esto conlleva efectos secundarios potencialmente graves, pero también una falta de cumplimiento del paciente y la aparición de resistencias. Por el contrario, una destrucción genética ("inactivación") de CCR5 (por medio de genoterapia), en el caso ideal en un solo tratamiento, sería suficiente, ya que la protección genética se transferiría a todas las células hijas. Esto no solo queda demostrado por la resistencia natural de los individuos homocigotos para CCR5 Δ 32, sino también por el caso clínico del éxito del tratamiento de una infección del VIH en el, así denominado, "paciente de Berlín", después del alotransplante de células madre con células donantes homocigotas para CCR5 Δ 32 (Hütter G. y col., Long-term

control of HIV by CCRDelta32/Delta32 stem-cell transplantation, N. Engl. J. Med. 2009, 360: 692-698; Allers K. y col., Evidence for the cure of HIV infection by CCR5 Δ 32/ Δ 32 stem cell transplantation, Blood 2011, 117: 2791-2799).

Sobre la base de estas observaciones se han desarrollado conceptos para la inactivación genética de CCR5 en los pacientes con VIH. Las estrategias más prometedoras hasta el momento se basan en las denominadas “nucleasas de diseño” (véase, por ejemplo, Manjunath N. y col., Newer Gene Editing Technologies toward HIV Gene Therapy, Viruses 2013, 5: 2748-2766). Estas nucleasas de diseño constan de dos componentes: un domino de reconocimiento, que determina la especificidad en el genoma y que puede diseñarse de forma casi completamente libre, y un dominio nucleásico, que induce una rotura bicatenaria en el lugar elegido del genoma. La reparación incorrecta de esta rotura bicatenaria da lugar a desplazamientos del marco de lectura abierto del gen diana y, por tanto, en el caso ideal, a una inactivación. Las primeras nucleasas de diseño de uso extendido fueron las nucleasas con dedos de cinc (ZFN). Por ejemplo, la empresa Sangamo BioSciences Inc. está ensayando actualmente una nucleasa con dedos de cinc específica para CCR5 desarrollada por la propia empresa con la denominación SB-728 (<http://www.sangamo.com/pipeline/sb-728.html>) para su uso clínico (Tebas y col., Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV, N. Engl. J. Med., 2014, 370: 901-10). El estudio clínico demostró la factibilidad de la estrategia, pero solamente pudo observarse un beneficio clínico a largo plazo en relación con la carga vírica en una de las personas de experimentación, que resultó ser heterocigota para la mutación CCR5 Δ 32 natural.

Las nucleasas efectoras TAL (nucleasas efectoras de tipo activador de transcripción, TALEN) constituyen la siguiente

generación de nucleasas de diseño (véanse, por ejemplo, Mussolino C., Cathomen T., TALE nucleases: tailored genome engineering made easy, Curr. Opin. Biotechnol. 2012, 23(5): 644-50; y los documentos WO 2011/072246 A2, EP 2510096 A2, WO 2011/154393 A1, WO 2011/159369 A1, WO 2012/093833 A2 y WO 2013/182910 A2). En comparación con las ZFN, se caracterizan, por ejemplo, por una mayor especificidad, de manera que se reduce considerablemente el riesgo de efectos fuera de diana, es decir, la producción de mutaciones en otro lugar del genoma distinto del deseado (Händel E.-M., Cathomen T., Zinc-finger nuclease based genome surgery: it's all about specificity. Curr. Gene Ther. 2011, 11: 28-37; Mussolino C. y col., A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity, Nucleic Acids Res. 2011, 39: 9283-9293).

Ya se conocen TALEN específicas para CCR5 (véanse, por ejemplo, Mussolino C. y col., A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity, Nucleic Acids Res. 2011, 39: 9283-9293; y los documentos WO 2011/146121 A1, WO 2012/093833 A2 y US 2013/0217131 A1), sin embargo, hasta ahora no se ha descrito ningún beneficio clínico.

La necesidad de agentes para el tratamiento eficaz de las infecciones de VIH persiste. Por consiguiente, el objetivo de la presente invención es hacer disponible un agente de este tipo. En particular, el objetivo de la presente invención es hacer disponible un agente, por medio del cual pueda conseguirse una inactivación del correceptor CCR5 del VIH más eficiente que hasta ahora.

El objetivo se consigue mediante el objeto de la reivindicación 1 y de las reivindicaciones subordinadas siguientes. En las reivindicaciones subordinadas se indican configuraciones recomendadas de la solución de acuerdo con la invención.

En un primer aspecto, la invención hace disponible un par de nucleasas efectoras TAL que comprende un primer y un segundo monómero de nucleasa efectora TAL, en que cada monómero de nucleasa efectora TAL comprende un dominio endonucleásico con actividad de endonucleasa de tipo II y un dominio de unión a ADN del efector TAL con una pluralidad de unidades repetitivas que presentan en cada caso un par de aminoácidos variables RVD y en que

a) el dominio de unión a ADN del efector TAL del primer monómero de nucleasa efectora TAL se une a la secuencia diana GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) y/o comprende la secuencia de RVD NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN

y

b) el dominio de unión a ADN del efector TAL del segundo monómero de nucleasa efectora TAL se une a la secuencia diana AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) y/o comprende la secuencia de RVD NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG.

El par de nucleasas efectoras TAL de acuerdo con la invención es capaz de producir la inactivación de CCR5 en linfocitos T primarios con una alta eficiencia, inalcanzada hasta el momento, de más del 50 %. A este respecto, sorprendentemente, la invención hace posible también una inactivación consistentemente bialélica de los dos alelos de CCR5 y de este modo una protección completa de las células modificadas frente a la entrada del VIH, en contraposición con una opinión expresada por los principales expertos en el estado de la técnica, según la cual esto no es posible en la actualidad ("Consistent nuclease-mediated biallelic knockdown is not yet tenable", véase Kay M. A. y Walker B. D., 2014, Engineering Cellular Resistance to HIV, N. Engl. J. Med. 370: 968-969). Además, se ha comprobado que el par de TALEN de acuerdo con la invención se adecúa perfectamente a una transferencia

génica basada en la transfección de ARNm (y, por tanto, especialmente cuidadosa, segura y compatible con las prácticas correctas de fabricación). Por consiguiente, la invención hace disponible por primera vez un agente para el tratamiento del VIH basado en una nucleasa de diseño que combina una alta eficiencia y selectividad de inactivación con escasos efectos fuera de diana y otras propiedades ventajosas desde el punto de vista farmacológico.

En este documento, por “nucleasa efectora TAL” o “TALEN” (nucleasa efectora de tipo activador de transcripción) se entiende una proteína de fusión que contiene el dominio de unión a ADN de un efector TAL (TALE) y el dominio de corte de ADN de una endonucleasa de restricción. Los efectores TAL son proteínas que se unen a ADN producidas por el patógeno de plantas *Xanthomonas* spp. La unión a ADN de los efectores TAL tiene lugar a través de un dominio con un número variable (generalmente de cinco a 30) de unidades repetitivas (“repeticiones”), en su mayoría de 33 a 35 aminoácidos. En ello, cada unidad repetitiva cuenta con dos restos aminoacídicos hipervariables (par de aminoácidos variables de la repetición, RVD), por lo general en las posiciones 12 y 13, que se unen exactamente a una base de una secuencia de ADN diana. La relación entre RVD y el nucleótido diana en el ADN se indica a continuación.

RVD (código de una letra)	RVD (código de tres letras)	de Nucleótido(s)
NH	Asn-His	G
HD	His-Asp	C
NG	Asn-Gly	T
NI	Asn-Ile	A
NN	Asn-Asn	R (G, A)
NK	Asn-Lys	G
NS	Asn-Ser	N (A, C, G, T)

En este documento, por "secuencia de RVD" se entiende una sucesión coherente de RVD en el dominio de unión de un efector TAL, en lo que, siempre que no se diga lo contrario, la secuencia se indica en el sentido N-C, es decir del extremo N al extremo C. Por supuesto, el experto será consciente de que los RVD de una "secuencia de RVD" no se suceden entre sí inmediatamente, en el sentido de que estén directamente unidos covalentemente entre sí, sino que son las unidades repetitivas ("repeticiones") que contienen los RVD las que están unidas directamente entre sí, de manera que los RVD están separados respectivamente por los aminoácidos de la estructura básica de las unidades repetitivas.

Por "secuencia diana" se entiende en este documento una secuencia nucleotídica, por lo general una secuencia de ADN, a la que se une el dominio de unión del efector TAL.

Las secuencias de RVD aquí desveladas, compuestas respectivamente de 19 RVD, se corresponden con las siguientes secuencias diana (en sentido 5'-3'; código de una letra para los aminoácidos):

NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN
GCTGGTCATCCTCATCCTG (SEQ ID NO: 1)

NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG
AGATGTCAGTCATGCTCTT (SEQ ID NO: 2)

Por "unidad repetitiva" ("repetición"), en relación con el dominio de unión de un efector TAL, se entiende en este documento una sucesión coherente de generalmente 33-35, en su mayoría de 34 aminoácidos, que, con excepción de los hipervariables RVD en las posiciones 12 y 13, presentan una secuencia aminoacídica esencialmente idéntica. También es posible que dentro de la estructura básica conservada de la unidad repetitiva, es decir, en la estructura esencialmente

invariable en la que se integran los hipervariables RVD, haya algunos aminoácidos variables, por ejemplo en las posiciones 4, 10 y/o 32 de una unidad repetitiva de 34 aminoácidos. Por ejemplo, una unidad repetitiva típica puede tener la secuencia aminoacídica siguiente (los subíndices indican la posición dentro de la unidad repetitiva):

LTPX₄QVVAIX₁₀SX₁₂X₁₃GGKQALETVQRLLPVLCQX₃₂HG (SEQ ID NO: 5)

X representa un aminoácido cualquiera, que en las posiciones 12 y 13 corresponde a los aminoácidos hipervariables RVD. En la posición 4 (X₄) pueden encontrarse, por ejemplo, los aminoácidos E, Q, D o A y en la posición 32 (X₃₂) los aminoácidos A o D. En la posición 10 pueden encontrarse, por ejemplo, A o V. A continuación se reproducen ejemplos de unidades repetitivas (XX representa los aminoácidos hipervariables RVD; los aminoácidos variables están subrayados para su identificación):

LTPE0VVAIASXXGGKOALETVORLLPVLCOAHHG (SEQ ID NO: 6)

LTPQ0VVAIASXXGGKQALETVQRLLPVLCQAHHG (SEQ ID NO: 7)

LTPD0VVAIASXXGGKQALETVQRLLPVLCQDHHG (SEQ ID NO: 8)

LTPA0VVAIASXXGGKOALETVORLLPVLCODHHG (SEQ ID NO: 9)

LTPE0VVAIVSXGGKOALETVORLLPVLCOAHHG (SEQ ID NO: 10)

Los dominios de unión de los efectores TAL pueden incluir una o varias variantes de unidades repetitivas semejantes, en lo que también es posible una mezcla de distintas variantes.

La unidad repetitiva exterior, más próxima al dominio nucleásico, puede tener menos aminoácidos, por ejemplo, solo comprender los primeros 15, 16, 17, 18, 19 o 20 aminoácidos de las demás unidades repetitivas. Una unidad

repetitiva semejante se denomina también “media unidad repetitiva” o “media repetición”.

Por un “dominio de unión a ADN” se entiende en este documento una región de una proteína que ocasiona la unión de la proteína a un ADN. En el caso de un dominio de unión a ADN de un efector TAL, esta unión tiene lugar por medio de las unidades repetitivas (“repeticiones”) descritas con más detalle anteriormente.

La formulación según la cual un dominio de unión a ADN de un efector TAL se une a una secuencia diana ha de entenderse como que el dominio de unión a ADN del efector TAL se une específicamente a la secuencia diana gracias a su secuencia de RVD. A este respecto, aunque se prefiere, no es necesario que cada nucleótido de la secuencia diana se corresponda con un RVD en el dominio de unión. La relación entre la secuencia de RVD del dominio de unión a ADN del efector TAL y la secuencia diana solamente debe ser tal que tenga lugar una unión específica a la secuencia diana. En este contexto, “específica” significa que esencialmente la unión tiene lugar exclusivamente con la secuencia diana.

Por “monómero de nucleasa efectora TAL” se entiende una nucleasa efectora TAL formada por una sola cadena polipeptídica. Por “par de nucleasas efectoras TAL” o “par de TALEN” se entiende una TALEN compuesta de dos monómeros de nucleasa efectora TAL. Los monómeros representan los brazos izquierdo o derecho de una TALEN, que se unen a las cadenas opuestas de un ADN y conjuntamente producen la escisión del ADN en un punto.

Cuando en este documento en relación con un par de TALEN se habla de una TALEN “izquierda” o “derecha” o del “brazo izquierdo o derecho” de una TALEN, esto refleja el hecho de que en un par de TALEN se emplean monómeros de TALEN en

pares, es decir, producen una rotura de cadena en un ADN bicatenario al unirse un monómero a una secuencia diana en la cadena codificante, mientras el otro monómero de TALEN del par de TALEN se une a una secuencia diana en la cadena no codificante complementaria, y ello de tal modo que los dominios nucleásicos, orientados el uno al otro, se sitúan conjuntamente en una región de ADN entre las secuencias diana denominada "espaciador" y producen respectivamente en este punto una rotura monocatenaria. Por consiguiente, los monómeros de TALEN "izquierdo" y "derecho" son parte de uno de estos pares de TALEN, en lo que frecuentemente, como TALEN "izquierda" se denomina la TALEN que se une a la cadena codificante, mientras que la TALEN "derecha" se une a la cadena complementaria. Sin embargo, cuando en este documento se habla de una TALEN "izquierda" o "derecha", con ello no se debe dar por supuesto que la TALEN "izquierda" se une a la cadena codificante y la TALEN "derecha" se une a la cadena complementaria.

La presente invención se refiere también, por tanto, a un "par de TALEN", es decir, un par compuesto de dos monómeros de acuerdo con la invención, cada uno de los cuales representa un brazo izquierdo o derecho de una TALEN. Por consiguiente, la presente invención se refiere también a un par de nucleasas efectoras TAL que comprende un monómero de nucleasa efectora TAL cuyo dominio de unión a ADN del efector TAL se une a la secuencia diana GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) y/o comprende la secuencia de RVD NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN y un monómero de nucleasa efectora TAL cuyo dominio de unión a ADN del efector TAL se une a la secuencia diana AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) y/o comprende la secuencia de RVD NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG.

Por "dominio endonucleásico con actividad de endonucleasa de tipo II" se entiende un polipéptido que presenta la

actividad de escisión de ADN de una endonucleasa de restricción y corta el ADN dentro o en inmediata proximidad de una secuencia de reconocimiento, no requiere ATP y no tiene actividad de metiltransferasa. Por "dominio endonucleásico con actividad de endonucleasa de tipo IIS" se entiende un dominio de una endonucleasa de tipo II, cuyo punto de corte se encuentra en inmediata proximidad de la secuencia de reconocimiento, pero no dentro de la misma.

Por "CCR5" se entiende en este documento el receptor de quimiocinas con el motivo CC 5 (también denominado CD195, CMKBR5 o CC-CKR5). La secuencia del CCR5 humano se reproduce en SEQ ID NO: 11 (véase el n° de acceso NC_018914.2 del NCBI).

Por "vector" se entiende en este documento un vehículo de transporte para la transferencia de un ácido nucleico principalmente exógeno a una célula receptora viva mediante transfección o transducción. Por "vector de transferencia génica" se entiende en este documento un vector por medio del cual puede introducirse un gen en una célula. Los vectores (de transferencia génica) son bien conocidos por el experto. Algunos ejemplos de vectores de transferencia génica son plásmidos, vectores víricos o ARNm.

Por "ácido nucleico" se entiende un polímero cuyos monómeros son nucleótidos. Un nucleótido es un compuesto formado por un resto de azúcar, una base orgánica heterocíclica que contiene nitrógeno (base nucleotídica o nucleobase) y un grupo fosfato. Por lo general, el resto de azúcar es una pentosa, en el caso del ADN desoxirribosa y en el caso del ARN ribosa. La conexión de los nucleótidos tiene lugar a través del grupo fosfato, mediante un enlace fosfodiéster, por lo general, entre el átomo de C 3' del componente de azúcar de un nucleósido (compuesto formado por nucleobase y azúcar) y el átomo de C 5' del componente de azúcar del siguiente nucleósido. El término "ácido

nucleico" comprende, por ejemplo, ADN, ARN y secuencias mixtas de ADN/ARN. Según se usa en este documento, el término "ácido nucleico" se refiere especialmente a un ácido nucleico aislado. Por "ácido nucleico aislado" se entiende un ácido nucleico extraído de su entorno natural u originario o preparado sintéticamente.

El término "que comprende" se usa en este documento de manera que define tanto un objeto que presenta exclusivamente las características abarcadas por el término, como un objeto que, además de las características abarcadas por el término, presenta otras características adicionales. Por consiguiente, la definición de un objeto porque abarca determinadas características incluye también la definición de dicho objeto por el listado concluyente de dichas características, es decir, por la presencia exclusivamente de dichas características.

En una forma de realización preferida del par de nucleasas efectoras TAL de acuerdo con la invención, en cada uno de los monómeros de nucleasa efectora TAL, el dominio endonucleásico se encuentra en posición C terminal con respecto al dominio de unión a ADN del efector TAL. Preferentemente, cada unidad repetitiva, con excepción de la unidad repetitiva más próxima al dominio endonucleásico, comprende de 33 a 35 aminoácidos, con preferencia 34 aminoácidos, en que los RVD se encuentran en cada unidad repetitiva en las posiciones 12 y 13. Con preferencia especial, todas las unidades repetitivas, con excepción de la "media repetición", tienen la secuencia aminoacídica de acuerdo con SEQ ID NO: 5, en que la posición 4 puede corresponder a E, Q, D o A, la posición 10 a A o V y la posición 32 a A o D. La estructura básica de las unidades repetitivas puede ser igual o diferente para todas las unidades repetitivas. Puede haber una o varias unidades repetitivas en las que las posiciones dentro de la estructura básica varíen con respecto a los aminoácidos,

por ejemplo las posiciones 4, 10 y/o 32. La unidad repetitiva más próxima al dominio endonucleásico puede comprender un menor número de aminoácidos, por ejemplo 15, 16, 17, 18, 19 o 20 aminoácidos, en lo que, en este caso, los aminoácidos corresponden preferentemente a los primeros 15, 16, 17, 18, 19 o 20 aminoácidos de las demás unidades repetitivas. En ello, por ejemplo, el aminoácido en la posición 4 puede variar, por ejemplo, ser E, Q, D y A y/o el aminoácido en la posición 10 puede variar, por ejemplo, ser V en lugar de A.

Con preferencia especial, el dominio endonucleásico de los monómeros de nucleasa efectora TAL es un dominio endonucleásico de tipo IIS, con preferencia especial el dominio de escisión de ADN de la endonucleasa *FokI*. Una secuencia aminoacídica adecuada para un dominio de escisión para *FokI* se indica en SEQ ID NO: 12. Sin embargo también pueden considerarse otros dominios de escisión para endonucleasas de tipo II. Las endonucleasas de tipo II son conocidas por el experto y los dominios de escisión adecuados pueden determinarse mediante investigaciones rutinarias.

En una forma de realización especialmente preferida, el primer monómero de nucleasa efectora TAL presenta una secuencia aminoacídica de acuerdo con SEQ ID NO: 3 y el segundo monómero de nucleasa efectora TAL una secuencia aminoacídica de acuerdo con SEQ ID NO: 4. En ello, en SEQ ID NO: 3 se indica la TALEN izquierda (denominada también a continuación CCR5-Uco-L o brazo izquierdo de CCR5-Uco) y en SEQ ID NO: 4 la TALEN derecha (a continuación denominada también CCR5-Uco-R o brazo derecho de CCR5-Uco) de un par de TALEN, las cuales causan conjuntamente una rotura bicatenaria en la secuencia de ADN de CCR5 dentro del espaciador situado entre las secuencias diana de acuerdo con SEQ ID NO: 1 y SEQ ID NO: 2. La reparación de esta rotura bicatenaria por los sistemas de reparación celulares

(unión de extremos no homólogos, NHEJ) conduce muy probablemente a un desplazamiento del marco de lectura y por tanto a una inactivación de CCR5.

En un segundo aspecto, la presente invención se refiere también a un ácido nucleico que comprende

- a) un primer ácido nucleico que codifica un primer monómero de nucleasa efectora TAL, en que el primer monómero de nucleasa efectora TAL comprende un dominio endonucleásico con actividad de endonucleasa de tipo II y un dominio de unión a ADN del efector TAL con una pluralidad de unidades repetitivas que presentan en cada caso un par de aminoácidos variables RVD y en que el dominio de unión a ADN del efector TAL se une a la secuencia diana GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) y/o comprende la secuencia de RVD NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN y
- b) un segundo ácido nucleico que codifica un segundo monómero de nucleasa efectora TAL, en que el segundo monómero de nucleasa efectora TAL comprende un dominio endonucleásico con actividad de endonucleasa de tipo II y un dominio de unión a ADN del efector TAL con una pluralidad de unidades repetitivas que presentan en cada caso un par de aminoácidos variables RVD y en que el dominio de unión a ADN del efector TAL se une a la secuencia diana AGATGTGTCAGTCATGCTCTT (SEQ ID NO: 2) y/o comprende la secuencia de RVD NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG.

En este aspecto de la invención, los monómeros de TALEN que forman el par de TALEN de acuerdo con la invención están codificados conjuntamente en un solo ácido nucleico. El ácido nucleico puede ser, por ejemplo, un plásmido u otro vector (de transferencia génica) adecuado. Los vectores adecuados, así como los procedimientos para su preparación y su uso son bien conocidos en el estado de la técnica. Dado el caso, el ácido nucleico puede contener otros

elementos, además del código de la TALEN, por ejemplo uno o varios promotores, señales de poliadenilación, etc.

Los monómeros de TALEN que forman el par de TALEN de acuerdo con la invención también pueden estar codificados por separado en dos ácidos nucleicos. En un tercer aspecto, la presente invención se refiere por tanto también a una composición de ácidos nucleicos que comprende

- a) un primer ácido nucleico que codifica un primer monómero de nucleasa efectora TAL, en que el primer monómero de nucleasa efectora TAL comprende un dominio endonucleásico con actividad de endonucleasa de tipo II y un dominio de unión a ADN del efector TAL con una pluralidad de unidades repetitivas que presentan en cada caso un par de aminoácidos variables RVD y en que el dominio de unión a ADN del efector TAL se une a la secuencia diana GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) y/o comprende la secuencia de RVD NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN y
- b) un segundo ácido nucleico que codifica un segundo monómero de nucleasa efectora TAL, en que el segundo monómero de nucleasa efectora TAL comprende un dominio endonucleásico con actividad de endonucleasa de tipo II y un dominio de unión a ADN del efector TAL con una pluralidad de unidades repetitivas que presentan en cada caso un par de aminoácidos variables RVD y en que el dominio de unión a ADN del efector TAL se une a la secuencia diana AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) y/o comprende la secuencia de RVD NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG.

Preferentemente, el primer y el segundo ácido nucleico son respectivamente ARNm, con preferencia especial ARNm estabilizado (véanse, por ejemplo, Kallen K.-J. y col., A novel, disruptive vaccination technology, Hum. Vaccin. Immunother. 1 de octubre de 2013, 9(10): 2263-2276, iod: 10.4161/hv.25181; Kallen K.-J. y Theß A., A development

that may evolve into a revolution in medicine: mRNA as the basis for novel, nucleotide-based vaccines and drugs, Ther. Adv. Vaccines, enero de 2014; 2(1): 10-31, iod:10.1177/2051013613508729). Dado el caso, el primer y el segundo ácido nucleico pueden contener otros elementos, además del código de la TALEN, por ejemplo uno o varios promotores, señales de poliadenilación, etc. En SEQ ID NO: 17 y 18 se indican a modo de ejemplo ARNm adecuados para los brazos izquierdo y derecho de una TALEN de acuerdo con la invención.

En el caso de un ARNm, su transporte a la(s) célula(s) diana, por ejemplo, linfocitos T, tiene lugar por medio del procedimiento descrito por Berdien y col. (Berdien B. y col., TALEN-mediated editing of endogenous T-cell receptors facilitates efficient reprogramming of T lymphocytes by lentiviral gene transfer, Gene Therapy, 2014, iod:10.1038/gt.2014.26). Con preferencia especial, los dos brazos de la TALEN (brazos izquierdo y derecho) se introducen en una célula simultáneamente.

La introducción del par de TALEN de acuerdo con la invención mediante ARNm presenta una serie de ventajas decisivas para el uso clínico. El empleo de un vector de transferencia génica a base de ADN puede evitarse, lo que simplifica definitivamente la preparación y la aplicación práctica. La expresión de la TALEN a partir de ARNm solo tiene lugar de manera comparativamente breve, ya que el ARNm se degrada rápidamente en la célula diana. De este modo se reduce aún más el riesgo de efectos fuera de diana. Además, las células diana sólo tienen que cultivarse *in vitro* durante muy poco tiempo. La tecnología correspondiente es fácil de adaptar a los requisitos de las prácticas correctas de fabricación. Al contrario que con los vectores víricos o plasmídicos, no son de esperar efectos secundarios por la transferencia génica misma (por ejemplo, mutaciones de inserción) o por una expresión de la

TALEN a largo plazo (efectos fuera de diana, activación de una respuesta inmunitaria específica para TALEN) a consecuencia de una inserción del vector indeseada.

En otro aspecto, la presente invención se refiere también a un vector, especialmente un vector de transferencia génica, que comprende un ácido nucleico de acuerdo con la invención y una célula hospedadora aislada que comprende un vector de acuerdo con la invención, un ácido nucleico o una composición de ácidos nucleicos de acuerdo con la invención, en que la célula hospedadora aislada no es una célula de línea germinal de un ser humano, en particular, no es una célula germinal humana ni una célula germinal embrionaria humana y tampoco es una célula madre embrionaria humana en cuya obtención se haya destruido o se destruya un embrión humano.

En otro aspecto más, la presente invención se refiere a una composición farmacéutica que comprende un ácido nucleico, una composición de ácidos nucleicos o un vector de acuerdo con la presente invención. La composición farmacéutica puede comprender adyuvantes, por ejemplo, disolventes, solubilizantes, aceleradores de disolución, salificantes, sales, sustancias tamponantes, modificadores de la viscosidad y la consistencia, gelificantes, emulsionantes, solubilizadores, humectantes, expansores, antioxidantes, conservantes, cargas y soportes, etc.

En otro aspecto más, la presente invención se refiere a un medicamento que comprende un ácido nucleico, una composición de ácidos nucleicos, un vector o una composición farmacéutica de acuerdo con la presente invención.

A continuación, la invención se explicará en más detalle con fines ilustrativos mediante ejemplos de realización y las figuras adjuntas.

Figura 1. Representación esquemática de los dominios de unión a ADN de un par de TALEN específico para CCR5 ("CCR5-Uco") y sus secuencias diana en el gen CCR5. En cada caso, la línea inferior muestra la secuencia diana para a) el brazo de TALEN izquierdo y b) el brazo de TALEN derecho, y la línea superior, los RVD relevantes (pares de aminoácidos variables de las repeticiones) de los correspondientes monómeros de TALEN en los recuadros (los aminoácidos se indican en el código de una letra); c) sección del ADN de CCR5 (nucleótidos 135-221 de la secuencia de acuerdo con SEQ ID NO: 11) con la cadena complementaria. Las secuencias diana del brazo izquierdo (arriba) y derecho (abajo, en la cadena complementaria) de la TALEN CCR5-Uco se resaltan encuadradas.

Figura 2. Comparación de eficiencia entre la TALEN para CCR5 de acuerdo con la invención ("Uco") y una TALEN para CCR5 de control ("Mco") del estado de la técnica. El ensayo se llevó a cabo mediante transfección plasmídica en una línea celular indicadora positiva para CCR5, basada en las células 293T. Para todas las construcciones ensayadas se observaron eficiencias de transfección similares (mediante la cotransfección de eGFP). La inactivación de CCR5 se determinó seis días después de la transfección de los clones celulares 293T CCR5+ por medio de un anticuerpo específico (anticuerpo anti-CD195-APC-Cy-7) (n = 2). Para los controles simulados, las células se transfectaron con un plásmido de control irrelevante (pUC) que no codifica ninguna TALEN.

Figura 3. Inactivación de CCR5 en linfocitos T primarios con CCR5-Uco tras la transfección de ARNm. Después de su activación *ex vivo*, aproximadamente la mitad de los linfocitos T primarios de una persona de experimentación sana expresaron CCR5 (= las células a la derecha de la línea rayada en comparación con las células "sin

transfectar"). a) Después de la transfección de la TALEN específica para CCR5 (Uco), la proporción de células positivas para CCR5 disminuye al aumentar la cantidad de ARNm transfectado (proporción inversa). b) Por el contrario, la transfección de control de los brazos de TALEN individuales no conlleva una disminución de la proporción de células positivas para CCR5. La inactivación de CCR5 se determinó seis días después de la transfección del ARNm por medio de un anticuerpo específico (anticuerpo anti-CD195-PerCP-Cy5.5). c) Un análisis de la diana de la TALEN CCR5-Uco muestra una inactivación génica en nueve de 17 (>50 %) de los linfocitos T primarios analizados.

EJEMPLOS

Se preparó y analizó una TALEN específica para CCR5 de acuerdo con la invención (a continuación "CCR5-Uco"). La TALEN de acuerdo con la invención se diferencia de las TALEN descritas hasta ahora en la secuencia diana reconocida en el gen CCR5 (véase la figura 1).

Frente a una TALEN específica para CCR5 con codones optimizados ("Mco"; véanse SEQ ID NO: 13, 14), basada en los trabajos publicados por el laboratorio del profesor Toni Cathomen (Friburgo, Alemania) (véase Mussolino C. y col., A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity, Nucleic Acids Res. 2011, 39: 9283-9293), la TALEN CCR5-Uco de acuerdo con la invención muestra una tasa de inducción de la inactivación de CCR5 después de su transfección plasmídica en una línea celular indicadora (véase la figura 2) claramente superior. En ello, la secuencia de ácido nucleico de los componentes de la TALEN con codones optimizados para su empleo en células humanas se basa en una publicación del grupo de Feng Zhang (Zhang y col., Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription, Nature

Biotechnology, 2011, 29: 149-153; Sanjana N. E. y col., A TAL Effector Toolbox for Genome Engineering, Nature Protocols, 2012, 7: 171-192).

Las secuencias de RVD de la TALEN CCR5-Mco de acuerdo con el estado de la técnica son las siguientes:

brazo izquierdo (L) = NN NG NN NN NN HD NI NI HD NI NG NN HD NG NN NN NG HD;

brazo derecho (R) = HD NG NG HD NI NN HD HD NG NG NG NG NN HD NI NN NG NG.

La correspondiente secuencia de reconocimiento a nivel de ADN:

L en la cadena codificante = GTGGGCAACATGCTGGTC (SEQ ID NO: 15);

R en la cadena no codificante = CTTCAGCCTTTTGCAGTT (SEQ ID NO: 16).

La longitud del espaciador es de 15 nucleótidos. También aquí se basa la preparación de los plásmidos con las TALEN en la publicación de Zhang F. o bien Sanjana N. E. y col. (véase anteriormente).

Mediante la transfección de ARNm (véase Berdien B. y col., TALEN-mediated editing of endogenous T-cell receptors facilitates efficient reprogramming of T lymphocytes by lentiviral gene transfer, Gene Therapy, 2014, doi:10.1038/gt.2014.26) con la CCR5-Uco de acuerdo con la invención se logró la inactivación de CCR5 en linfocitos T primarios (véase la figura 3). Las secuencias de ácido nucleico de los ARNm usados para ello se reproducen en SEQ ID NO: 17 y 18. SEQ ID NO: 17 corresponde al ARNm del brazo de TALEN izquierdo y SEQ ID NO: 18 al brazo de TALEN

derecho. Los nucleótidos 10-3.225 de los ARNm en SEQ ID NO: 17 y 18 codifican respectivamente los brazos de TALEN (monómeros) cuyas secuencias aminoacídicas se indican en SEQ ID NO: 3 y 4.

El ARNm transfectado se preparó mediante el promotor T7, después de la linealización del vector con *AvrII*. Dado que el punto de corte de *AvrII* se encuentra 563 pb después del codón de parada, la secuencia indicada es mayor que el marco de lectura abierto. Después de la linealización, el ADN de la TALEN Uco respectivo se usó como molde para la preparación de ARNm mediante el sistema de producción de ARNm estándar T7 mScript™ de Cellscript (Madison, WI 53713, EE. UU.). De acuerdo con las indicaciones del fabricante, se incorporó al ARNm una caperuza 5' y una cola de poli A. La transfección del ARNm se llevó a cabo mediante electroporación de las células T primarias durante 10 ms a 300 V. Por el contrario, con la TALEN CCR5-Mco no se logró la inactivación de CCR5 en células T primarias ni en líneas celulares T (mientras sí fue posible la inactivación del receptor de las células T, véase Berdien y col., cita anterior). Solo con la TALEN CCR5-Uco de acuerdo con la invención fue posible inactivar una parte significativa superior al 50 % de los alelos de CCR5 mediante la transfección de ARNm (figura 3c). Esto sugiere que solamente TALEN con la suficiente actividad son capaces de desarrollar su función en células T primarias después de una transfección de ARNm. Por consiguiente, la TALEN para CCR5 de acuerdo con la invención es especialmente adecuada para su empleo mediante transfección de ARNm, lo que la hace extraordinariamente atractiva para el uso clínico.

Resumen de secuencias:

SEQ ID NO	Tipo	Descripción
01	ADN	Secuencia diana de TALEN CCR5-Uco L
02	ADN	Secuencia diana de TALEN CCR5-Uco R

03	Proteína	TALEN CCR5-Uco L
04	Proteína	TALEN CCR5-Uco R
05	Proteína	Secuencia repetida (consenso)
06	Proteína	Secuencia repetida
07	Proteína	Secuencia repetida
08	Proteína	Secuencia repetida
09	Proteína	Secuencia repetida
10	Proteína	Secuencia repetida
11	ADN	hCCR5
12	Proteína	Dominio de escisión de <i>FokI</i>
13	Proteína	TALEN CCR5-Mco L
14	Proteína	TALEN CCR5-Mco R
15	ADN	Secuencia diana de TALEN CCR5-Mco L
16	ADN	Secuencia diana de TALEN CCR5-Mco R
17	ARNm	ARNm de CCR5-Uco L
18	ARNm	ARNm de CCR5-Uco R

PROTOCOLO DE SECUENCIAS - TEXTO LIBRE

TALEN repeat = repetición de TALEN

Any amino acid, or E, Q, D o A = cualquier aminoácido o E, Q, D o A

Any amino acid, or A or V = cualquier aminoácido o A o V

Any amino acid, or A or D = cualquier aminoácido o A o D

Repeat variable diresidue (RVD) = par de aminoácidos variable de la repetición (RVD)

FokI cleavage domain = dominio de escisión de *FokI*