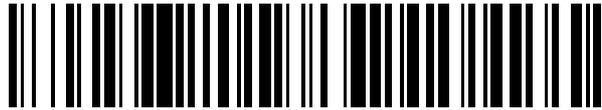


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 557 633**

51 Int. Cl.:

**C12N 9/22**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.06.2012 E 12735348 (0)**

97 Fecha y número de publicación de la concesión europea: **09.12.2015 EP 2718430**

54 Título: **Ribonucleasa H modificada por ingeniería específica de secuencia y el método para determinar la preferencia de secuencia de proteínas de unión de híbrido de ADN-ARN**

30 Prioridad:

**08.06.2011 PL 39517911**

**08.06.2011 PL 39518011**

**08.06.2011 US 201161494494 P**

**08.06.2011 US 201161494481 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**27.01.2016**

73 Titular/es:

**MIEDZYNARODOWY INSTYTUT BIOLOGII  
MOLEKULARNEJ I KOMORKOWEJ (100.0%)  
Ul. Ks. Trojdena 4  
02-109 Warszawa, PL**

72 Inventor/es:

**BUJNICKI, JANUSZ MAREK;  
SULEJ, AGATA AGNIESZKA;  
SKOWRONEK, KRZYSZTOF JERZY y  
NOWOTNY, MARCIN**

74 Agente/Representante:

**ZEA CHECA, Bernabé**

**ES 2 557 633 T3**

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

## DESCRIPCIÓN

Ribonucleasa H modificada por ingeniería específica de secuencia y el método para determinar la preferencia de secuencia de proteínas de unión de híbrido de ADN-ARN

5

**Campo técnico**

**[0001]** La invención se refiere a una ribonucleasa H específica de secuencia, que comprende una fusión de un dominio de ribonucleasa HI (RNasa HI) modificada por ingeniería con un dominio de dedos de cinc. La presente invención se puede aplicar en genética. La ribonucleasa específica de secuencia que actúa en híbridos de ADN-ARN es útil en cualquier aplicación, en la que se puede realizar la escisión de la hebra de ARN en un híbrido de ADN-ARN, por ejemplo en manipulación de ácidos nucleicos *in vitro*, en particular híbridos de ADN-ARN con una secuencia especificada. Además del uso *in vitro*, algunas enzimas que escinden híbridos de ADN-ARN de una manera específica de secuencia también se pueden usar en la terapia de ciertas infecciones de virus de ARN (por ejemplo, virus oncogénicos, retrovirus, hepatitis B, gripe), que se replican mediante la formación transitoria de una hebra de ADN en un molde de ARN o para escindir otros híbridos de ADN-ARN que se forman *in vivo*.

**[0002]** También se desvela un método para la determinación de la preferencia de secuencia del híbrido de ADN-ARN que se une a proteína(s) o sus dominio(s) y mediante esto la determinación de la secuencia reconocida mediante una proteína de unión a híbrido de ADN-ARN. El método mencionado anteriormente se puede aplicar en genética. El método para la determinación de la preferencia de secuencia proteínas de unión a híbrido de ADN-ARN se puede aplicar en la determinación de la preferencia de secuencia de cualquier proteína o su dominio que se une a secuencias específicas en el híbrido de ADN-ARN y mediante esto la determinación de la secuencia reconocida por una proteína(s) de unión a híbrido de ADN-ARN o su dominio. El método se puede aplicar de forma complementaria con cualquier técnica de modificación por ingeniería específica de tales proteínas. Las proteínas de unión a híbrido de ADN-ARN específicas se pueden usar en el diagnóstico de ciertos virus de ARN (por ejemplo, virus oncogénicos, retrovirus, hepatitis B, gripe), que se replican mediante la transición para la creación de un molde de ARN con hebra de ADN. Tales dominios también se pueden aplicar en modificación por ingeniería de proteínas para obtener enzimas con nueva especificidad tales como fusiones de proteína de unión ADN-ARN con dominios enzimáticos, tales como nucleasa, enzimas de modificación de ARN o de modificación de ADN y otros.

**Antecedentes de la técnica**

**[0003]** La escisión de ácidos nucleicos en una ubicación específica se usa frecuentemente en muchas técnicas de ingeniería genética. Existen muchos métodos para fragmentar moléculas de ADN, que incluyen las enzimas de restricción disponibles en el mercado usadas ampliamente. No se conocen muchas enzimas de procesamiento de ARN y la mayoría de ellas se caracterizan por una baja especificidad secuencial o por su ausencia total.

**[0004]** El documento de patente WO 2010076939 A1 describe las composiciones y métodos para realizar recombinación o mutación genética dirigida usando la nucleasa de dedos de cinc quimérica. El documento de patente WO 03087341 A2 describe el uso de una nucleasa de dedos de cinc para la edición dirigida del gen regulador de la conductancia transmembrana de la fibrosis quística humana, proporcionando de este modo una terapia potencial para la fibrosis quística. El documento de patente WO 2009146179 A1 describe el desarrollo de un método de ensamblaje modular altamente eficaz y fácil de poner en práctica que usa dedos de cinc disponibles al público para preparar nucleasas de dedos de cinc que son capaces de modificar las secuencias de ADN de varios sitios genómicos en las células humanas. Ninguno de los documentos de patente WO 2010076939 A1, WO 03087341 A2 o WO 2009146179 A1 describe o sugiere el uso de la fusión de la RNasa HI o parte de la misma con un dedo de cinc, en particular con ZfQQR ni en particular desvelan ni sugieren la posibilidad de la escisión dirigida de híbridos de ADN-ARN ni cualquier tipo de escisión dirigida de híbridos de ADN-ARN mediante cualquier proteína. A partir de las descripciones de los documentos de patente WO2007014181A2 y WO2007014182A2 se conocen fusiones de muchos dominios de dedos de cinc y nucleasas FokI que facilitan la edición dirigida del genoma. Sin embargo, a diferencia del primer caso, la fusión de las proteínas nativas solas, la RNasa HI y el dedo de cinc, no conduce a una enzima específica de secuencia.

**[0005]** El método para la determinación de la preferencia de secuencia de proteína(s) de unión a híbrido de ADN-ARN o su dominio(s) y mediante esto la determinación de la secuencia reconocida por una proteína de unión híbrido de ADN-ARN es una modificación del procedimiento de SELEX. SELEX representa la evolución sistemática de ligandos mediante enriquecimiento exponencial. El principio de este método se basa en la selección iterativa y enriquecimiento de moléculas a partir de una gran biblioteca bifurcada de secuencias de ácidos nucleicos que presentan una afinidad elevada hacia un ligando. La etapa del enriquecimiento se consigue mediante la unión de los ácidos nucleicos a un ligando y la retirada de las secuencias no unidas. Hasta ahora este método se ha usado para obtener aptámeros de ARN y de ADN que se unen a ligandos con especificidad elevadas (Ellington, A.D., Szostak, J.W., 1990. Nature 346, 818-822; Huizenga DE, Szostak JW. 1995. Biochemistry 34 (2): 656-65"), para determinar una preferencia de secuencia de una proteína de unión a ADN o a ARN (Blackwell TK y Weintraub H. 1990. Science 250: 1104-1110), pero nunca se usó para proteínas que se unen a híbridos de ADN-ARN. Las bibliotecas de

oligonucleótidos descritas usadas en modificaciones conocidas de SELEX consistían en oligonucleótidos monocatenarios (ARN, ADNss, ARN modificado, o ADNss modificado, PNA), o ADN bicatenario (ADNds), pero ninguna caracteriza el uso de biblioteca de híbrido de ADN-ARN.

5 **[0006]** El documento de patente WO 2010076939 A1 describe las composiciones y métodos para realizar la recombinación o mutación genética dirigida usando la nucleasa de dedos de cinc química. El documento de patente WO 03087341 A2 describe el uso de una nucleasa de dedos de cinc para la edición dirigida del gen regulador de la conductancia transmembrana de la fibrosis quística humana, proporcionando de este modo una  
10 método de ensamblaje modular altamente eficaz y fácil de poner en práctica que usa dedos de cinc disponibles al público para preparar nucleasas de dedos de cinc que son capaces de modificar las secuencias de ADN de varios sitios genómicos en las células humanas. Ninguna de las publicaciones mencionadas anteriormente describe el uso del método SELEX para determinar la preferencia de sustrato de proteínas de unión a híbrido de ADN-ARN o el uso de dedos de cinc para la unión de secuencias específicas en híbridos de ADN-ARN o el uso de ZfQQR en particular  
15 para esa finalidad.

**[0007]** Herskovitz M.A. *et al.*, Mol Microbiol. 2000 Dec: 38 (5): 1027-33, "Endoribonuclease RNase III is essential in *Bacillus subtilis*." describe el crecimiento de una cepa en la que la expresión de Bs-RNasa III (*rncS*) era dependiente de la transcripción de *rncS* a partir de un plásmido sensible a la temperatura y a la temperatura no permisiva daba  
20 como resultado un 90-95 % de muerte celular, y prácticamente todas las células que sobrevivían retenían el plásmido de expresión de *rncS*. Por lo tanto, los autores concluyeron en que *rncS* es esencial en *B. subtilis*. Dasgupta S. *et al.*, Mol Microbiol. 1998; 28 (3): 629-40, "Genetic uncoupling of the dsRNA-binding and RNA cleavage activities of the *Escherichia coli* endoribonuclease RNase III -- the effect of dsRNA binding on gene expression." describe los fenotípicos de bacterias que portan mutaciones puntuales en *rnc*, el gen que codifica la RNasa III. Karen Shahbadian *et al.*, The EMBO Journal (2009) 28, 3523-3533, "RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*" describe una proteína esencial de función desconocida temprana, YmdA, identificada como una nueva endorribonucleasa (denominada RNasa Y en la actualidad) que era capaz de escindir de forma preferente *in vitro* la corriente arriba de ribointercambio de yitJ 5'-monofosforilado del dominio del aptámero de unión a SAM. Ninguno de Herskovitz M.A. *et al.*, Dasgupta S. *et al.*, o Karen Shahbadian *et al.*, menciona o  
25 sugiere el uso del método SELEX para determinar la preferencia de sustrato de proteínas de unión a híbrido de ADN-ARN o de dos de cinc que se unen a híbridos de ADN-ARN o cualquier método para obtener híbridos de ADN-ARN.  
30

**[0008]** El documento de patente US 2006057590 describe la generación de una molécula de ARN bicatenario que cubre básicamente la región transcrita completa de un gen y la escisión de esta molécula usando una ARN endonucleasa para generar moléculas pequeñas de ARN, que ya están marcadas o se pueden marcar posteriormente. La patente JP54059392 describe una nueva nucleasa B-1 que ataca la zona de una sola escisión de cadena de una de las cadenas de ADN de un ácido desoxirribonucleico bicatenario, y divide de forma específica la otra cadena de ADN en su zona complementaria. Los documentos de patente US 2006057590 y JP54059392  
40 permanecen en silencio con respecto al uso del método SELEX para determinar la preferencia de sustrato de proteínas de unión a híbrido de ADN-ARN o el uso de dedos de cinc para la unión de secuencias específicas en híbridos de ADN-ARN o el uso de ZfQQR en particular para esa finalidad.

**[0009]** Por lo tanto, existe una necesidad de una ribonucleasa que escinda la era de ARN de los híbridos de ADN-ARN en una ubicación específica. También existe una necesidad de un nuevo método mejorado para determinar la preferencia de sustrato de proteínas de unión a híbrido de ADN-ARN o el uso de dedos de cinc para la unión de secuencias específicas en híbridos de ADN-ARN o en particular del uso de ZfQQR para esa finalidad.  
45

### Divulgación de la invención

50 **[0010]** A la vista del estado de la técnica de descrito, el objeto de la presente invención es superar las desventajas indicadas y proporcionar una ribonucleasa que es sin dar la hebra de ARN de dos híbridos de ADN-ARN en una ubicación específica. Por lo tanto, el objeto de la invención es proporcionar enzimas modificadas por ingeniería basadas en la combinación de dominio catalítico de RNasa HI y el dominio de dedos de cinc que reconoce la secuencia en híbridos de ADN-ARN y obtener una enzima específica de secuencia que escinda solamente la hebra de ARN de los híbridos de ADN-ARN.  
55

**[0011]** También se desvela un nuevo método, mejorado para obtener una biblioteca de híbridos de ADN-ARN con una secuencia aleatoria y su uso para determinar la preferencia de secuencia de la proteína de unión a híbrido de ADN-ARN, preferentemente preferencia de secuencia de híbrido de ADN-ARN para unión a dedo de cinc, más preferentemente mediante ZfQQR. Tal método se puede usar de forma eficaz a identificar sistemáticamente biblioteca de híbridos de ADN-ARN. Además, se desvela un método nuevo y mejorado para obtener una biblioteca de híbridos de ADN-ARN.  
60

65 **[0012]** Los inventores han encontrado de forma inesperada que se podía obtener una ribonucleasa H específica de

secuencia modificada por ingeniería que escinde la hebra de ARN de los híbridos de ADN-ARN en una ubicación específica fusionando un dominio catalítico de la RNasa HI o un derivado de la misma con un dominio de unión de híbrido de ADN-ARN a dedos de cinc, en la que el dominio de unión a dedo de cinc tiene la capacidad de unirse a las secuencias específicas en el híbrido de ADN-ARN.

5

**[0013]** En el primer aspecto, la invención proporciona una ribonucleasa que escinde la era de ARN en los híbridos de ADN-ARN, en la que dicha ribonucleasa es una proteína de fusión que comprende un derivado del dominio catalítico de la ribonucleasa HI (RNasa HI) en la que el derivado del dominio catalítico de la RNasa HI es de *Bacillus halodurans*, que comprende polipéptido codificado por los nucleótidos 175 a 588 del gen *mhA* que se muestra en la  
 10 SEC ID N°: 1, y en la que el derivado del dominio catalítico de la RNasa HI comprende al menos una sustitución de un resto de aminoácido en el dominio de unión al sustrato seleccionada entre: K81A, K89E y K123A que hace referencia a la secuencia del polipéptido codificado por la SEC ID N°: 1, preferentemente contiene todas las sustituciones K81A, K89E y K123A que hacen referencia a la secuencia del polipéptido codificada por la SEC ID N°: 1, y en la que el derivado del dominio catalítico de la RNasa HI no comprende dominio de unión a híbrido de RNasa HI,  
 15 con un dominio de unión de híbrido de ADN-ARN a dedos de cinc, en la que el dominio de unión a dedo de cinc tiene la capacidad de unirse a la secuencia específica en el híbrido de ADN-ARN.

**[0014]** En la ribonucleasa preferente, el dominio de dedos de cinc es un derivado del dedo de cinc ZfQQR, preferentemente un polipéptido codificado por los nucleótidos de 19 a 303 de la secuencia ZfQQR que se muestra en la SEC ID N°: 2. La ribonucleasa comprende preferentemente la proteína de fusión catAEA-ZfQQR tal como se muestra en la SEC ID N°: 4. La ribonucleasa comprende preferentemente la proteína de fusión GQ tal como se muestra en la SEC ID N°: 6. La ribonucleasa también comprende preferentemente la proteína de fusión GGKKQ tal como se muestra en la SEC ID N°: 8.

En el siguiente aspecto, la invención proporciona la composición comprende una ribonucleasa de acuerdo con la  
 25 invención.

La invención también se refiere al uso de la ribonucleasa de acuerdo con la invención o una composición de acuerdo con la invención para la escisión *in vitro* de la hebra de ARN en los híbridos de ADN-ARN. En tal uso, la escisión preferente de la hebra de ARN en los híbridos de ADN-ARN se localiza con 2-16 nucleótidos, preferentemente con 5-7 nucleótidos, de separación del sitio de unión con respecto al dedo de cinc.

30

**[0015]** La invención también proporciona el método para obtener variante RNasa HI modificada por ingeniería, que escindir la hebra de ARN en híbridos de ADN-ARN que comprende las siguientes etapas:

a) obtener un derivado de dominio catalítico de la RNasa HI, en el que el derivado del dominio catalítico de la  
 35 RNasa HI es de *B. halodurans*, que comprende polipéptido codificado por los nucleótidos 175 a 588 del gen *mhA* que se muestra en la SEC ID N°: 1, y en el que el derivado del dominio catalítico de la RNasa HI comprende al menos una sustitución de un resto de aminoácido en el dominio de unión al sustrato seleccionada entre: K81A, K89E y K123A que hacen referencia a la secuencia del polipéptido codificada por la SEC ID N°: 1, preferentemente contiene todas las sustituciones K81A, K89E y K123A que hacen referencia a la secuencia del  
 40 polipéptido codificada por la SEC ID N°: 1, y en el que el derivado del dominio catalítico de la RNasa HI no comprende dominio de unión a híbrido de RNasa HI,

b) obtener una RNasa HI modificada por ingeniería mediante la preparación de proteína de fusión que comprende el derivado del dominio catalítico de la RNasa HI obtenido en la etapa a) con un dominio de unión a  
 45 dedo de cinc que tiene la capacidad de unirse a las secuencias específicas en el híbrido de ADN-ARN. En el método preferente para obtener una variante modificada por ingeniería de la RNasa HI, el dominio de dedos de cinc es un derivado del dedo de cinc ZfQQR, preferentemente un polipéptido codificado por los nucleótidos de 19 a 303 de la secuencia ZfQQR que se muestra en la SEC ID N°: 2. El dominio catalítico de la RNasa HI comprende cambios en la región de unión al sustrato.

50

**[0016]** La ribonucleasa que escinde la hebra de ARN de los híbridos de ADN-ARN, contiene una fusión de RNasa HI o derivado de la misma y un dedo de cinc que se une a híbrido de ADN-ARN, en la que la RNasa HI es de *B. halodurans* y el dedo de cinc tiene la capacidad de unirse a las secuencias específicas en el híbrido de ADN-ARN. La ribonucleasa descrita se caracteriza por que es un derivado de la RNasa HI de *B. halodurans*, el péptido del  
 55 dominio catalítico, que comprende un polipéptido codificado por los nucleótidos 175 a 588 del gen *mhA* mostrado en la SEC ID N°: 1. La ribonucleasa descrita se caracteriza por que es un derivado de la RNasa HI nativa, que contiene una supresión del dominio de unión al híbrido de ADN-ARN. La ribonucleasa descrita se caracteriza por que es un derivado del dominio catalítico de la RNasa HI, que contiene una sustitución de aminoácido en la región de unión al sustrato: K81A, K89E y K123A. La ribonucleasa descrita se caracteriza por que es una fusión del polipéptido  
 60 codificado por los nucleótidos 175 a 588 del gen *mhA* que se muestra en la SEC ID N°: 1 y el dominio de dedos de cinc siendo un derivado del dedo de cinc ZfQQR, preferentemente un polipéptido codificado por los nucleótidos de 19 a 303 del dedo de cinc ZfQQR que se muestra en la SEC ID N°: 2.

**[0017]** La ribonucleasa H específica de secuencia se puede usar en un fragmento específico y localizado de ácidos nucleicos para espectrometría de masas de ARN, incluyendo estudios sobre la modificación del ARN. La invención

se puede usar para generar fragmentos de ARN para secuenciación de tercera generación. La RNasa H modificada por ingeniería específica de secuencia se puede usar para detectar o formar mapas de ARN viral con secuencias específicas, en la que el ARN monocatenario se hibrida con ADN, y el híbrido resultante se escinde. La invención se puede usar para la modificación por ingeniería de proteínas mediante fragmentos de barajado y ligación de ARNm, en la que los fragmentos se obtienen por digestión con la ribonucleasa H modificada por ingeniería específica. La invención puede usarse para la escisión específica del sitio directa de híbridos de ADN-ARN persistentes *in vivo*.

**[0018]** Se pueden obtener enzimas con nuevas características mediante la construcción de fusiones de varios dominios con diferente funcionalidad. La modificación por ingeniería de una ribonucleasa que es fin de la hebra de ARN del híbrido de ADN-ARN una manera dependiente de la secuencia se basa en la fusión de los dominios de proteínas: RNasa HI modificada por ingeniería y un dedo de cinc que se une a híbrido de ADN-ARN. La RNasa HI de *B. halodurans* es una enzima que hidroliza la hebra de ARN en un híbrido de ADN-ARN de una manera independiente de la secuencia. El dedo de cinc ZfQQR tiene la capacidad de unirse a una secuencia bien definida en el híbrido de ADN-ARN. En una de las realizaciones de la enzima de fusión, el dominio que presenta una actividad de ribonucleasa frente a la hebra de ARN en un híbrido de ADN-ARN es un fragmento de un gen de *B. halodurans*, *rnhA*, que codifica el dominio catalítico (denominado *cat*). Se trata de un fragmento de los nucleótidos 175 a 588 del gen *rnhA*, que corresponde a la región de 59 a 196 restos de aminoácidos de la proteína nativa RNasa HI de *B. halodurans*. A partir del gen nativo, se retiró un fragmento que codifica el dominio de unión a híbrido (HBD), debido a su capacidad para unirse a híbridos de ADN-ARN independientemente de la secuencia. La fusión del dominio catalítico con ZfQQR se denominó **cat-ZfQQR**. El dominio catalítico modificado por ingeniería tiene tres sustituciones de aminoácidos introducidas en la región de unión al sustrato: K81A, K89E y K123A. Las sustituciones implican a las lisinas cargadas de forma positiva, que se localizan cerca de la región conocida de unión al sustrato. No se pretende que las sustituciones en el sitio de unión inactiva en a la enzima, sino que hagan que la unión al sustrato de la enzima sea dependiente de la presencia del dominio de unión al híbrido de ADN-ARN adicional. El dominio, que transmite la especificidad de secuencia de la enzima de fusión, es el dedo de cinc ZfQQR. En la enzima de fusión, se usó el fragmento genético que codifica un dedo de cinc ZfQQR de 19 a 303 nucleótidos, que corresponde a una región de 7 a 101 restos de aminoácidos de una proteína. Además, la región conectora de interdominios de la enzima de fusión se modificó para producir dos variantes, denominadas **GQ** y **GGKKQ**. GQ es la fusión del dominio catalítico con sustituciones en K81A, K89E y K123A con ZfQQR que carece de un fragmento de codificación de aminoácidos en las posiciones 138-148 de la enzima de fusión. GGKKQ es la fusión del dominio catalítico con sustituciones en K81A, K89E y K123A con ZfQQR que carece de un fragmento de codificación de aminoácidos en las posiciones 138-139 y 141-146 de la enzima de fusión. Las descripciones de las construcciones generadas se resumen en la Tabla 1.

**Tabla 1.** Descripción de construcciones, sus abreviaturas usadas adicionalmente y en los ejemplos y sus referencias a las SEC ID N<sup>os</sup>.

Descripción	Abreviatura	Referencias a las SEC ID N <sup>os</sup>
dominio catalítico de la RNasa HI de <i>B. halodurans</i>	<b>cat</b>	fragmento codificado por los nucleótidos 175 a 588 de la SEC ID N <sup>o</sup> : 1
la fusión del dominio catalítico con ZfQQR	<b>cat-ZfQQR</b>	fragmento codificado por los nucleótidos 175 a 588 de la SEC ID N <sup>o</sup> : 1 y los nucleótidos 19 a 303 de la SEC ID N <sup>o</sup> : 2
la fusión del dominio catalítico con sustituciones de K81A, K89E y K123A con ZfQQR	<b>catAEA-ZfQQR</b>	construcción codificada por la SEC ID N <sup>o</sup> : 3 (secuencia de nucleótidos) y la SEC ID N <sup>o</sup> : 4 mostrada (secuencia de aminoácido)
la fusión del dominio catalítico con sustituciones que introducen cambios en la secuencia de aminoácidos de una proteína en posiciones K81A, K89E y K123A con el dedo de cinc ZfQQR y el conector interdominios acortado con un fragmento que codifica aminoácidos en las posiciones 138-148	<b>GQ</b>	construcción codificada por la SEC ID N <sup>o</sup> : 5 (secuencia de nucleótidos) y la SEC ID N <sup>o</sup> : 6 (secuencia de aminoácidos)
la fusión del dominio catalítico con sustituciones que introducen cambios en la secuencia de aminoácidos de una proteína en posiciones K81A, K89E y K123A con el dedo de cinc ZfQQR y el conector interdominios acortado con un fragmento que codifica aminoácidos en las posiciones 138-139 y 141-146	<b>GGKKQ</b>	construcción codificada por la SEC ID N <sup>o</sup> : 7 (secuencia de nucleótidos) y la SEC ID N <sup>o</sup> : 8 (secuencia de aminoácidos)

**[0019]** También se desvela el método para determinar la preferencia de sustrato de la proteína(s) de unión a híbrido de ADN-ARN o su dominio(s). Tal método permite determinar la secuencia preferente reconocida por una proteína(s) o su dominio(s) y/o la unión mediante un dominio de proteína en el híbrido de ADN-ARN.

5 **[0020]** En el método desvelado para determinar la preferencia de secuencia de la proteína(s) de unión a híbrido de ADN-ARN o su dominio(s), el método comprende las siguientes etapas:

- 10 a) poner en contacto la proteína purificada o su dominio, con una mezcla de una biblioteca de sustratos de híbridos de ADN-ARN, en la que los sustratos de híbridos de ADN-ARN comprenden secuencias aleatorias en la parte central, preferentemente ordenadas al azar en posiciones de 9 o 10 nucleótidos, con secuencias de flanco fijadas y que permite que una proteína o su dominio sometidos a ensayo se unan con la secuencia hacia la que tiene una afinidad;
- 15 b) separar los híbridos de ADN-ARN no unidos mediante inmovilización de proteína o su dominio, con híbrido de ADN-ARN unido, a la resina, preferentemente glutatión agarosa;
- 15 c) retirar los híbridos sin unir;
- d) aislar la proteína recombinante, junto con los híbridos de ADN-ARN asociados, preferentemente mediante la adición de un tampón que contiene glutatión;
- 20 e) amplificación del individuo aislado usando PCR, preferentemente RT-PCR, en la que en la reacción de amplificación, se usan los cebadores complementarios a secuencias conocidas en las secuencias de flanco de la región clasificada al azar, y en la que durante la reacción de PCR en uno de los cebadores, se añade la secuencia de promotor de ARN polimerasa para obtener ADN bicatenario para transcripción *in vitro* con ARN polimerasa;
- 25 f) la transcripción inversa se realiza usando transcriptasa inversa que no posee la actividad de RNasa H y un cebador de ADN complementario con el extremo en la posición 3' del molde de ARN para obtener un híbrido de ADN - ARN, y las etapas de a) a f) se repiten preferentemente.

Por ejemplo, en el método, la secuencia promotora comprende la secuencia promotora para la T7 ARN polimerasa y la ARN polimerasa es T7 ARN polimerasa. Por ejemplo, la proteína o su dominio es una proteína recombinante o dominio recombinante. La proteína o su dominio comprenden preferentemente dedo(s) de cinc. Por ejemplo, la proteína o su dominio es la fusión del dominio de dedos de cinc y GST, preferentemente la fusión del dominio de dedos de cinc y GST codificada por la SEC ID N°: 37. La proteína o su dominio a modo de ejemplo es la fusión de la RNasa HI o un derivado de la misma, por ejemplo la RNasa HI es de *B. halodurans*, que comprende por ejemplo polipéptido codificado por los nucleótidos 175 a 588 del gen *mha* mostrado en la SEC ID N°: 1, y un dedo de cinc. El dedo de cinc tiene por ejemplo la capacidad para unirse a las secuencias específicas en el híbrido de ADN-ARN, el 35 dedo de cinc a modo de ejemplo es ZfQQR, por ejemplo con secuencia codificada por los nucleótidos de 19 a 303 de la secuencia ZfQQR que se muestra en la SEC ID N°: 2.

**[0021]** La descripción también desvela un método para obtener biblioteca de híbridos de ADN-ARN que comprende las siguientes etapas:

- 40 a) amplificación con PCR de la biblioteca de oligonucleótidos de ADN que contiene una secuencia degenerada en la posición central, flanqueada con secuencias invariables, que incluyen una secuencia promotora para la ARN polimerasa,
- 45 b) síntesis de hebra de ARN usando ARN polimerasa en la obtenida en a) ADN bicatenario como un molde,
- c) transcripción inversa del cebador complementario con la secuencia invariable presente en el extremo en la posición 3' del ARN obtenido en b) con una transcriptasa inversa que no posee una actividad de RNasa H,

50 en el que durante la transcripción inversa, se obtienen la hebra de ARN no se degrada, y el híbrido que consiste en hebras de ARN y ADN totalmente complementarias. Por ejemplo, en tal método, los oligonucleótidos contienen una secuencia degenerada en la posición central que comprende posiciones clasificadas al azar de 9 o 10 nucleótidos. Por ejemplo, la secuencia promotora comprende la secuencia promotora para la T7 ARN polimerasa y la ARN polimerasa es la T7 ARN polimerasa.

**[0022]** Para determinar el intervalo completo de secuencias reconocidas por una proteína de unión al híbrido de ADN-ARN específica de secuencia y la preferencia de sustrato de la proteína(s) o sus dominio(s) de unión a híbrido de ADN-ARN, en los híbridos de ADN-ARN, tal como se describe, se desarrolló una modificación del método SELEX, que permite ciclos iterativos de selección de una secuencia de híbrido de ADN-ARN unidad de forma preferente mediante una proteína (Figura 10). En este proceso, un nuevo elemento es el modo de amplificación y recreación de una biblioteca de híbridos de ADN-ARN, para su uso en el siguiente. Esto implica el uso de 60 amplificación con PCR de una biblioteca de oligonucleótidos de ADN que contienen una secuencia degenerada en la posición central, flanqueada por secuencias invariables, que incluyen una secuencia promotora para la ARN polimerasa, preferentemente la T7 ARN polimerasa. El ADN bicatenario resultante se usa como un molde para la síntesis de la hebra de ARN usando ARN polimerasa, preferentemente la T7 ARN polimerasa. Se obtiene una combinación de ARN monocatenario y sirve como un molde para transcripción inversa para obtener una biblioteca 65 de híbridos de ADN-ARN. La transcripta inversa extiende el cebador complementario a la secuencia invariable

presente en el extremo en la posición 3' del ARN. Durante la transcripción inversa con una enzima con la actividad de ribonucleasa H eliminada (la transcriptasa inversa que no posee actividad de ribonucleasa H), la hebra de ARN no se degrada, lo que permite la creación de un híbrido que consiste en una hebra de ARN y de ADN totalmente complementarias. De esta manera, se obtiene una biblioteca de híbridos de ADN-ARN, porque el molde de ARN era heterogéneo en la parte central de la secuencia. La eficacia de la modificación desarrollada del método SELEX se confirmó usando el dedo de cinc ZfQQR como ligando, que se une al híbrido de ADN-ARN. Como resultado de la modificación por ingeniería se obtuvo ZFQQR y se une a la secuencia 5' 'GGGAAGAA-3' en la hebra de ADN de los híbridos de ADN-ARN (Shi y Berg, Specific DNA-RNA hybrid binding by zinc finger proteins. 1995. Science, vol. 268). Para este ejemplo, se usó la fusión de ZFQQR con dominio de Glutatión S-transferasa (denominado GST).

10

### Breve descripción de las figuras

[0023] Para entender mejor la invención, en las figuras se muestran varios ejemplos, en los que

15 La **Figura 1** muestra la secuencia de los cebadores de la SEC ID N°: 11-25 usados en la preparación de la construcción de ADN final de una fusión genética del gen *mhA* con el dedo de cinc ZfQQR y la preparación de sustituciones en la RNasa HI de la secuencia genética de *B. halodurans*.

20 La **Figura 2** es una secuencia del híbrido de ADN-ARN mostrado en la SEC ID N°: 9 (Fig. 2(A) -hebra de ARN) y la SEC ID N°: 10 (Fig. 2(B) - hebra de ADN) que contiene el sitio de unión para el dedo de cinc ZfQQR, que se usó en ensayos de digestión como sustrato específico.

25 La **Figura 3** muestra los productos de digestión resultantes del corte del sustrato de híbrido de ADN-ARN que contiene el sitio de unión para el dedo de cinc, y la dependencia de la escisión en presencia de iones magnesio y cinc. La reacción de digestión mostrada en las Calles 1-14 se realizó con 0,05  $\mu$ M del sustrato marcado con radiactividad, 0,5  $\mu$ M del sustrato no marcado, Tris 25 mM (pH 8,0), KCl 100 mM, DTT 2 mM, 30 min a 37 °C con presencia o no de diferentes enzimas y presencia de MgCl<sub>2</sub> 5 mM y/o ZnSO<sub>4</sub> 20  $\mu$ M como se indica a continuación. Calle 1: sustrato sin escindir, Calle 2: 12,5 nM de RNasa HI de *B. halodurans*, MgCl<sub>2</sub> 5 mM, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 3: 625 nM de **cat**, MgCl<sub>2</sub> 5 mM, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 4: **cat-ZfQQR** 5 nM, MgCl<sub>2</sub> 5 mM, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 5: **catAEA-ZfQQR** 25 nM, MgCl<sub>2</sub> 5 mM, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 6: 12,5 nM de RNasa HI de *B. halodurans*, MgCl<sub>2</sub> 5 mM, Calle 7: 625 nM de **cat**, MgCl<sub>2</sub> 5 mM, Calle 8: **cat-ZfQQR** 5 nM, MgCl<sub>2</sub> 5 mM, Calle 9: 25 nM de **catAEA-ZfQQR**, MgCl<sub>2</sub> 5 mM, Calle 10: 12,5 nM de RNasa HI de *B. halodurans*, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 11: 625 nM de **cat**, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 12: 5 nM de **cat-ZfQQR**, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 13: 25 nM de **catAEA-ZfQQR**, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 14: Marcador con un tamaño de 10-100 nucleótidos de ARN monocatenario marcado con radiactividad con isótopo de fósforo 33 (USB) en el que un asterisco (\*) marcaba un único sitio de escisión.

35 La **Figura 4** muestra los productos de digestión resultantes del corte del sustrato de híbrido de ADN-ARN, que contiene el sitio de unión para el dedo de cinc, y la dependencia de la escisión mediante la variante GQ del **catAEA-ZfQQR** en presencia de iones magnesio y cinc. La reacción de digestión mostrada en las Calles 1-14 se realizó con 50 nM de la variante **GQ**, 0,05  $\mu$ M del sustrato marcado con radiactividad, 0,5  $\mu$ M de sustrato sin marcar, Tris 25 mM (pH 8,0), ZnSO<sub>4</sub> 20  $\mu$ M, KCl 100 mM, DTT 2 mM, 30 min a 37 °C con diversas concentraciones de MgCl<sub>2</sub> como se indica a continuación. Calle 1: sustrato sin escindir, Calle 2: MgCl<sub>2</sub> 0,05 mM, Calle 3: 0,1 mM MgCl<sub>2</sub> Calle 4: MgCl<sub>2</sub> 0,2 mM, Calle 5: MgCl<sub>2</sub> 0,5 mM, Calle 6: MgCl<sub>2</sub> 1 mM, Calle 7: MgCl<sub>2</sub> 2 mM, Calle 8: MgCl<sub>2</sub> 5 mM, Calle 9: MgCl<sub>2</sub> 10 mM, Calle 10: Marcador con un tamaño de 10-100 nucleótidos de ARN monocatenario marcado con radiactividad con isótopo de fósforo 33 (USB).

40 La **Figura 5** muestra los productos de digestión resultantes del corte del sustrato de híbrido de ADN-ARN, que contiene el sitio de unión para el dedo de cinc, y la dependencia de la escisión mediante la variante GGKKQ del **catAEA-ZfQQR** en presencia de iones magnesio y cinc. La reacción de digestión mostrada en las Calles 1-14 se realizó con variante de GGKKQ 50 nM, 0,05  $\mu$ M del sustrato marcado con radiactividad, 0,5  $\mu$ M de sustrato sin marcar, Tris 25 mM (pH 8,0), ZnSO<sub>4</sub> 20  $\mu$ M, KCl 100 mM, DTT 2 mM, 30 min a 37 °C con diversas concentraciones de MgCl<sub>2</sub> como se indica a continuación: Calle 1: sustrato sin escindir, Calle 2: MgCl<sub>2</sub> 0,05 mM, Calle 3: MgCl<sub>2</sub> 0,1 mM, Calle 4: MgCl<sub>2</sub> 0,2 mM, Calle 5: MgCl<sub>2</sub> 0,5 mM, Calle 6: MgCl<sub>2</sub> 1 mM, Calle 7: MgCl<sub>2</sub> 2 mM, Calle 8: MgCl<sub>2</sub> 5 mM, Calle 9: MgCl<sub>2</sub> 10 mM, Calle 10: Marcador con un tamaño de 10-100 nucleótidos de ARN monocatenario marcado con radiactividad con isótopo de fósforo 33 (USB).

50 La **Figura 6** muestra la formación de mapas del sitio de escisión generado por **catAEA-ZfQQR**, la variante **GQ** y la variante de **GGKKQ** en el ARN monocatenario de sustrato del híbrido de ADN-ARN que contiene un sitio de unión para el dedo de cinc ZfQQR. La reacción de digestión mostrada en las Calles 1-14 se realizó con 0,05  $\mu$ M del sustrato marcado con radiactividad, 0,5  $\mu$ M de sustrato sin marcar, Tris 25 mM (pH 8,0), KCl 100 mM, DTT 2 mM, 30 min a 37 °C con presencia o no de diferentes enzimas y presencia de concentración apropiada de MgCl<sub>2</sub> y/o ZnSO<sub>4</sub> 20  $\mu$ M como se indica a continuación. Calle 1: sustrato sin escindir, Calle 2: 35 nM de **catAEA-ZfQQR**, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 3: 50 nM de la variante **GQ**, MgCl<sub>2</sub> 1mM, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 4: 50 nM de la variante de **GGKKQ**, MgCl<sub>2</sub> 2 mM, ZnSO<sub>4</sub> 20  $\mu$ M, Calle 5: escisión de ARNss de ribonucleasa T 1, Calle 6: La hidrólisis alcalina del sustrato de ARNss.

65

La **Figura 7** ilustra posiciones de escisión de los híbridos de ADN-ARN de sustrato que contienen el sitio de unión para el dedo de cinc ZfQQR mediante el **catAEA-ZfQQR**, indicado con flechas en la parte superior de la secuencia, la flecha de color negro más grande indica el sitio de escisión principal, las flechas más pequeñas indican sitios adicionales (hebra de ARN de la parte superior, la hebra de ADN de la parte inferior, el sitio de unión para ZfQQR marcado con caja).

La **Figura 8** ilustra posiciones de escisión de los híbridos de ADN-ARN de sustrato que contienen el sitio de unión para el dedo de cinc ZfQQR con la variante **GQ** del catAEA-ZfQQR, indicado con flechas en la parte superior de la secuencia, la flecha de color negro más grande indica el sitio de escisión principal, las flechas más pequeñas indican sitios adicionales (hebra de ARN de la parte superior, la hebra de ADN de la parte inferior, el sitio de unión para ZfQQR marcado con caja).

La **Figura 9** ilustra posiciones de escisión de los híbridos de ADN-ARN de sustrato que contienen el sitio de unión para el dedo de cinc ZfQQR con la variante de **GGKKQ** del catAEA-ZfQQR, indicado con flechas en la parte superior de la secuencia, la flecha de color negro más grande indica el sitio de escisión principal, las flechas más pequeñas indican sitios adicionales (hebra de ARN de la parte superior, la hebra de ADN de la parte inferior, el sitio de unión para ZfQQR marcado con caja).

La **Figura 10** muestra un diagrama que representa las diferentes etapas de una ronda del procedimiento de SELEX modificado.

La **Figura 11** muestran las secuencias de moldes de oligonucleótido de ADN para obtener ARN monocatenario en el proceso de transcripción *in vitro* usando la T7 ARN polimerasa de bacteriófago. El molde A (SEC ID N°: 26) contiene un sitio de unión para el dedo de cinc ZfQQR. El molde B (SEC ID N°: 27), no contiene sitio de unión para el dedo de cinc ZfQQR y tiene un sitio de escisión único para la enzima de restricción XhoI. El molde 9N (SEC ID N°: 28), contiene nueve regiones degeneradas de nucleótidos (el nucleótido degenerado individual marcado como "N").

La **Figura 12** ilustra el proceso de obtención de un híbrido de ADN-ARN. Los fragmentos de ADN se resolvieron en gel poliacrilamida de desnaturalización al 15 % que contenía urea 6 M. Los extremos en la posición 5' de oligonucleótido de ADN con una longitud de 55 nucleótidos y el cebador 2 se marcaron con isótopo radiactivo de fósforo 33. Calle 1: 0,5 pmol del cebador 2 marcado, Calle 2: 0,25 pmol de producto de transcripción inversa, Calle 3: 0,25 pmol de oligonucleótido de ADN marcado de 55 nt.

La **Figura 13** ilustra el proceso de obtención de un híbrido de ADN-ARN. Los fragmentos de ADN se resolvieron en gel poliacrilamida nativo al 15 %. Los extremos en la posición 5' de oligonucleótido de ADN con una longitud de 55 nucleótidos y el cebador 2 se marcaron con isótopo radiactivo de fósforo 33. Calle 1: 0,5 pmol del cebador 2 marcado, Calle 2: 0,25 pmol de producto de transcripción inversa, Calle 3: 0,25 pmol de productos de transcripción inversa que digieren 5 unidades de ribonucleasa H, Calle 4: 0,25 pmol de productos de transcripción inversa que digieren una unidad de ribonucleasa H, Calle 5: 0,25 pmol de oligonucleótido de ADN marcado de 55 nt.

La **Figura 14** muestra un gráfico de barras que ilustra el cambio en la proporción entre el ADN bicatenario que contiene un sitio de unión para el dedo de cinc ZfQQR (A) y el ADN bicatenario que no contiene un sitio de unión para ZfQQR (B) en rondas sucesivas del procedimiento de SELEX ("O" significa la entrada, mezcla de partida del híbrido A con B en la proporción de 1:10000, I-V son números de rondas sucesivas).

La **Figura 15** ilustra una secuencia logo derivada de 40 secuencias usando WebLogo.

La **Figura 16** muestran las secuencias de los híbridos de ADN-ARN usadas para determinar la constante  $K_D$  para el dedo de cinc ZfQQR, en la que la Figura 16A es un híbrido de ADN-ARN que contiene el sitio de unión que se describe en la bibliografía y la Figura 16B es un híbrido de ADN-ARN con la secuencia consenso después de cinco rondas de SELEX.

La **Figura 17** muestra un gráfico que representa el cambio de la cantidad de sustrato marcado con radiactividad retenida en un filtro de nitrocelulosa, dependiendo de la concentración del dedo de cinc ZfQQR en la mezcla de reacción (expresado como porcentaje unido).

### Descripción de realizaciones

**[0024]** Los ejemplos que siguen a continuación se presentan solamente para ilustrar la invención y para explicar ciertos aspectos de la misma y no para limitar la invención, por lo tanto, no se deberían identificar con su alcance completo, que se define en las reivindicaciones adjuntas.

**[0025]** En los siguientes ejemplos, a menos que se indique de otro modo, se usan materiales y métodos

convencionales tal como se describe en Sambrook J. *et al.*, "Molecular Cloning: A Laboratory Manual, 2ª edición. 1989. Cold Spring Harbor, N.Y. Cold Spring Harbor Laboratory Press, o procediendo de acuerdo con las recomendaciones de los fabricantes para materiales y métodos específicos. Como se usa en el presente documento, a menos que se indique de otro modo con abreviaturas convencionales para aminoácidos y nucleótidos o  
5 ribonucleótidos.

#### **Ejemplo 1 Clonación del fragmento del gen *rnhA* de *Bacillus halodurans* en el vector de expresión pET 30b**

**[0026]** El vector pET15a que porta el gen *rnhA* (SEC ID N°:1) (la SEC ID N°:1 es la secuencia de ADN de la RNasa HI de *rnhA* (ID del Gen 893801, BH0863) de *B. halodurans* y se obtuvo a partir de fuentes privadas. Un fragmento del gen que codifica *cat* (dominio catalítico) se anticipó usando la técnica de PCR en condiciones convencionales realizadas en el pET15a que portan un gen *rnhA* como el molde y 1 U de Phusion polimerasa (New England Biolabs) con 50 pmol de cada ADN de cebador Bhcatr (SEC ID N°: 12) y Bhcatf (SEC ID N°: 11) (véase la Figura 1) que codifica el fragmento del gen *rnhA* que corresponde a 175 - 588 nucleótidos y el vector pET 30b de ADN (compañía Novagen) se digirió con NdeI y KpnI (Fermentas). La mezcla de reacción que contenía el ADN se separó en gel de agarosa al 0,7 % en tampón de TAE y los fragmentos correspondientes a los tamaños esperados de 430 pb y 5313 pb, respectivamente, se volvieron a aislar a partir del gel usando un kit para reisolamiento (Gel out, A & A Biotechnology). A continuación, se volvieron a aislar 50 ng del fragmento del gen *rnhA* y 25 ng del vector de corte pET 30b se trataron con T4 ADN ligasa de bacteriófago (Fermentas) y la ligasa se inactivó con calor. A continuación, la mezcla de ligación se usó para transformación de células bacterianas. Se transformaron bacterias competentes de *E. coli* (Top 10 (Invitrogen)), con mezcla de ligación (20 - 200 ng por 50 µl de células bacterianas). Los transformantes se seleccionaron en placas de LB-agar complementado con 30 µg/ml kanamicina. La selección de transformantes que contenía los recombinantes deseados se basó en análisis de mapas de restricción, y a continuación las muestras se sentenciaron para confirmar la secuencia de las construcciones.  
10  
15  
20  
25

#### **Ejemplo 2 Clonación del gen que codifica un dedo de cinc ZfQQR para el vector de expresión pET28b**

**[0027]** La síntesis del gen que codifica el dedo de cinc ZfQQR se encargó en Epoch's Life Sciences (véase la SEC ID N°: 2 que es la secuencia del ADN sintetizado de un gen de dedo de cinc ZFQQR) sobre la base de la secuencia de aminoácidos en el artículo "Shi Y, Berg JM. Specific DNA-RNA hybrid binding proteins. Science. 1995. vol. 268, 282-284). El ADN que codifica ZfQQR y el vector pET28b de ADN (compañía Novagen) se digirió con las enzimas de restricción NcoI y XhoI (las enzimas eran de la Compañía Fermentas, las reacciones se realizaron en un tampón 2X Yellow de acuerdo con las instrucciones del fabricante, 1 unidad de enzima por 1 mg de ADN durante 1 h a 37 °C). La mezcla de reacción se separó en gel de agarosa al 0,7 % en tampón de TAE y se trató tal como en el Ejemplo 1. A continuación, se ligaron 50 ng del gen de ZfQQR reaislado y 25 ng del vector pET28b digerido con T4 ADN Ligasa (Fermentas, la reacción se realizó en tampón proporcionado por el fabricante) a temperatura ambiente durante 1 hora. La ligasa se inactivó con calor mediante incubación a 75 °C durante 10 min. A continuación, la mezcla de ligación se usó para transformar células bacterianas tal como en el Ejemplo 1.  
30  
35  
40

#### **Ejemplo 3 Clonación del gen que codifica el dedo de cinc ZfQQR para el vector de expresión pET 30b que contiene un fragmento de un gen *rnhA* de *Bacillus halodurans***

**[0028]** Un fragmento del gen que codifica ZfQQR a partir del vector pET28b se amplificó con la técnica de PCR. La reacción se realizó en condiciones convencionales realizadas en el pET28b que el gen de ZfQQR como el molde y 1 U de Phusion polimerasa (New England Biolabs) con 50 pmol de cada cebador, BhZf (SEC ID N°: 13) y Kmr (SEC ID N°: 15) (véase la Figura 1). Un fragmento del gen que codifica el dominio catalítico de la RNasa HI en el vector pET 30b se amplificó con la técnica de PCR. La reacción se realizó en condiciones convencionales realizadas en el vector pET 30b que porta el gen de *cat* como el molde y 1 U de Phusion polimerasa (New England Biolabs) con 50 pmol de cada cebador Kmf (SEC ID N°: 14) y Bhcatr (SEC ID N°: 12) (véase la Figura 1). Los productos de PCR purificados se fosforilaron con T4 polinucleótido quinasa de bacteriófago (Fermentas). El ADN de ambas reacciones (20 ng) se combinó, se ligó y las mezclas de ligación se transformaron en células bacterianas tal como en el Ejemplo 1.  
45  
50

#### **Ejemplo 4 Mutagénesis del gen que codifica la enzima de fusión que introduce sustituciones de aminoácidos en las posiciones de K81A, K89E y K123A**

**[0029]** La ligación de los fragmentos generados en el Ejemplo 3 creó una construcción en la que el marco de lectura abierto contenía al inicio una parte del gen que codifica *cat* y el fragmento genético que codifica ZfQQR al final. La construcción de ADN obtenida servía como un molde para introducir sustituciones usando la técnica de PCR en los nucleótidos que codifican los restos de aminoácidos en las posiciones 81, 89 y 123 de la proteína. Se introdujeron sustituciones en etapas, en primer lugar se sustituyeron los nucleótidos que codifican los restos 81 (sustitución de lisina por alanina) y 89 (sustitución de lisina por ácido glutámico), solamente después de obtener tal construcción, se realizó PCR para convertir los nucleótidos que codifican el resto en la posición 123 (sustitución de lisina por alanina). La reacción de PCR se realizó en condiciones convencionales realizadas en el pET30b que porta el gen de *cat-ZfQQR* como el molde y 1 U de Phusion polimerasa (New England Biolabs) con 50 pmol de cada cebador K81Af  
60  
65

(SEC ID N°: 17), K81Ar (SEC ID N°: 16), K89Ef (SEC ID N°: 19), K89Er (SEC ID N°: 18), K123Af (SEC ID N°: 21) y K123Ar (SEC ID N°: 20) (véase la Figura1).

5 **[0030]** Los productos de PCR purificados se trataron tal como en el Ejemplo 1 y las mezclas de ligación se transformaron en células bacterianas. La secuencia final del gen que codifica el **catAEA-ZfQQR** se muestra en la SEC ID N°: 3 y la secuencia de aminoácidos resultante de la secuencia de proteínas se muestra en la SEC ID N°: 4.

**Ejemplo 5 Mutagénesis del gen que codifica la enzima de fusión con sustituciones de aminoácidos en las posiciones de K81A, K89E y K123A que acortar la longitud de la región conectora de interdominios**

10

**[0031]** La construcción generada en el Ejemplo 4 porta el marco de lectura abierto que codifica el **catAEA-ZfQQR** en el extremo y sirvió como un molde para acortar el conector de inter dominios usando la técnica de PCR en la región que codifica el conector interdominios, 409-449, en el gen en la SEC ID N°: 3. La reacción de PCR se realizó en condiciones convencionales realizadas en el pET30b que exporta el gen de **catAEA-ZfQQR** como el molde y 1 U de Phusion polimerasa (New England Biolabs) con 50 pmol de cada cebador, dell1f (SEC ID N°: 22) y del11r (SEC ID N°: 23) (que se usaron para generar la variante denominada GQ), o los cebadores del5f (SEC ID N°: 24) y del5r (SEC ID N°: 25) (que se usaron para generar la variante denominada GGKKQ) respectivamente (véase la Figura 1 para secuencias de cebadores).

20 **[0032]** Los productos de PCR purificados se trataron tal como en el Ejemplo 1 y las mezclas de ligación se transformaron en células bacterianas. La secuencia final del gen que codifica la variante de GQ que es el catAEA-ZfQQR con conector inter dominios acortado con un fragmento que codifica los aminoácidos en las posiciones 138-148 mostrados en la SEC ID N°: 5 y la secuencia de aminoácidos resultantes de la secuencia de proteínas mostrada en la SEC ID N°: 6. La secuencia final del gen que codifica la variante de GGKKQ que es el catAEA- ZfQQR con el conector inter dominios acortado con un fragmento que codifica aminoácidos en las posiciones 138-139 y 141-146 mostradas en la SEC ID N°: 7 y la secuencia de aminoácidos resultante de la secuencia de proteínas mostrada en la SEC ID N°: 8.

**Ejemplo 6 Expresión y purificación de proteínas mediante cromatografía de intercambio iónico**

30

**[0033]** Cada uno de, el plásmido pET15a con el gen que codifica la RNasa HI de longitud completa de *B. halodurans*, pET 30b con el gen que codifica **cat**, pET 30b con el gen que codifica el **catAEA-ZfQQR** y pET 30b con el gen que codifica la variante de **GQ** y pET30 con el gen que codifica la variante de **GGKKQ** se transformó en la cepa ER2566 de en *E. coli* (New England Biolabs). La expresión de proteínas se indujo con IPTG y las proteínas se purificaron de acuerdo con el protocolo convencional para resina de Ni-NTA (Sigma Aldrich).

35

**Ejemplo 7 Ensayo de actividad de nucleasa de la enzima de fusión**

40 **[0034]** Se sometió a ensayo el efecto de la presencia de iones magnesio y cinc en reacción sobre la actividad y la especificidad enzimática de la RNasa HI de longitud completa de *B. halodurans*, **cat**, **cat-ZfQQR** y **catAEA-ZfQQR**. El ensayo de actividad incluía la presencia de MgCl<sub>2</sub> 5 mM y/o ZnSO<sub>4</sub> 20 μM. El extremo la posición 5' de la hebra de ARN en el sustrato de híbrido de ADN-ARN se marcó con isótopo radiactivo de fósforo 33. La reacción de digestión contenía sustrato marcado con radiactividad 0,05 μM y sustrato sin marcar 0,5 μM:

45 ARN:

AGAACUAGUGGAUCAACCGGGCUGCAGGAAUUCGAUAUCAAGCUUAUCGAUACC  
GUGGCGGUUCUUCCTCAAGCC (SEC ID N°: 9)

ADN:

50

GCTTGGGGAAGAACCGCCACGGTATCGATAAGCTTGATATCGAATTCCTGCAGCCC  
GGTTGATCCACTAGTTCT (SEC ID N°: 9)

55 y Tris 25 mM (pH 8,0), KCl 100 mM, 5 mM MgCl<sub>2</sub> y/o ZnSO<sub>4</sub> 20 μM, DTT 2 mM. Las reacciones contenían: 12,5 nM de la RNasa HI de longitud completa de *B. halodurans*, 625 nM de **cat**, 5 nM del **cat-ZfQQR** o 25 nM de **catAEA-ZfQQR**. Las reacciones de digestión se realizaron a 37 °C y se detuvieron después de 30 min mediante la adición de formamida hasta una concentración final de un 50 %. Los productos de digestión se resolvieron en gel poliacrilamida de desnaturalización al 12 % que contenía urea 6 M (véase la Figura 3). El gel se secó en un secador de vacío y se expuso a Storage Phosphor Screens (GE Healthcare) durante una noche. El autoradiograma se escaneó en un escáner Typhoon Trio (GE Healthcare).

En presencia de ZnSO<sub>4</sub> 20 μM, y en ausencia de MgCl<sub>2</sub>, la digestión del sustrato con el **catAEA-ZfQQR** genera productos de un tamaño único. La escisión en este lugar no se produce cuando tal sustrato se difiere con la RNasa HI de longitud completa de *B. halodurans*, **cat** y **cat-ZfQQR**.

5 **[0035]** Se determinó en efecto de la reducción de la longitud del conector sobre la especificidad de la escisión del sustrato con las variantes **GQ** y **GGKKQ**. Las condiciones sometidas a ensayo incluían un intervalo de concentración de iones magnesio en presencia de cinc en la reacción. El ensayo de actividad incluía la presencia de un intervalo de 0,05 mM a 10 mM de MgCl<sub>2</sub> y ZnSO<sub>4</sub> 20 μM y composición de tampón tal como se ha mencionado anteriormente. El extremo en la posición 5' de la hebra de ARN en el sustrato de híbrido de ADN-ARN marcado con isótopo radiactivo  
10 de fósforo 33. La reacción de digestión contenía sustrato marcado con radiactividad 0,05 μM y sustrato sin marcar 0,5 μM (Figura 2), Tris 25 mM (pH 8,0), KCl 100 mM, MgCl<sub>2</sub> 5 mM y ZnSO<sub>4</sub> 20 μM, DTT 2 mM. En la reacción, se usaron 50 nM de la variante de GQ y 50 nM de la variante de GGKKQ. Las reacciones de digestión se realizaron como anteriormente y los resultados se muestran en la Figura 4 para GQ y en la Figura 5 para GGKKQ. Para la variante de **GQ** del conector, la concentración óptima de Mg<sup>2+</sup> en la reacción era 1 mM y para la variante de **GGKKQ**  
15 era 2 mM.

**[0036]** Para determinar de forma precisa el sitio de escisión en el sustrato de híbrido de ADN-ARN que contiene un sitio de unión para el dedo de cinc ZfQQR generado con el **catAEA-ZfQQR** y sus variantes de **GQ** y **GGKKQ**, los productos de escisión se separaron en un gel de poli(acrilamida) gel con una resolución que permitía la separación de  
20 fragmentos que diferían en un nucleótido. El extremo en la posición 5' de la hebra de ARN en el híbrido de ADN-ARN sustrato se marcó con el isótopo radiactivo de fósforo 33. La reacción de digestión contenía sustrato marcado con radiactividad 0,05 μM y sustrato sin marcar 0,5 μM (véase la Figura 2), Tris 25 mM (pH 8,0), KCl 100 mM, ZnSO<sub>4</sub> 20 μM, MgCl<sub>2</sub> 1 mM (en el caso de la variante de GQ) y MgCl<sub>2</sub> 2 mM (en el caso de la variante de GGKKQ) y DTT 2 mM. La reacción contenía catAEA-ZfQQR 0,1 μM, 50 nM de la variante de GQ y 50 nM de la variante de  
25 GGKKQ. Las reacciones de digestión se realizaron a 37 °C y se detuvieron después de 1 h mediante la adición de formamida hasta una concentración final de un 50 %. El tamaño del marcador con fragmentos que diferían en un nucleótido se obtuvo mediante hidrólisis alcalina de acuerdo con las reacciones de Ambion (www.ambion.com). Los productos de digestión se resolvieron en gel de poli(acrilamida) de desnaturalización al 12 % que contenía urea 6 M (véase la Figura 6). Se formaron mapas de los fragmentos resultantes de la longitud de la secuencia identificada en la secuencia del sustrato y se determinó que el producto de escisión principal para el **catAEA-ZfQQR** tenía una longitud de 54 nucleótidos (véase la Figura 7). Este se ubica en la hebra opuesta a 7 nucleótidos del sitio de unión para el dedo de cinc ZfQQR. En el caso de **GQ** y **GGKKQ**, el producto de escisión principal tiene una longitud de 56 nucleótidos (véase la Figura 8 para GQ y la Figura 9 para GGKKQ) y se ubica en la hebra opuesta a 5 nucleótidos del sitio de unión para el dedo de cinc ZfQQR.  
30  
35

**[0037]** La nueva ribonucleasa escinde la hebra de ARN en los híbridos de ADN-ARN, y es una proteína de fusión que comprende el dominio catalítico de la ribonucleasa HI (RNasa HI) con un dominio de unión de híbrido de ADN-ARN a dedos de cinc, en la que el dominio de unión a dedo de cinc tiene la capacidad de unirse a secuencias específicas en el híbrido de ADN-ARN. Las proteínas de fusión obtenidas preferentes son **catAEA-ZfQQR**, **GQ**,  
40 **GGKKQ**.

**[0038]** No era evidente que se podría obtener una RNasa H modificada por ingeniería específica de secuencia mediante la fusión de un dominio de RNasa HI y un dedo de cinc porque el dominio de nucleasa es una enzima de procesamiento que después de la unión al sustrato se escinde múltiples veces en varios sitios del sustrato. Además,  
45 como se ha descrito anteriormente, la fusión de una RNasa H sin cambiar por un dedo de cinc no permite la obtención de una enzima específica de secuencia así como la fusión del fragmento **cat** de la RNasa HI no permite la obtención de una enzima específica de secuencia. Sin desear quedar ligado por teoría alguna, existen algunas razones posibles para esto - la RNasa HI es una enzima de procesamiento que siempre se escindirá en múltiples ocasiones, el dedo de cinc en fusión con otros dominios no se une a su secuencia de unión en híbridos de ADN-  
50 ARN, las condiciones de escisión no eran óptimas, el dominio de cat se une al sustrato independientemente de la presencia del dedo de cinc o todas ellas eran importantes.

**[0039]** Además, no era evidente que las sustituciones en la región de unión al sustrato: K81A, K89E y K123A de cat condujeran a una enzima específica de secuencia. Existen varias hipótesis que pueden proporcionar una explicación  
55 para la no obtención de un específico: el dominio cat se puede unir al sustrato independientemente de la presencia del dominio de cinc, o las sustituciones en la región de unión podrían haber influido en la catálisis de la enzima y haber producido una variante que es inactiva, o las sustituciones de los aminoácidos en proteínas pueden conducir a variantes que no pueden plegarse de forma apropiada y por lo tanto no son solubles. No había datos ni sugerencias disponibles en la técnica anterior que permitieran estar seguro de que la desestabilización de la unión mediante el  
60 dominio de cat permitiera obtener una enzima que se escindiera en una distancia definida desde el sitio de unión.

### **Ejemplo 8 Preparación de híbridos de ADN-ARN de sustrato**

**[0040]** Se sintetizaron tres ADN monocatenarios de 78 nucleótidos de longitud denominados: molde A (SEC ID N°: 26), molde B (SEC ID N°: 27) y molde 9N (SEC ID N°: 28), respectivamente.  
65

**[0041]** Estos oligonucleótido se usaron como un molde para crear una ADN bicatenario usando la técnica de PCR. La reacción se realizó en condiciones convencionales, aproximadamente 0,1 pmol de molde y 1 unidad de Phusion polimerasa (New England Biolabs) con 10 pmol de cada cebador 1 y 2:

- 5     **cebador 1:** AATTTAATACGACTCACTATAGGGCTCTAGATCTCACTAAGCATAG (SEC ID N°: 33)  
       **cebador 2:** GAGATCTAGACGGAACATG (SEC ID N°: 34)

**[0042]** Los cebadores 1 y 2 se usaron para la amplificación de ADN bicatenario en la reacción de PCR y el cebador 2 también se usó en la reacción de transcripción inversa.

10

**[0043]** Condiciones de reacción: desnaturalización inicial a 98 °C durante 1 min, seguido de 15 ciclos de desnaturalización de 15 s a 98 °C, 15 s, renaturalización a 72 °C (una reducción de la temperatura de renaturalización con cada ciclo con 1 °C hasta una temperatura final de 58 °C) y 2 s de extensión a 72 °C, seguido de 10 ciclos de desnaturalización de 15 s a 98 °C, 15 s de renaturalización a 58 °C y 2 s de extensión a 72 °C. La mezcla de reacción se extrajo con fenol y la fase acuosa recogida se precipitó con acetato sódico 0,5 M a pH 4,5 y etanol. El precipitado se secó al aire y se volvió a suspender en agua. Los transcritos de ARN monocatenario se obtuvieron en un molde de ADN bicatenario con la secuencia promotora para la T7 ARN polimerasa de bacteriófago usando un kit para transcripción MEGAscript™ T7 (Ambion). La reacción se realizó en condiciones convencionales con 200 - 500 ng de ADN. La reacción de transcripción se resolvió en gel de poli(acrilamida de desnaturalización al 15 % que contenía urea 6 M y el tampón de TBE y se visualizó usando bromuro de etidio con luz UV. Una banda que correspondía a la longitud de 55 nucleótidos se escindió y la elución del ARN a partir del gel se realizó de acuerdo con procedimientos convencionales usando columnas Costar Spin X (Corning Life Sciences).

15

20

**[0044]** El híbrido de ADN-ARN se obtuvo en un molde de ARN monocatenario usando transcriptasa inversa con actividad de RNasa H eliminada. Se incubaron 200 pmol del cebador 2 ((SEC ID N°: 34) y 200 pmol de la hebra de ARN monocatenario durante 2 min a 70 °C, seguido de 2 min en hielo. El molde preparado se sometió a transcripción inversa. La reacción se realizó en 40 µl en un tampón proporcionado por el fabricante, con una concentración final de mezcla de dNTP 1 mM, 200 pmol de molde hibridado con cebador, 40 unidades de RiboLock (Fermentas) y 400 unidades de RevertAid H menos Transcriptasa inversa (Fermentas) durante 2 h a 42 °C. Los híbridos de ADN-ARN de la mezcla de reacción se separaron del tampón y los dNTP usando la resina G-25 (Sigma Aldrich) en un procedimiento convencional.

25

30

**[0045]** Para confirmar que el protocolo desarrollado conduce a la formación de híbridos de ADN-ARN complementarios, se realizó transcripción inversa (descripción tal como anteriormente) usando un cebador 2 marcado de forma radiactiva con el isótopo de fósforo 33 (SEC ID N°: 34). La síntesis de ADN del tamaño apropiado (la longitud esperada es de 55 nucleótidos) se confirmó mediante la comparación del tamaño del ADN resultante con un control de oligonucleótido de ADN de 55 nucleótidos de longitud. Las muestras se resolvieron en gel de poli(acrilamida desnaturalización al 15 % que contenía urea 6 M en el tampón de TBE (véase la Figura 12). El gel se secó en un secador de vacío y se expuso al Storage Phosphor Screens (GE Healthcare) durante una noche. El audiorradiograma se escaneó en un escáner Typhoon Trio (GE Healthcare). La creación de híbridos de ADN-ARN se confirmó mediante digestión del producto de transcripción inversa con la ribonucleasa H y mediante comparación de la migración del producto de la reacción con un control de oligonucleótido de ADN de 55 nucleótidos de longitud en gel de poli(acrilamida nativo al 15 % en el tampón de TBE (véase la Figura 13). La digestión de las hebras de ARN de los híbridos de ADN-ARN se realizó en 10 µl en el tampón proporcionado por el fabricante con 1 y 5 unidades de RNasa H (Fermentas) durante 30 min a 37 °C.

35

40

45

### **Ejemplo 9 Clonación, expresión y purificación de proteína de fusión de GST con el dedo de cinc ZFQQR**

**[0046]** Para obtener la fusión del dominio de dedos de cinc con GST, se clonó un gen que codifica ZFQQR en el vector de expresión pGEX-4T-1 (Amersham). La síntesis del gen que codifica el dedo de cinc ZfQQR se encargó en Epoch's Life Sciences sobre la base de la secuencia de aminoácidos en el artículo "Shi Y, Berg JM. Specific DNA-RNA hybrid binding proteins. Science. 1995. vol. 268, 282-284). El ADN que codifica ZfQQR se amplificó usando la técnica de PCR. Los cebadores ZFf y ZFr se sintetizaron y se usaron para la preparación de construcción de ADN de la fusión de genes que codifican el dominio de GST y el dedo de cinc ZfQQR. La reacción se realizó en condiciones convencionales y 1 unidad de Phusion polimerasa (New England Biolabs) con 50 pmol de cada cebador, ZFf y ZFr:

50

55

ZFf: GGTTCTGGTGACCCGGG (SEC ID N°: 35)

ZFr: CGGGAAAACAGCATTCCAGGTATTAG (SEC ID N°: 36)

**[0047]** El ADN que codifica el dedo de cinc ZfQQR y el vector pGEX-4T-1 DNA se digirieron con las enzimas de restricción SmaI y XhoI (Fermentas). La digestión de SmaI se realizó en un tampón de 1X Yellow de acuerdo con las recomendaciones del fabricante. Después de este tiempo, el tampón se añadió hasta una concentración final de 2X Yellow en la reacción, 1 unidad de la enzima XhoI para 1 mg de ADN durante 1 h a 37 °C. La reacción se separó en gel de agarosa al 0,7 % en tampón de TAE y los fragmentos que correspondían a los tamaños esperados se volvieron a aislar a partir del gel usando un kit para el nuevo aislamiento a partir de los geles de agarosa (Gel Out, A

60

65

& A Biotechnology). A continuación, se ligaron 50 ng del ADN aislado de nuevo que codifica el dedo de cinc ZfQQR y 25 ng del vector pGEX-4T-1 de corte con T4 ADN ligasa de bacteriófago (Fermentas, la reacción se realizó en tampón proporcionado por el fabricante) a temperatura ambiente durante 1 hora. La ligasa se inactivó con calor incubando a 75 °C durante 10 min. A continuación, la mezcla de ligación se usó a la transformación en una cepa bacteriana, 5 Top10 F, de *E. coli* (Invitrogen). Los transformantes se seleccionaron en 100 µg/ml de ampicilina. La selección de los transformantes adecuados que contenían los recombinantes deseados se basó en el análisis de mapas de restricción, y a continuación los clones se secuencian para confirmar la secuencia de las construcciones.

**[0048]** El gen pGEX-4T-1 de plásmido que codifica el dedo de cinc ZfQQR se transformó en la cepa BL21 de *E. coli* 10 (DE3): (Promega). Los transformantes de un cultivo de una noche se inocularon en líquido LB con 100 mg/ml de ampicilina, y se incubaron a 37 °C durante 2 h. Después de este periodo de tiempo, se añadió IPTG (concentración final 1 mM) y el cultivo se incubó con agitación a 25 °C durante un periodo adicional de 5 horas. Después de la inducción, los cultivos se centrifugaron a 4000 rpm a 4 °C durante 10 min se lavaron con 2 ml de STE, se centrifugaron a 10.000 g durante 10 min. El sedimento resultante del cultivo se volvió a suspender en 35 ml de PBS 15 y a continuación las células bacterianas se desintegraron mediante un solo pasaje a través de la prensa francesa (Constant Systems LTD) a 1360 atmósferas. Los lisados se eliminaron mediante centrifugación en la ultracentrífuga a 20 000 g a 4 °C durante 20 min. La proteína se purificó en resina de Glutación Sepharose (GE Healthcare) de acuerdo con procedimientos convencionales.

## 20 **Ejemplo 10 Unión de proteína a híbrido de ADN-ARN, separación de híbridos sin unir y elución del complejo de proteína-híbrido de ADN-ARN a partir de la resina**

**[0049]** La reacción de unión del híbrido de ADN-ARN con el dedo de cinc ZfQQR se realizó en 40 µl y contenía: 200 pmol de híbridos de ADN-ARN, 1 pmol de fusión de GST con el dedo de cinc ZfQQR, Tris 25 mM a pH 8,0, KCl 25 100 mM, ZnSO<sub>4</sub> 20 µM, DTT 2 mM. La reacción se realizó en temperatura ambiente durante 30 min. Después de este periodo de tiempo, la reacción se añadió a 7,5 µl de resina de Glutación Sepharose y se incubó durante 30 min en Thermomixer compact (Eppendorf) a 22 °C con agitación (1.400 rpm). En la siguiente etapa, la resina se lavó tres veces con 100 µl de tampón P (Tris 25 mM a pH 8,0, KCl 100 mM, ZnSO<sub>4</sub> 20 µM, DTT 2 mM), en cada ocasión, la muestra se centrifugaba a 1000 g durante 2 min a temperatura ambiente y el sobrenadante se retiraba. La siguiente 30 etapa era la elución, que se repitió dos veces con 30 µl de tampón E y un periodo de incubación de 10 min a temperatura ambiente. Después de cada periodo de incubación, la muestra se centrifugó a 1000 g durante 2 min a temperatura ambiente y el sobrenadante se transfirió a un tubo. 5 µl del eluato que contenía la fusión de GST con el dedo de cinc finger asociado con el híbrido de ADN-ARN se sometieron a amplificación usando PCR para obtener ADN bicatenario, a continuación el ARN se transcribió y los híbridos de ADN-ARN se sintetizaron mediante 35 transcripción inversa tal como en el Ejemplo 8. El híbrido de ADN-ARN obtenido era el material de partida para la siguiente ronda de SELEX.

## **Ejemplo 11 SELEX de control**

40 **[0050]** Un control del método SELEX se realizó en una mezcla de híbridos de ADN-ARN que contenían el sitio de unión para el dedo de cinc ZfQQR (híbrido A, se forma en un molde A de la SEC ID N°: 26) y un híbrido, que no tiene tal sitio de unión, en lugar de un sitio de restricción de XhoI (híbrido B, se crea el molde B de la SEC ID N°: 27). Se realizaron cinco rondas de SELEX en una combinación de control que consistía en una mezcla de los híbridos A y B en la proporción de 1:10000. Para distinguir las dos secuencias, se digirieron 100 ng de ADN resultantes de 45 amplificación con PCR del eluato después de cada ronda con 2 unidades de enzima XhoI durante 16 h a 37 °C. Los productos de digestión se separaron en gel de poliacrilamida nativo al 15 % en el tampón de TBE. El gel se incubó 2 min en una solución de bromuro de etidio hasta una concentración final de 0,5 µg/ml. Las bandas se visualizaron con luz UV usando una cámara CCD para el archivo digital de imágenes Multilmager Fluoro-S (BioRad). La intensidad de las bandas que corresponden al tamaño del ADN sin escindir (78 pb) y los productos de digestión (55 y 23 pares de 50 bases) se midió usando Quantity One (BioRad) y se calcularon las proporciones relativas del ADN obtenido en un molde de híbridos de ADN-ARN después de cada ronda (véase la Figura 14). Después de cinco rondas de selección con la fusión de GST y el dedo de cinc ZfQQR, las proporciones de ADN, que portan el sitio de unión para ZfQQR y sin tal secuencia, han cambiado a 1:1,7. Esto significa que mediante la realización del procedimiento de SELEX, la combinación de las secuencias de control se enriquecía 8500 veces con una secuencia unida de forma específica 55 por ZfQQR.

## **Ejemplo 12 en los híbridos de ARN-ADN de la biblioteca SELEX**

60 **[0051]** En la siguiente etapa, la biblioteca de híbridos de ADN-ARN que contiene una secuencia aleatoria central (generada en el molde 9N - SEC ID N°: 28) se usó para realizar el procedimiento de SELEX y determinar la preferencia de la secuencia del dedo de cinc ZfQQR tal como en el Ejemplo 11. Para determinar la secuencia de los híbridos seleccionados, se digirieron 500 ng de producto de PCR de ADN obtenido en el molde del eluato después de las cinco rondas de SELEX con 5 unidades de la enzima XbaI (Fermentas) en tampón proporcionado por el fabricante. Los productos de digestión se separaron en gel de poliacrilamida nativo al 15 % en el tampón de TBE. La 65 banda correspondiente al fragmento con una longitud de 43 pares de bases se escindió y se aisló del gel de acuerdo

con procedimientos convencionales. Para obtener concatémicos, 100 ng de fragmentos de ADN digerido se ligaron con T4 ADN ligasa de bacteriófago (Fermentas) en tampón proporcionado por el fabricante a temperatura ambiente durante 3 horas. La ligasa se inactivó con calor incubando a 75 °C durante 10 min. Se digirieron 200 ng del vector pUC18 de plásmido de ADN (Fermentas) con la enzima XbaI en el tampón proporcionado por el fabricante durante 3 h a 37 °C. La mezcla de reacción se separó en gel de agarosa al 0,7 % en TAE y los fragmentos correspondientes a los tamaños esperados se volvieron a aislar a partir del gel usando un kit para reisolamiento de geles de agarosa (Gel out, A & A Biotechnology). El vector y los concatémicos preparados de este modo se ligaron y a continuación se transformaron en una cepa bacteriana. La selección de los transformantes adecuados que contenían los recombinantes deseados se basó en un análisis del color de la colonia y secuenciaron 12 clones y se encontró que contenían un total de 42 fragmentos obtenidos con el procedimiento de SELEX como se indica en la Tabla 2.

**Tabla 2.** Las secuencias obtenidas en el procedimiento de SELEX adquiridas después de la secuenciación del ADN obtenido en un molde de híbrido de ADN-ARN al final de cinco rondas del procedimiento de SELEX modificado realizado en el dedo de cinc ZfQQR y una biblioteca de secuencias que contiene 9 regiones degeneradas de nucleótidos.

Número	Secuencia de la biblioteca	Número	Secuencia de la biblioteca
1	TGGGGACGC	22	GTCAGATGT
2	AGTGCTCGA	23	GAGATCAGT
3	GACGCATGG	24	GCCGAGCGG
4	GGTCTGGAG	25	GCTCGGTGA
5	GAGCGGGAA	26	CGTAGGGAA
6	TCGTTGCAT	27	GTCGGATGG
7	CTAGCGGGT	28	CAGCGGGGT
8	GCCAGCGAG	29	GGGGGGAGG
9	AGAGGGGCA	30	TCGGCCAAG
10	TGGCAGGTT	31	CATGTGTTG
11	GGGTGGGAG	32	GACTGTAGA
12	GGTCCGGGC	33	CTGCCTGAA
13	CTCCGATCA	34	CGGCGGAGA
14	GCAAATGAA	35	CGTACAAGG
15	GGGCGGCGT	36	GGTGGTTGA
16	ACTCTGAGA	37	ATCAGGGGG
17	TGGGAATGG	38	TTGGCGGGG
18	AGCGGAGGG	39	GGGGTCCGA
19	GGATGGAAA	40	GAGAGAGCG
20	GGGCTGTCA	41	TGGCAGCTT
21	ACTCCTGAG	42	CGAGATGGA

[0052] Sobre esta base, la secuencia consenso de unión a ZFQQR, 5'-GGNCGNGGG-3' (Figura 15), se obtuvo usando WebLogo (Crooks GE, Hon G, Chandonia JM, Brenner SE WebLogo: A sequence logo generator, Genome Research, 14: 1188-1190, (2004), Schneider TD, Stephens RM. 1990. Sequence Logos: A New Way to Display Consensus Sequences. Nucleic Acids Res. 18: 6097-6100).

### **Ejemplo 13 Unión del dedo de cinc ZFQQR a la secuencia consenso de SELEX**

[0053] Para confirmar que la secuencia consenso, obtenida después de cinco rondas de SELEX, se une mediante el dedo de cinc, se midió una constante  $K_D$  usando el método de unión a filtro de nitrocelulosa. El ensayo se realizó con sustratos marcados con radiactividad que se describen en la bibliografía (LDNA, LRNA) y con una secuencia consenso (CDNA, CRNA), véase la Figura 16:

**LDNA:** TCACTGGGGAAGAAGAATCCTC (SEC ID N°: 29)

**LRNA:** GAGGAUUCUUCUUCUCCCCAGUGA (SEC ID N°: 30)  
**CDNA:** TCACTGGTCGGTGGGAATCCTC (SEC ID N°: 31)  
**CRNA:** GAGGAUCCCCACCGACCAGUGA (SEC ID N°: 32)

5 **[0054]** La reacción de unión contenía 2 nM de sustrato marcado, 2 µg de poli dI-dC como un competidor no específico, Tris 25 mM a pH 8,0, KCl 100 mM, ZnSO<sub>4</sub> 20 µM, DTT 2 mM. La reacción de unión se incubó durante 30 min a temperatura ambiente y a continuación se filtró inmediatamente a través de un filtro de nitrocelulosa y se lavó con 8 volúmenes de tampón de reacción de Tris 25 mM a pH 8,0, KCl 100 mM, ZnSO<sub>4</sub> 20 µM, DTT 2 mM. La medida de la intensidad de la señal que proviene de la retención de sustrato marcado con radiactividad se realizó en un audiorradiograma del filtro de nitrocelulosa usando un escáner Typhoon Trio y el software ImageQuant TL. Las constantes de unión  $K_D$  se midieron y para ambos sustratos eran similares,  $188 \pm 38$  nM para la unión del sustrato que se describe en la bibliografía y  $155 \pm 32$  nM para la secuencia consenso de SELEX (Figura 17).

**El listado de secuencias identificadas en la descripción:**

15

**[0055]**

**SEC ID N°: 1** - secuencia de ADN del gen *mhA* de ribonucleasa HI (ID del Gen 893801, BH0863) de *B. halodurans* (longitud completa);

20 **SEC ID N°: 2** - secuencia de nucleótidos del dedo de cinc ZfQQR;

**SEC ID N°: 3** - secuencia de nucleótidos de un gen que codifica la fusión enzimática que contiene un fragmento de la ribonucleasa HI de *B. halodurans* con las sustituciones K81A, K89E y K123A con el dedo de cinc ZfQQR - denominada **catAEA-ZfQQR**;

25 **SEC ID N°: 4** - secuencia de fusión de aminoácidos que contiene un fragmento de ribonucleasa HI de *B. halodurans* con las sustituciones K81A, K89E, K123A y con el dedo de cinc ZfQQR- denominada **catAEA-ZfQQR**;

**SEC ID N°: 5**- secuencia de nucleótidos del gen que codifica la variante **GQ** de catAEA-ZfQQR (la fusión enzimática que contiene un fragmento de ribonucleasa HI de *B. halodurans* con sustituciones que introducen cambios en la secuencia de aminoácidos de proteína en las posiciones K81A, K89E y K123A con el conector interdominios de dedo de cinc ZfQQR acortado con un fragmento que codifica los aminoácidos en las posiciones 138-148) - denominada GQ;

30 **SEC ID N°: 6** - secuencia de aminoácidos de la variante GQ de catAEA-ZfQQR, la fusión enzimática - denominada GQ;

**SEC ID N°: 7** - secuencia de nucleótidos de la variante **GGKKQ** de catAEA-ZfQQR (fusión enzimática que contiene un fragmento de ribonucleasa HI de *B. halodurans* con sustituciones que introducen cambios en la secuencia de aminoácidos de proteína en las posiciones K81A, K89E y K123A con el conector interdominios con un fragmento que codifica los aminoácidos en las posiciones 138-139 y 141-146)- denominada GGKKQ;

35 **SEC ID N°: 8** - secuencia de aminoácidos de la variante GGKKQ de la función enzimática - denominada GGKKQ;

**SEC ID N°: 9** - hebra de ARN del sustrato que contiene la secuencia de unión para ZfQQR;

**SEC ID N°: 10** - hebra de ADN del sustrato que contiene la secuencia de unión para ZfQQR;

**SEC ID N°: 11** - secuencia de nucleótidos del cebador Bhcatf;

**SEC ID N°: 12** - secuencia de nucleótidos del cebador Bhcatr;

**SEC ID N°: 13** - secuencia de nucleótidos del cebador BhZf;

45 **SEC ID N°: 14** - secuencia de nucleótidos del cebador Kmf;

**SEC ID N°: 15** - secuencia de nucleótidos del cebador Kmr;

**SEC ID N°: 16** - secuencia de nucleótidos del cebador K81 Ar;

**SEC ID N°: 17** - secuencia de nucleótidos del cebador K81 Af;

**SEC ID N°: 18** - secuencia de nucleótidos del cebador K89Er;

50 **SEC ID N°: 19** - secuencia de nucleótidos del cebador K89Ef;

**SEC ID N°: 20** - secuencia de nucleótidos del cebador K123Ar;

**SEC ID N°: 21** - secuencia de nucleótidos del cebador K123Af;

**SEC ID N°: 22** - secuencia de nucleótidos del cebador del11f;

**SEC ID N°: 23** - secuencia de nucleótidos del cebador del11r;

55 **SEC ID N°: 24** - secuencia de nucleótidos del cebador del5f;

**SEC ID N°: 25** - secuencia de nucleótidos del cebador del5r;

**SEC ID N°: 26** - secuencia de la hebra de ADN del sustrato usado para control del método de SELEX que contiene el sitio de unión para ZfQQR (Molde A);

**SEC ID N°: 27** - secuencia de la hebra de ADN del sustrato usado para control del método SELEX que no tienen un sitio de unión para ZfQQR, en su lugar tiene un sitio de restricción XhoI (Molde B);

60 **SEC ID N°: 28** - secuencia de la hebra de ADN del sustrato usado para la construcción de la biblioteca de híbridos de ADN-ARN que contiene una secuencia aleatoria nonanucleótido central (Molde 9N);

**SEC ID N°: 29** - secuencia de nucleótidos de la hebra de ADN del híbrido de ADN-ARN que contiene la secuencia de unión a ZfQQR usada para la determinación de  $K_D$ ;

65 **SEC ID N°: 30** - secuencia de nucleótidos de la hebra de ARN del híbrido de ADN-ARN que contiene la

secuencia de unión de ZfQQR usada para la determinación de K<sub>D</sub>;

**SEC ID N°: 31** - secuencia de nucleótidos de la hebra de ADN del híbrido de ADN-ARN que contiene la secuencia consenso deducida a partir de SELEX usada para la determinación de K<sub>D</sub>;

5 **SEC ID N°: 32** - secuencia de nucleótidos de la hebra de ARN del híbrido de ADN-ARN que contiene la secuencia consenso deducida a partir de SELEX usada para la determinación de K<sub>D</sub>;

**SEC ID N°: 33** - cebador 1 - cebador usado para la amplificación del ADN de los moldes A, B y 9N en la reacción de PCR;

**SEC ID N°: 34** - cebador 2 - cebador usado para la amplificación del ADN de los moldes A, B y 9N en la reacción de PCR;

10 **SEC ID N°: 35** - cebador ZFf - cebador usado en PCR para la preparación de la construcción de ADN del gen de fusión que codifica el dominio GST y el gen que codifica el dedo de cinc Zf-QQ;

**SEC ID N°: 36** - cebador ZFr - cebador usado para la preparación de la construcción de ADN del gen de fusión que codifica el dominio GST y el gen que codifica el dedo de cinc Zf-Q;

15 **SEC ID N°: 37** - es la secuencia de nucleótidos del gen que codifica la fusión del dominio GST y el gen que codifica el dedo de cinc ZfQQR (GST-ZfQQR de fusión).

LISTADO DE SECUENCIAS

**[0056]**

20

<110> IBMIK

<120> Ribonucleasa H modificada por ingeniería específica de secuencia y el método para determinar la preferencia de secuencia de proteínas de unión de híbrido de ADN-ARN

25

<130> PZ/1574/AGR/PCT

<160> 37

30

<170> PatentIn versión 3.3

<210> 1

<211> 688

<212> ADN

35

<213> *Bacillus halodurans*

<400> 1

```

atggcaaaagt caaaatacta tgtcgtatgg aatggacgga agcccggcat ttatacgagc      60
tggctctgcat gtgaagctca agtaaaagga tataccggcg ccaaatttaa atcctatcct      120
tccaaggaag aagcagaggc tgcttttaga ggagaagagg caacaccgaa gcttgcaaaa      180
gaggagatta tttgggagag cctgtctgta gatgttggca gccaaaggaa tcccgggaatt      240
gtggaatata aaggcgttga tacgaaaacg ggagaagtcc tttttgaacg agagccgatt      300
ccgatcggga caaacaatat gggtagtatt ctgcgatcgc ttcacgggct tcgttacctt      360
aaggaacgga acagtcgtaa gccgatctat tctgattccc agacggcaat caaatgggtg      420
aaggataaaa aagcaaaatc gaccctcgtg cgcaatgaag aaacagcgct tatttggaag      480
cttgtagatg aagcggagga gtggctaaac actcatacct atgaaacgcc catcttaaaa      540
tggcagaccg ataagtgggg ggaaattaag gccgattacg ggagaaaagg taccgacgac      600
gacgacaagg ccatggcgat atcggatccg aattcgagct ccgtcgacaa gcttgcggcc      660
gcactcggagc accaccacca ccaccact
    
```

40

<210> 2

<211> 305

<212> ADN

<213> artificial

45

<220>

<223> Secuencia de nucleótidos ZfQQR

50

<400> 2

ES 2 557 633 T3

ccatggaaaa	actgcgtaac	ggttctggtg	acccgggtaa	aaaaaacag	cacgcgtgcc	60
cggaatgcgg	taaactcttc	tctcagtctt	ctaacctgca	gaaacaccag	cgtaccacaca	120
cgggtgaaaa	accgtacaaa	tgtccagaat	gtggcaaaag	ctttagtcaa	agttctaatac	180
ttcaaaaaaca	tcaacgcacg	cataccggcg	agaagccata	taagtgtccg	gagtgcggca	240
aaagcttctc	ccgctctgat	cacctccagc	gtcatcagcg	cactcatcag	aacaaaaaac	300
tcgag						305

<210> 3  
 <211> 720  
 <212> ADN  
 <213> artificial

5

<220>  
 <223> secuencia de nucleótidos - fusión catAEA-ZfQQR

10

<400> 3

atggcaaaaag	aggagattat	ttgggagagc	ctgtctgtag	atggttggcag	ccaagggaat	60
cccggaaattg	tggaatataa	aggcgttgat	acgaaaacgg	gagaagtcct	ttttgaacga	120
gagccgattc	cgatcgggac	aaacaatatg	ggtgagtttc	tcgcgatcgt	tcacgggctt	180

cgttacctta	aggaacggaa	cagtcgtaag	ccgatctatt	ctgattccca	gacggcaatc	240
gcatgggtga	aggataaaaa	agcagaatcg	accctcgtgc	gcaatgaaga	aacagcgctt	300
at ttggaagc	ttgtagatga	agcggaggag	tggtctaaaca	ctcataccta	tgaaacgccc	360
atcttagcat	ggcagaccga	taagtggggg	gaaattaagg	ccgattacgg	gagaaaaggt	420
tctggtgacc	cgggtaaaaa	aaaacagcac	gcgtgcccgg	aatgcggtaa	atctttctct	480
cagtcttcta	acctgcagaa	acaccagcgt	acccacaccg	gtgaaaaacc	gtacaaatgt	540
ccagaatgtg	gcaaaagctt	tagtcaaagt	tctaactctc	aaaaacatca	acgcacgcat	600
accggcgaga	agccatataa	gtgtccggag	tgcggcaaaa	gcttctcccg	ctctgatcac	660
ctccagcgtc	atcagcgcac	tcatcagaac	aaaaaactcg	agcaccacca	ccaccaccac	720

15

<210> 4  
 <211> 240  
 <212> PRT  
 <213> artificial

20

<220>  
 <223> secuencia de fusión de aminoácidos de catAEA-ZfQQR

25

<400> 4

ES 2 557 633 T3

Met Ala Lys Glu Glu Ile Ile Trp Glu Ser Leu Ser Val Asp Val Gly  
 1 5 10 15  
 Ser Gln Gly Asn Pro Gly Ile Val Glu Tyr Lys Gly Val Asp Thr Lys  
 20 25 30  
 Thr Gly Glu Val Leu Phe Glu Arg Glu Pro Ile Pro Ile Gly Thr Asn  
 35 40 45  
 Asn Met Gly Glu Phe Leu Ala Ile Val His Gly Leu Arg Tyr Leu Lys  
 50 55 60  
 Glu Arg Asn Ser Arg Lys Pro Ile Tyr Ser Asp Ser Gln Thr Ala Ile  
 65 70 75 80  
 Ala Trp Val Lys Asp Lys Lys Ala Glu Ser Thr Leu Val Arg Asn Glu  
 85 90 95  
 Glu Thr Ala Leu Ile Trp Lys Leu Val Asp Glu Ala Glu Glu Trp Leu  
 100 105 110  
 Asn Thr His Thr Tyr Glu Thr Pro Ile Leu Ala Trp Gln Thr Asp Lys  
 115 120 125  
 Trp Gly Glu Ile Lys Ala Asp Tyr Gly Arg Lys Gly Ser Gly Asp Pro  
 130 135 140  
 Gly Lys Lys Lys Gln His Ala Cys Pro Glu Cys Gly Lys Ser Phe Ser  
 145 150 155 160  
 Gln Ser Ser Asn Leu Gln Lys His Gln Arg Thr His Thr Gly Glu Lys  
 165 170 175  
 Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Ser Asn  
 180 185 190  
 Leu Gln Lys His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys  
 195 200 205  
 Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp His Leu Gln Arg His  
 210 215 220  
 Gln Arg Thr His Gln Asn Lys Lys Leu Glu His His His His His His  
 225 230 235 240

<210> 5  
 <211> 687  
 <212> ADN  
 <213> artificial

5

<220>  
 <223> secuencia de nucleótidos - fusión de la variante GQ de catAEA-ZfQQR

10

<400> 5

ES 2 557 633 T3

```

atggcaaaag aggagattat ttgggagagc ctgtctgtag atgttggcag ccaagggaaat      60
cccggaattg tggaatataa aggcggtgat acgaaaacgg gagaagtcct ttttgaacga      120
gagccgattc cgatcgggac aaacaatatg ggtgagtttc tcgcgatcgt tcacgggctt      180
cgttacctta aggaacggaa cagtcgtaag ccgatctatt ctgattccca gacggcaatc      240
gcatgggtga aggataaaaa agcagcatcg accctcgtgc gcaatgaaga aacagcgctt      300
atltggaagc ttgtagatga agcggaggag tggctaaaca ctcatacctt tgaacgccc      360
atcttagcat ggcagaccga taagtggggg gaaattaagg ccgattacgg gcagcacgcg      420
tgcccggaat gcggtaaatc tttctctcag tcttctaacc tgcagaaaca ccagcgtacc      480
cacaccggtg aaaaaccgta caaatgtcca gaatgtggca aaagctttag tcaaagttct      540
aatcttcaaa aacatcaacg cagcatatcc ggcgagaagc catataagtg tccggagtgc      600
ggcaaaagct tctcccgtc tgatcacctc cagcgtcacc agcgcactca tcagaacaaa      660
aaactcgagc accaccacca ccaccac      687

```

<210> 6

<211> 229

5 <212> PRT

<213> artificial

<220>

10 <223> secuencia de aminoácidos - fusión de la variante GQ de catAEA-ZfQQR

<400> 6

```

Met Ala Lys Glu Glu Ile Ile Trp Glu Ser Leu Ser Val Asp Val Gly
1          5          10          15
Ser Gln Gly Asn Pro Gly Ile Val Glu Tyr Lys Gly Val Asp Thr Lys
          20          25          30
Thr Gly Glu Val Leu Phe Glu Arg Glu Pro Ile Pro Ile Gly Thr Asn
          35          40          45
Asn Met Gly Glu Phe Leu Ala Ile Val His Gly Leu Arg Tyr Leu Lys
          50          55          60
Glu Arg Asn Ser Arg Lys Pro Ile Tyr Ser Asp Ser Gln Thr Ala Ile
65          70          75          80
Ala Trp Val Lys Asp Lys Lys Ala Ala Ser Thr Leu Val Arg Asn Glu
          85          90          95
Glu Thr Ala Leu Ile Trp Lys Leu Val Asp Glu Ala Glu Glu Trp Leu
          100          105          110
Asn Thr His Thr Tyr Glu Thr Pro Ile Leu Ala Trp Gln Thr Asp Lys
          115          120          125
Trp Gly Glu Ile Lys Ala Asp Tyr Gly Gln His Ala Cys Pro Glu Cys
          130          135          140

```

ES 2 557 633 T3

Gly Lys Ser Phe Ser Gln Ser Ser Asn Leu Gln Lys His Gln Arg Thr  
 145 150 155 160  
 His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe  
 165 170 175  
 Ser Gln Ser Ser Asn Leu Gln Lys His Gln Arg Thr His Thr Gly Glu  
 180 185 190  
 Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp  
 195 200 205  
 His Leu Gln Arg His Gln Arg Thr His Gln Asn Lys Lys Leu Glu His  
 210 215 220  
 His His His His His  
 225

5 <210> 7  
 <211> 696  
 <212> ADN  
 <213> artificial

10 <220>  
 <223> secuencia de nucleótidos - fusión de la variante GGKKQ de catAEA-ZfQQR  
 <400> 7

```

atggcaaaag aggagattat ttgggagagc ctgtctgtag atgttggcag ccaagggaaat 60
cccggaattg tggaatataa aggcgttgat acgaaaacgg gagaagtect ttttgaacga 120
gagccgattc cgatcgggac aaacaatatg ggtgagtttc tcgcatcgt tcacgggctt 180
cgttacctta aggaacggaa cagtcgtaag ccgatctatt ctgattccca gacggcaatc 240
gcatgggtga aggataaaaa agcagcatcg accctcgtgc gcaatgaaga aacagcgctt 300
atgttgaagc ttgtagatga agcggaggag tggctaaaca ctcatacctt tgaaacgccc 360
atcttagcat ggcagaccga taagtggggg gaaattaagg ccgattacgg gggcaaaaaa 420
cagcagcgtg gcccggaatg cggtaaactt ttctctcagt cttctaacct gcagaaacac 480
cagcgtaccc acaccggtga aaaaccgtac aaatgtccag aatgtggcaa aagctttagt 540
caaagtctta atcttcaaaa acatcaacgc acgcataacc gcgagaagcc atataagtgt 600
ccggagtgcg gcaaaagctt ctcccgtctt gatcacctcc agcgtcatca gcgcactcat 660
cagaacaaaa aactcgagca ccaccaccac caccac 696
  
```

15 <210> 8  
 <211> 232  
 <212> PRT  
 <213> Artificial

20 <220>  
 <223> secuencia de aminoácidos - fusión de la variante GGKKQ de catAEA-ZfQQR  
 <400> 8

ES 2 557 633 T3

Met Ala Lys Glu Glu Ile Ile Trp Glu Ser Leu Ser Val Asp Val Gly  
 1 5 10 15  
 Ser Gln Gly Asn Pro Gly Ile Val Glu Tyr Lys Gly Val Asp Thr Lys  
 20 25 30  
 Thr Gly Glu Val Leu Phe Glu Arg Glu Pro Ile Pro Ile Gly Thr Asn  
 35 40 45  
 Asn Met Gly Glu Phe Leu Ala Ile Val His Gly Leu Arg Tyr Leu Lys  
 50 55 60  
 Glu Arg Asn Ser Arg Lys Pro Ile Tyr Ser Asp Ser Gln Thr Ala Ile  
 65 70 75 80  
 Ala Trp Val Lys Asp Lys Lys Ala Ala Ser Thr Leu Val Arg Asn Glu  
 85 90 95  
 Glu Thr Ala Leu Ile Trp Lys Leu Val Asp Glu Ala Glu Glu Trp Leu  
 100 105 110  
 Asn Thr His Thr Tyr Glu Thr Pro Ile Leu Ala Trp Gln Thr Asp Lys  
 115 120 125  
 Trp Gly Glu Ile Lys Ala Asp Tyr Gly Gly Lys Lys Gln His Ala Cys  
 130 135 140  
 Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Ser Asn Leu Gln Lys His  
 145 150 155 160  
 Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly  
 165 170 175  
 Lys Ser Phe Ser Gln Ser Ser Asn Leu Gln Lys His Gln Arg Thr His  
 180 185 190  
 Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser  
 195 200 205  
 Arg Ser Asp His Leu Gln Arg His Gln Arg Thr His Gln Asn Lys Lys  
 210 215 220  
 Leu Glu His His His His His His  
 225 230

<210> 9  
 <211> 75  
 5 <212> ARN  
 <213> Artificial

<220>  
 <223> hebra de ARN que contiene la secuencia de unión para ZfQQR

10 <400> 9

agaacuagug gaucaaccgg gcugcaggaa uucgauauca agcuuauca uaccguggcg 60  
 guucuucccc aagcc 75

ES 2 557 633 T3

<210> 10  
 <211> 74  
 <212> ADN  
 5 <213> artificial  
  
 <220>  
 <223> hebra de ADN del sustrato que contiene la secuencia de unión para ZfQQR  
 10 <400> 10  
  
 gcttggggaa gaaccgccac ggtatcgata agcttgatat cgaattcctg cagcccggtt 60  
  
 gatccactag ttct 74  
  
 <210> 11  
 15 <211> 32  
 <212> ADN  
 <213> artificial  
  
 <220>  
 20 <223> cebador  
  
 <400> 11  
 gacgcatatg gcaaaagagg agattattg gg 32  
  
 25 <210> 12  
 <211> 26  
 <212> ADN  
 <213> artificial  
  
 30 <220>  
 <223> cebador  
  
 <400> 12  
 gtgtacctt ttctccgta atcggc 26  
 35  
 <210> 13  
 <211> 17  
 <212> ADN  
 <213> artificial  
 40  
 <220>  
 <223> cebador  
  
 <400> 13  
 45 ggttctggtg acccggg 17  
  
 <210> 14  
 <211> 21  
 <212> ADN  
 50 <213> artificial  
  
 <220>  
 <223> cebador  
  
 <400> 14  
 55 gggatcgag tggtagtaa c 21  
  
 <210> 15  
 <211> 26  
 60 <212> ADN  
 <213> artificial

ES 2 557 633 T3

<220>  
<223> cebador

<400> 15  
5 cgggaaaaca gcattccagg tattag 26

<210> 16  
<211> 23  
<212> ADN  
10 <213> artificial

<220>  
<223> cebador

<400> 16  
15 ccgtctggga atcagaatag atc 23

<210> 17  
<211> 27  
20 <212> ADN  
<213> artificial

<220>  
<223> cebador

<400> 17  
25 caatcgcgatg ggtgaaggat aaaaaag 27

<210> 18  
30 <211> 21  
<212> ADN  
<213> artificial

<220>  
35 <223> cebador

<400> 18  
ctgcttttt atccttcacc c 21

<210> 19  
40 <211> 27  
<212> ADN  
<213> artificial

<220>  
45 <223> cebador

<400> 19  
50 caatcgcgatg ggtgaaggat aaaaaag 27

<210> 20  
<211> 24  
<212> ADN  
55 <213> artificial

<220>  
<223> cebador

<400> 20  
60 taagatgggc gttcatagg tatg 24

<210> 21  
<211> 20  
<212> ADN  
65 <213> artificial

ES 2 557 633 T3

<220>  
 <223> cebador  
 5 <400> 21  
 gcatggcaga ccgataagtg 20  
 <210> 22  
 <211> 25  
 10 <212> ADN  
 <213> artificial  
 <220>  
 <223> cebador  
 15 <400> 22  
 ggctccggcc agcacgctg cccgg 25  
 <210> 23  
 <211> 31  
 20 <212> ADN  
 <213> artificial  
 <220>  
 <223> cebador  
 25 <400> 23  
 agaaccgctc ccgtaatcg ccttaattc c 31  
 30 <210> 24  
 <211> 21  
 <212> ADN  
 <213> artificial  
 35 <220>  
 <223> cebador  
 <400> 24  
 caaaaaacag cacgctgcc c 21  
 40 <210> 25  
 <211> 25  
 <212> ADN  
 <213> artificial  
 45 <220>  
 <223> cebador  
 <400> 25  
 50 cccccgtaat cggcctaat ttcc 25  
 <210> 26  
 <211> 78  
 <212> ADN  
 55 <213> artificial  
 <220>  
 <223> contiene el sitio de unión para ZfQQR  
 60 <400> 26  
 gagatctaga cggaacatga aggggaagaa ttctatgctt agtgagatct agagccctat 60  
 agtgagtcgt attaaatt 78

ES 2 557 633 T3

<210> 27  
 <211> 78  
 <212> ADN  
 5 <213> artificial  
  
 <220>  
 <223> no contiene el sitio de unión para ZfQQR  
 10 <400> 27  
  
 gagatctaga cggaacatga agctcgagcc ttctatgctt agtgagatct agagccctat 60  
  
 agtgagtcgt attaaatt 78  
  
 <210> 28  
 <211> 78  
 <212> ADN  
 <213> artificial  
 15  
  
 <220>  
 20 <223> Molde con secuencia aleatoria nonanucleótido  
  
 <220>  
 <221> misc\_feature  
 <222> (22)..(30)  
 25 <223> n e s a, c, g o t  
  
 <400> 28  
  
 gagatctaga cggaacatgt annnnnnnnn tactatgctt agtgagatct agagccctat 60  
  
 agtgagtcgt attaaatt 78  
 30  
 <210> 29  
 <211> 22  
 <212> ADN  
 <213> artificial  
 35  
 <220>  
 <223> hebra de ADN del híbrido de ADN-ARN  
  
 <400> 29  
 40 tctctgggga agaagaatcc tc 22  
  
 <210> 30  
 <211> 22  
 <212> ARN  
 45 <213> artificial  
  
 <220>  
 <223> hebra de ARN del híbrido ADN-ARN  
 50  
 <400> 30  
 gaggauucuu cuucccagu ga 22  
  
 <210> 31  
 <211> 22  
 55 <212> ADN  
 <213> artificial  
  
 <220>  
 <223> hebra de ADN del híbrido ADN-ARN  
 60

# ES 2 557 633 T3

<400> 31  
tcactggtcg gtgggaatcc tc 22

5 <210> 32  
<211> 22  
<212> ARN  
<213> artificial

10 <220>  
<223> hebra de ARN del híbrido ADN-ARN

<400> 32  
gaggauuccc accgaccagu ga 22

15 <210> 33  
<211> 46  
<212> ADN  
<213> artificial

20 <220>  
<223> iniciador

<400> 33  
aatttaatac gactcactat agggctctag atctcactaa gcatag 46

25 <210> 34  
<211> 19  
<212> ADN  
<213> artificial

30 <220>  
<223> iniciador

<400> 34  
gagatctaga cggaacatg 19

35 <210> 35  
<211> 17  
<212> ADN  
<213> artificial

40 <220>  
<223> iniciador

45 <400> 35  
ggttctggtg acccggg 17

50 <210> 36  
<211> 26  
<212> ADN  
<213> artificial

55 <220>  
<223> iniciador

<400> 36  
cgggaaaaca gcattccagg tattag 26

60 <210> 37  
<211> 975  
<212> ADN  
<213> artificial

65 <220>  
<223> fusión del dominio GST y dedo de cinc ZfQQR

ES 2 557 633 T3

<400> 37

atgtccccta	tactaggtta	ttggaaaatt	aagggccttg	tgcaaccac	tcgacttctt	60
ttggaatata	ttgaagaaaa	atatgaagag	catttgtatg	agcgcgatga	aggtgataaa	120
tggcgaaaca	aaaagtttga	attgggtttg	gagtttccca	atcttcctta	ttatattgat	180
ggtgatgtta	aattaacaca	gtctatggcc	atcatacgtt	atatagctga	caagcacaac	240
atggtgggtg	gttgtccaaa	agagcgtgca	gagatttcaa	tgcttgaagg	agcggttttg	300
gatattagat	acggtgtttc	gagaattgca	tatagtaaag	actttgaaac	tctcaaagtt	360
gattttctta	gcaagctacc	tgaaatgctg	aaaatgttcg	aagatcgttt	atgtcataaa	420
acataatata	atggtgatca	tgtaacccat	cctgacttca	tgttgtatga	cgctcttgat	480
gttgttttat	acatggacc	aatgtgcctg	gatgcgttcc	caaaattagt	ttgttttaaa	540
aaacgtattg	aagctatccc	acaaattgat	aagtacttga	aatccagcaa	gtatatagca	600
tggcctttgc	agggtggca	agccacgttt	ggtggtggcg	accatcctcc	aaaatcggat	660
ctggttccgc	gtggatcccc	ggaattcccg	ggtaaaaaaa	aacagcacgc	gtgcccggaa	720
tgcggtaaat	ctttctctca	gtcttctaac	ctgcagaaac	accagcgtac	ccacaccggt	780
gaaaaaccgt	acaaatgtcc	agaatgtggc	aaaagcttta	gtcaaagtcc	taatcttcaa	840
aaacatcaac	gcacgcatac	cggcgagaag	ccatataagt	gtccggagtg	cggcaaaagc	900
ttctcccgt	ctgatcacct	ccagcgtcat	cagcgcactc	atcagaacaa	aaaactcgag	960
cggccgcatac	gtgac					975

5

## REIVINDICACIONES

1. Una ribonucleasa, que escinde la hebra de ARN en híbridos de ADN-ARN, en la que dicha ribonucleasa es una proteína de fusión que comprende  
 5 un derivado del dominio catalítico de la ribonucleasa HI (RNasa HI),
- en la que el derivado del dominio catalítico de la RNasa HI es de *Bacillus halodurans*, que comprende el polipéptido codificado por los nucleótidos 175 a 588 del gen *mhA* mostrado en la SEC ID N°: 1, y  
 10 en la que el derivado del dominio catalítico de la RNasa HI comprende al menos una sustitución de un residuo de aminoácido en el dominio de unión al sustrato seleccionada entre: K81A, K89E y K123A en referencia a la secuencia del polipéptido codificada por la SEC ID N°: 1, preferentemente contiene todas las sustituciones K81A, K89E y K123A en referencia a la secuencia del polipéptido codificada por la SEC ID N°: 1, y  
 15 en la que el derivado del dominio catalítico de la RNasa HI no comprende dominio de unión a híbrido de RNasa HI,
- con un dominio de unión de dedo de zinc híbrido de ADN-ARN, en el que el dominio de unión de dedo de zinc tiene la capacidad de unirse a una secuencia específica en el híbrido de ADN-ARN.
2. La ribonucleasa de acuerdo con la reivindicación 1, en la que el dominio de dedo de zinc es un derivado del dedo de cinc ZfQQR, preferentemente un polipéptido codificado por los nucleótidos de 19 a 303 de la secuencia ZfQQR que se muestra en la SEC ID N°: 2.
3. La ribonucleasa de acuerdo con las reivindicaciones 1-2, que comprende una proteína de fusión catAEA-ZfQQR como se muestra en la SEC ID N°: 4, o una proteína de fusión GQ como se muestra en la SEC ID N°: 6, o  
 25 comprende una proteína de fusión GGKKQ como se muestra en la SEC ID N°: 8.
4. Una composición que comprende la ribonucleasa de acuerdo con las reivindicaciones 1-3.
5. Uso de una ribonucleasa de acuerdo con las reivindicaciones 1-3 o de una composición de acuerdo con la  
 30 reivindicación 4 para la escisión *in vitro* de la hebra de ARN en híbridos de ADN-ARN, preferentemente en el que la escisión de la hebra de ARN en híbridos de ADN-ARN se localiza a 2-16 nucleótidos, preferentemente a 5-7 nucleótidos, del sitio de unión para el dedo de zinc.
6. Un método para obtener una variante modificada por ingeniería de la RNasa HI que escinde la hebra de ARN en  
 35 híbridos de ADN-ARN que comprende las siguientes etapas:
- a) obtener un derivado del dominio catalítico de la RNasa HI, en el que el derivado del dominio catalítico de la RNasa HI es de *Bacillus halodurans*, que comprende el polipéptido codificado por los nucleótidos 175 a 588 del gen *mhA* que se muestra en la SEC ID N°: 1, y  
 40 en el que el derivado del dominio catalítico de la RNasa HI comprende al menos una sustitución de un residuo de aminoácido en el dominio de unión al sustrato seleccionada entre: K81A, K89E y K123A en referencia a la secuencia del polipéptido codificada por la SEC ID N°: 1, preferentemente contiene todas las sustituciones K81A, K89E y K123A en referencia a la secuencia del polipéptido codificada por la SEC ID N°: 1, y  
 45 en el que el derivado del dominio catalítico de la RNasa HI no comprende dominio de unión a híbrido de RNasa HI,
- b) obtener una RNasa HI modificada por ingeniería mediante la preparación de una proteína de fusión que comprende el derivado del dominio catalítico de la RNasa HI obtenido en la etapa a) con un dominio de unión de dedo de zinc que tiene la capacidad de unirse a secuencias específicas en el híbrido de ADN-ARN, preferentemente en el que el dominio de dedo de zinc es un derivado del dedo de zinc ZfQQR, preferentemente  
 50 un polipéptido codificado por los nucleótidos de 19 a 303 de la secuencia ZfQQR que se muestra en la SEC ID N°: 2.

**Fig. 1**

nombre del cebador	Secuencia	Número de SEC ID N°
Bhcatf	GACGCATATGGCAAAAGAGGAGATTATTTGGG	SEC ID N°: 11
Bhcatr	GTGGTACCTTTTCTCCGTAATCGGC	SEC ID N°: 12
BhZf	GGTCTGGTGACCCGGG	SEC ID N°: 13
Kmf	GGGATCGCAGTGGTGAGTAAC	SEC ID N°: 14
Kmr	CGGGAAAACAGCATTCCAGGTATTAG	SEC ID N°: 15
K81Ar	CCGTCTGGGAATCAGAATAGATC	SEC ID N°: 16
K81Af	CAATCGCATGGGTGAAGGATAAAAAAG	SEC ID N°: 17
K89Er	CTGCTTTTTTATCCTTCACCC	SEC ID N°: 18
K89Ef	CAATCGCATGGGTGAAGGATAAAAAAG	SEC ID N°: 19
K123Ar	TAAGATGGGCGTTTCATAGGTATG	SEC ID N°: 20
K123Af	GCATGGCAGACCGATAAGTG	SEC ID N°: 21
del11f	GGCTCCGGCCAGCACGCGTGCCCGG	SEC ID N°: 22
del11r	AGAACCGCTCCCGTAATCGGCCTTAATTTCC	SEC ID N°: 23
del5f	CAAAAAACAGCACGCGTGCCC	SEC ID N°: 24
del5r	CCCCCGTAATCGGCCTTAATTTCCC	SEC ID N°: 25

**Fig. 2****(A) ARN:**

AGAACUAGUGGAUCAACCGGGCUGCAGGAAUUCGAUAUCAAGCUUAUCGAUA  
CCGUGGCGGUUCUCCCCAAGCC (SEC ID N°: 9)

**(B) ADN:**

GCTTGGGGAAGAACCGCCACGGTATCGATAAGCTTGATATCGAATTCCTGCAG  
CCCGTTGATCCACTAGTTCT (SEC ID N°: 10)

Fig. 3

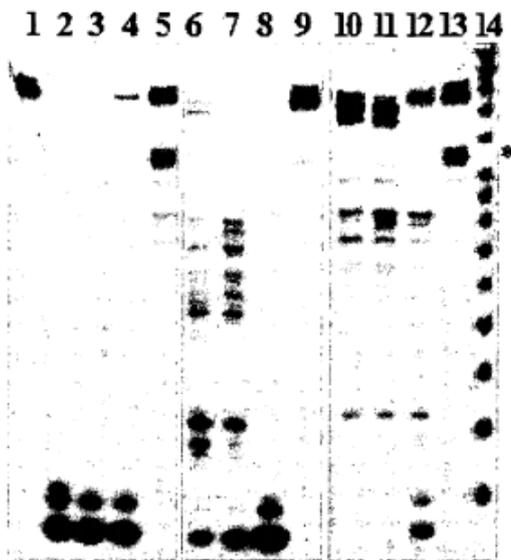


Fig. 4

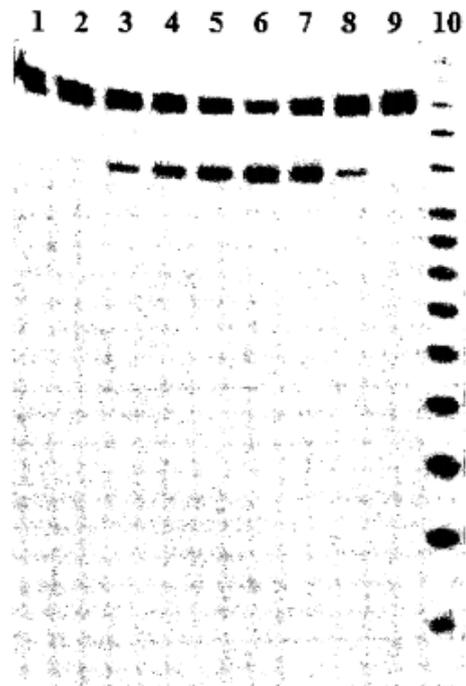


Fig. 5

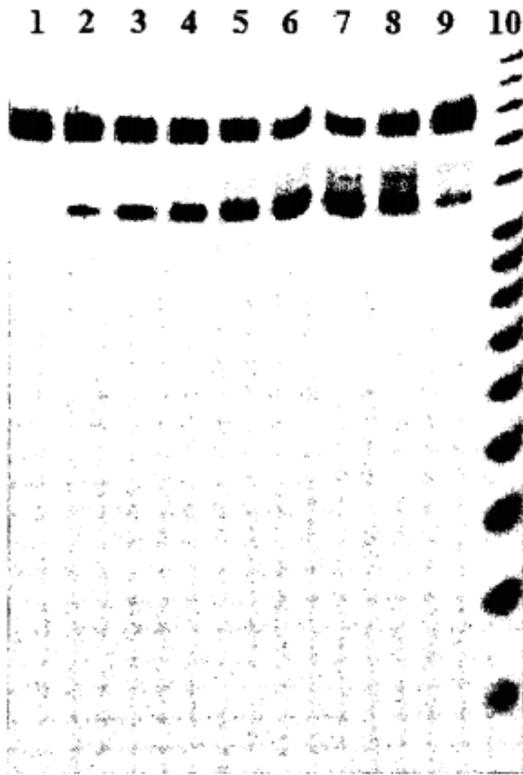
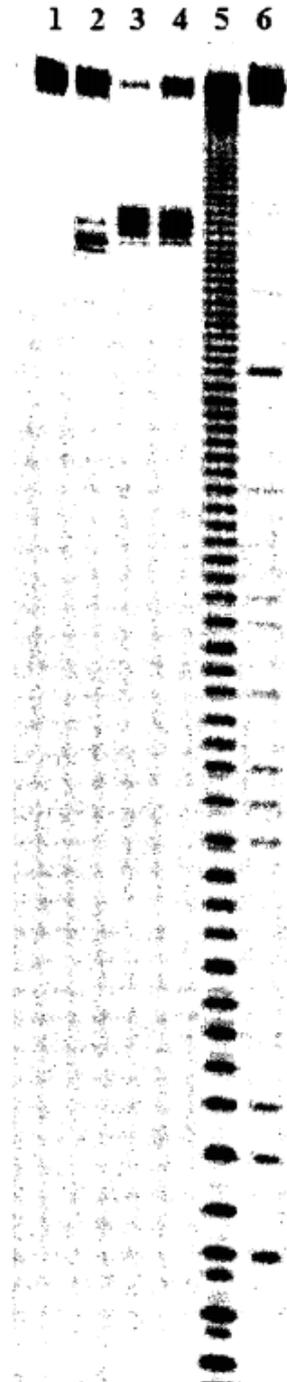


Fig. 6



**Fig.7**



**Fig. 8**



**Fig. 9**



**Fig. 10**



**Fig. 11**

**Molde A:**

GAGATCTAGACGGAACATGAAGGGGAAGAATTCTATGCTTAGTGAGATCTAGAGCC  
CTATAGTGAGTCGTATTAAT (SEC ID N°: 26)

**Molde B:**

GAGATCTAGACGGAACATGAAGCTCGAGCCTTCTATGCTTAGTGAGATCTAGAGCC  
CTATAGTGAGTCGTATTAAT (SEC ID N°: 27)

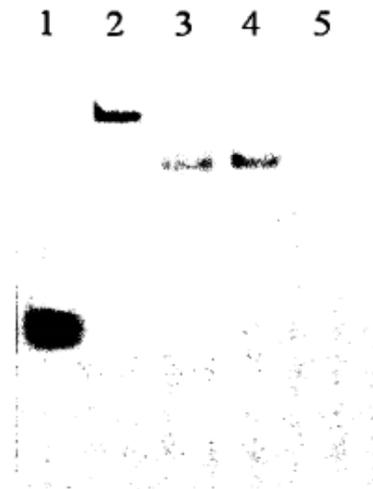
**Molde 9N:**

GAGATCTAGACGGAACATGTANNNNNNNNTACTATGCTTAGTGAGATCTAGAGCC  
CTATAGTGAGTCGTATTAAT (SEC ID N°: 28)

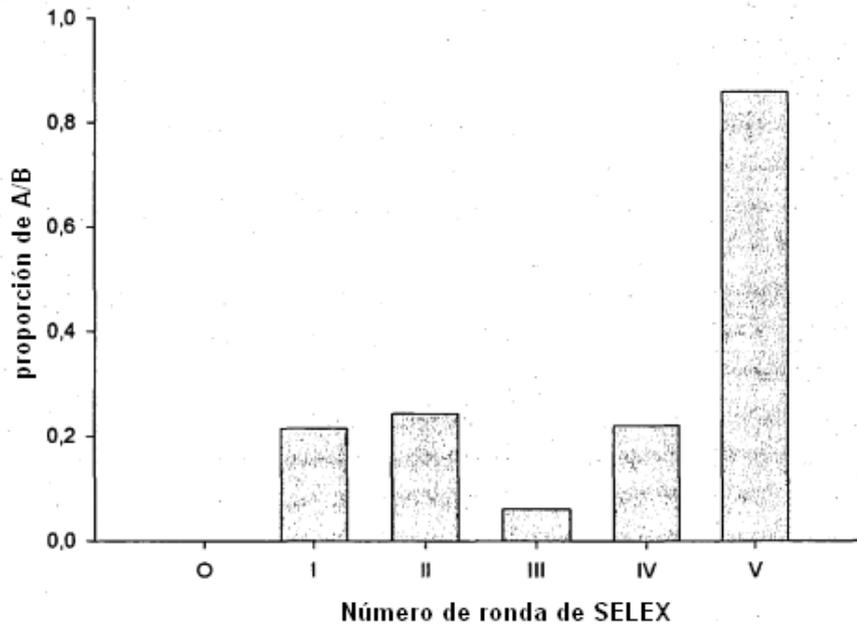
**Fig. 12**



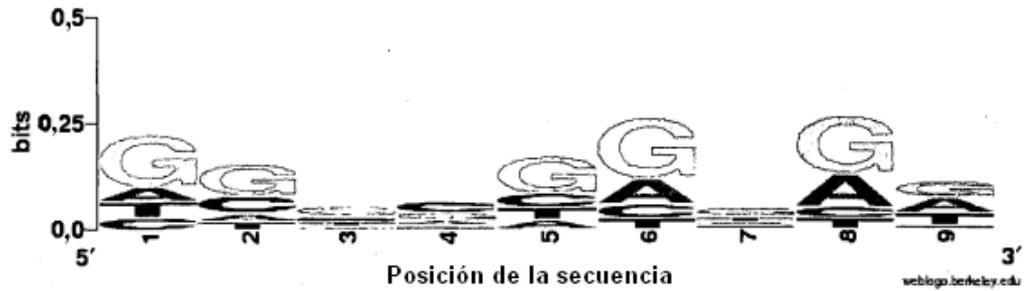
**Fig. 13**



**Fig. 14**



**Fig. 15**



**Fig. 16**

**A)**

**LDNA: TCACTGGGGAAGAAGAATCCTC (SEC ID N°: 29)**

**LRNA: GAGGAUUCUUCUCCCCAGUGA (SEC ID N°: 30)**

**B)**

**CDNA: TCACTGGTCGGTGGGAATCCTC (SEC ID N°: 31)**

**CRNA: GAGGAUUCCCACCGACCAGUGA (SEC ID N°: 32)**

## REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante es únicamente para la comodidad del lector. No forma parte del documento de la patente europea. A pesar del cuidado tenido en la recopilación de las referencias, no se pueden excluir errores u omisiones y la EPO niega toda responsabilidad en este sentido.

## Documentos de patentes citados en la descripción

10

- WO 2010076939 A1 [0004] [0006]
- WO 03087341 A2 [0004] [0006]
- WO 2009146179 A1 [0004] [0006]
- WO 2007014181 A2 [0004]
- WO 2007014182 A2 [0004]
- US 2006057590 A [0008]
- JP 54059392 B [0008]

15

## Literatura diferente de patentes citada en la descripción

20

- ELLINGTON, A.D. ; SZOSTAK, J.W. *Nature*, 1990, vol. 346, 818-822 [0005]
- HUIZENGA DE ; SZOSTAK JW. *Biochemistry*, 1995, vol. 34 (2), 656-65 [0005]
- BLACKWELL TK ; WEINTRAUB H. *Science*, 1990, vol. 250, 1104-1110 [0005]
- HERSKOVITZ M.A. et al. Endoribonuclease RNase III is essential in *Bacillus subtilis*. *Mol Microbiol.*, December 2000, vol. 38 (5), 1027-33 [0007]
- DASGUPTA S. et al. Genetic uncoupling of the dsRNA-binding and RNA cleavage activities of the *Escherichia coli* endoribonuclease RNase III--the effect of dsRNA binding on gene expression. *Mol Microbiol.*, 1998, vol. 28 (3), 629-40 [0007]
- SHAHBABIAN et al. RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *The EMBO Journal*, 2009, vol. 28, 3523-3533 [0007]
- SHI ; BERG. Specific DNA-RNA hybrid binding by zinc finger proteins. *Science*, 1995, vol. 268 [0022]
- SHI Y ; BERG JM. Specific DNA-RNA hybrid binding proteins. *Science*, 1995, vol. 268, 282-284 [0027] [0046]
- CROOKS GE ; HON G ; CHANDONIA JM ; BRENNER SE. WebLogo: A sequence logo generator. *Genome Research*, 2004, vol. 14, 1188-1190 [0052]
- SCHNEIDER TD ; STEPHENS RM. Sequence Logos: A New Way to Display Consensus Sequences. *Nucleic Acids Res.*, 1990, vol. 18, 6097-6100 [0052]

35