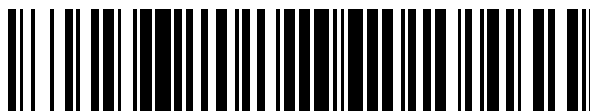


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 550 609**

51 Int. Cl.:

C12N 15/11 (2006.01)

C07K 14/47 (2006.01)

C12Q 1/68 (2006.01)

A61K 48/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.07.2003 E 03763812 (9)**

97 Fecha y número de publicación de la concesión europea: **22.07.2015 EP 1520022**

54 Título: **Interferencia de ARN mediante de moléculas de ARN de cadena sencilla**

30 Prioridad:

10.07.2002 EP 02015532

23.08.2002 EP 02018906

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.11.2015

73 Titular/es:

**MAX-PLANCK-GESELLSCHAFT ZUR
FÖRDERUNG DER WISSENSCHAFTEN E.V.
(100.0%)
HOFGARTENSTRASSE 8
80539 MÜNCHEN, DE**

72 Inventor/es:

**TUSCHL, THOMAS;
MARTINEZ, JAVIER;
PATKANIOWSKA, AGNIESZKA;
URLAUB, HENNING y
LÜHRMANN, REINHARD**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 550 609 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

Interferencia de ARN mediante de moléculas de ARN de cadena sencilla

La presente invención se refiere a características de la secuencia y estructurales de moléculas de ARN(ss) de cadena sencilla necesarias para mediar en modificaciones de ácido nucleico específicas para la diana mediante interferencia del ARN (ARNi) tales como la degradación del ARNm diana y/o la metilación del ADN.

La mayoría de los eucariotas poseen un sistema de defensa celular que protege sus genomas contra la invasión de elementos genéticos extraños. Se cree que la inserción de elementos extraños va acompañada generalmente de la formación de ARNds que es interpretado por la célula como una señal para la actividad génica indeseada (p. ej., Ahlquist, Science 296 (2002), 1270-1273; Fire et al., Nature 391 (1998), 806-811). Dicer RNasa III procesa rápidamente ARNds a pequeños fragmentos de ARNds de tamaño y estructura distintos (p. ej., Bernstein et al., Nature 409 (2001), 363-366), los ARNs pequeños de interferencia (ARNsi) (Elbashir et al., Genes & Dev. 15 (2001 b), 188-200), que dirigen la degradación específica para la secuencia de los ARNms de cadena sencilla de los genes invasores. ARNsi dúplex tienen extremos colgantes en 3' de 2 a 3 nt (nucleótidos) y contienen extremos 5' fosfato y 3' hidroxilo libres (documento WO 02/44321). El proceso de silenciamiento génico pos-transcripcional dependiente de ARNds se conoce comúnmente como interferencia de ARN (ARNi), y en algunos casos también está ligado al silenciamiento transcripcional.

La introducción experimental de ARNsi dúplex en células de mamífero se utiliza ahora ampliamente para interrumpir la actividad de genes celulares homólogos en secuencia al ARNds introducido. Utilizado como un enfoque genético inverso, el silenciamiento génico inducido por ARNsi acelera el enlace de la secuencia del gen a la función biológica. ARNsi dúplex son lo suficientemente cortos como para evitar efectos no específicos inducidos por ARNds generales en células de animales vertebrados y mamíferos. Los ARNsi también se pueden expresar intracelularmente a partir de plásmidos de expresión introducidas o vectores virales que proporcionan una alternativa a la síntesis química de ARN. Por lo tanto, una comprensión de cómo actúan los ARNsi en sistemas de mamíferos es importante para el perfeccionamiento de esta tecnología de silenciamiento génico y para la producción de agentes terapéuticos específicos para genes.

Estudios bioquímicos han comenzado a desvelar los detalles mecánicos de ARNi. Los primeros sistemas libres de células se desarrollaron utilizando extractos de células o embriones de *D. melanogaster*, y fueron seguidos por el desarrollo de sistemas *in vitro* de células de carcinoma embrionario de *C.elegans* y embrionario de ratón. Mientras que lisados de *D. melanogaster* apoyan las etapas de procesamiento de ARNds y la fijación como objetivo de ARNm específico para la secuencia, estos dos últimos sistemas sólo recapitulan la primera etapa.

El ARNi en extractos de *D. melanogaster* se inicia por el procesamiento dependiente de ATP de ARNds largo a ARNsi mediante Dicer RNasa III (p. ej., Bernstein et al., (2001), supra). A partir de entonces, ARNsi dúplex se ensamblan en un complejo multi-componente que guía al reconocimiento específico para la secuencia del ARNm diana y cataliza su escisión (p. ej., Elbashir (2001 b), supra). A este complejo se le alude como complejo de silenciamiento inducido por ARN (RISC) (Hammond et al., Nature 404 (2000), 293-296). Los ARNsi en *D. melanogaster* son predominantemente de 21 y 22 nt, y cuando se combinan de manera que contengan una estructura colgante en 3' de 2 nt se introducen de manera efectiva en el RISC (Elbashir et al., EMBO J. 20 (2001 c), 6877-6888). Sistemas de mamíferos tienen ARNsi de un tamaño similar, y ARNsi de 21 y 22 nt también representan los tamaños más eficaces para silenciar los genes expresados en células de mamífero (p. ej., Elbashir et al., Nature 411 (2001 a), 494-498, Elbashir et al., Methods 26 (2002), 199-213).

El RISC ensamblado en ARNsi dúplex en lisados de embrión de *D. melanogaster* fija como objetivo ARNs de cadena sencilla sentido así como antisentido homólogos para la degradación. Los sitios de escisión para los ARNs diana sentido y antisentido están situados en el centro de la región abarcada por el ARNsi dúplex. De manera importante, el extremo 5', y no el extremo 3', del ARNsi guía establece la regla para la posición de la escisión del ARN diana. Además, se requiere un fosfato 5' en la cadena complementaria de la diana de un ARNsi dúplex para la actividad de RISC, y el ATP se utiliza para mantener los 5' fosfatos de los ARNsi (Nykänen et al., Cell 107 (2001), 309-321). ARNsi dúplex sintéticos con colgantes 5' hidroxilo y 3' de 2 nt libres son fosforilados tan fácilmente en lisados de embrión de *D. melanogaster* que las eficiencias de ARNi de ARNsi 5'-fosforilados y no fosforilados no son significativamente diferentes (Elbashir et al. (2001 c), supra).

El desenrollamiento del ARNsi dúplex debe producirse antes del reconocimiento del ARN diana. Análisis de los requisitos de ATP revelaron que la formación de RISC en ARNsi dúplex requería ATP en lisados de *D. melanogaster*. Una vez formado, el RSIC escinde el ARN diana en ausencia de ATP. La necesidad de ATP

probablemente refleja la etapa de desenrollamiento y/u otros reordenamientos conformacionales. Sin embargo, actualmente se desconoce si las cadenas desenrolladas de un ARNsi dúplex permanecen asociadas con RISC o si RISC sólo contiene un ARNsi de cadena sencilla.

Un componente asociado con el RISC fue identificado como Argonaute2 a partir de células Schneider 2 (S2) de *D. melanogaster* (Hammond et al., Science 293 (2001 a), 1146-1150), y es un miembro de una gran familia de proteínas. La familia se conoce como familia Argonaute o PPD y se caracteriza por la presencia de un dominio PAZ y un dominio Piwi C-terminal, tanto de función desconocida (Cerutti et al., Trends Biochem. Sci. (2000), 481-482); Schwarz y Zamore, Genes & Dev. 16 (2002), 1025-1031). El dominio PAZ también se encuentra en Dicer. Debido a que Dicer y Argonaute2 interactúan en células S2, PAZ puede funcionar como un motivo de interacción proteína-proteína. Posiblemente, la interacción entre Dicer y Argonaute2 facilita la incorporación ARNsi en RISC. En *D. melanogaster*, la familia Argonaute tiene cinco miembros, la mayoría de los cuales se demostró que está implicada en el silenciamiento y el desarrollo de genes. Los miembros de mamífero de la familia Argonaute están mal caracterizados, y algunos de ellos han sido implicados en el control de la traducción, procesamiento de microARN y el desarrollo. Queda por establecer la función bioquímica de las proteínas Argonaute y el desarrollo de más sistemas bioquímicos es crucial.

En esta memoria, los autores de la invención informan sobre el análisis de RISC humano en extractos preparados a partir de células HeLa. La reconstitución de RISC y la etapa de fijación como objetivo del ARNm revelaron que RISC es un complejo de ribonucleoproteína que se compone de un ARNsi de cadena sencilla. Una vez que se forma el RISC, el ARNsi incorporado ya no puede intercambiarse con ARNsi libres. Sorprendentemente, RISC puede ser reconstituido en extractos HeLa S100, proporcionando ARNsi de cadena sencilla. La introducción de ARNsi antisentido de cadena sencilla 5'-fosforilados en células HeLa silencia potentemente un gen endógeno con una eficacia similar al ARNsi dúplex.

El objeto subyacente de la presente invención es proporcionar nuevos agentes capaces de mediar en el ARNi específico para la diana.

La solución de este problema se proporciona mediante el uso de una molécula de ARN de cadena sencilla según se define en la reivindicación 1. Sorprendentemente, se encontró que moléculas de ARN de cadena sencilla son capaces de inhibir la expresión de transcritos diana mediante la interferencia del ARN (ARNi).

La longitud de las moléculas de ARN de cadena sencilla es preferiblemente de 14-50 nt, en donde al menos del 14 al 20 en 5' la mayoría de los nucleótidos son sustancialmente complementarios al transcrito de ARN diana. Los oligonucleótidos de ARN están modificados mediante análogos 5'-monofosfato según se define en la reivindicación 1.

La inhibición de la expresión del transcrito diana puede producirse in vitro, p. ej. en cultivos celulares o extractos de células eucarióticas, particularmente de mamíferos. Por otro lado, la inhibición también puede producirse in vivo, es decir, en organismos eucarióticos, particularmente de mamíferos incluyendo los seres humanos.

Preferiblemente, la molécula de ARN de cadena sencilla tiene una longitud de 15 - 29 nucleótidos. La cadena de ARN puede tener un grupo 3'-hidroxilo. En algunos casos, sin embargo, puede ser preferible modificar el extremo 3' para hacerlo resistente contra 3' a 5' exonucleasas. Modificaciones 3' toleradas son, por ejemplo, nucleótidos 2'-desoxi terminales, 3'-fosfato, 2',3'-fosfato cíclico, aminoenzalizador C3 (o C6, C7, C12), enlazadores tiol, enlazadores carboxilo, espaciadores no nucleotídicos (C3, C6, C9, C12 abásicos, trietilenglicol, hexaetilenglicol), biotina, fluoresceína, etc.

El extremo 5' comprende un análogo fosfato. Modificaciones 5'-fosfato preferidas son 5'-monofosfato ((HO)₂(O)P-O-5'), 5'-difosfato ((HO)₂(O)P-O-P(HO)(O)-O-5'), 5'-trifosfato ((HO)₂(O)P-O-(HO)(O)P-OP(HO)(O)-O-5'), remate 5'-guanosina (7-metilado o no metilado) (7m-G-O-5'-(HO)(O)P-O-(HO)(O)P-OP(HO)(O)-O-5'), remate 5'-adenosina (Appp), y cualquier estructura de remate de nucleótidos modificada o no modificada (N-O-5'-(HO)(O)PO-(HO)(O)P-O-P(HO)(O)-O-5'), 5'-monotiofosfato (fosforotioato; (HO)₂(S)P-O-5'), 5'-monoditiofosfato (fosforoditioato; (HO)(HS)(S)P-O-5'), 5'-fosforotiolato ((HO)₂(O)P-S-5'); cualquier combinación adicional de monofosfato, difosfato y trifosfatos reemplazados por oxígeno/azufre (p. ej. 5'-alfa-tiotrifosfato, 5'-gamma-tiotrifosfato, etc.), 5'-fosforoamidatos ((HO)₂(O)P-NH-5', (HO)(NH₂)(O)P-O-5'), 5'-alquilfosfonatos (R=alquilo=metilo, etilo, isopropilo, propilo, etc., p. ej., RP(OH)(O)-O-5'-, (OH)₂(O)P-5'-CH₂-), 5'-alquileterfosfonatos (R=alquileter=metoximetilo (MeOCH₂-), etoximetilo, etc., p. ej. RP(OH)(O)-O-5'-).

La secuencia de la molécula de ARN de la presente invención tiene que tener una identidad suficiente con una molécula de ácido nucleico diana con el fin de mediar en el ARNi específico para la diana. Por lo tanto, la molécula de ARN de cadena sencilla de la presente invención es sustancialmente complementaria al transcrito diana.

La reacción de escisión del ARN diana guiada por las moléculas de ARN de una sola hebra de la presente invención es altamente específico para la secuencia. Sin embargo, no todas las posiciones de la molécula de ARN contribuyen igualmente al reconocimiento de la diana. Se toleran emparejamientos erróneos, particularmente en el extremo 3' de

la molécula de ARN de cadena sencilla, más particularmente los residuos 3' con respecto a los primeros 20 nt de la molécula de ARN de cadena sencilla. Especialmente preferidas son moléculas de ARN de cadena sencilla que tienen en el extremo 5' al menos 15 y preferiblemente al menos 20 nucleótidos que son completamente complementarios a un transcrito diana predeterminado o tienen un emparejamiento erróneo sólo y opcionalmente hasta 35 nucleótidos en el extremo 3' que puede contener 1 o varios, p. ej. 2, 3 o más emparejamientos erróneos.

Con el fin de mejorar la estabilidad de las moléculas de ARN de cadena sencilla, los extremos 3' pueden estabilizarse contra la degradación, p. ej., pueden seleccionarse de manera que consistan en nucleótidos de purina, particularmente nucleótidos de adenosina o guanosina. Alternativa o adicionalmente, 3' nucleótidos pueden estar sustituidos con análogos de nucleótidos modificados, incluyendo modificaciones de la cadena principal de residuos ribosa y/o fosfato.

En una realización especialmente preferida de la presente invención, la molécula de ARN puede contener al menos un análogo de nucleótido modificado. Los análogos de nucleótidos pueden estar localizados en posiciones en las que la actividad específica de la diana, p. ej., la actividad mediadora de ARNi no se ve afectada sustancialmente, p. ej., en una región en el extremo 5' y/o el extremo 3' de la molécula de ARN. En particular, el extremo 3' puede ser estabilizado mediante la incorporación de análogos de nucleótidos modificados tales como derivados químicos no nucleotídicos tales como aminoenzalizador C3 (o C6, C7, C12), enlazadores tiol, enlazadores carboxilo, espaciadores no nucleotídicos (C3, C6, C9, C12, abásicos, trietilenglicol, hexaetilenglicol), biotina, fluoresceína, etc. Una modificación adicional mediante la cual puede aumentarse la resistencia nucleasa de la molécula de ARN, es mediante el acoplamiento covalente de nucleótidos invertidos, p. ej., 2'-desoxirribonucleótidos o ribonucleótidos al extremo 3' de la molécula de ARN. Una estructura de molécula de ARN preferida comprende: ARNsi-3'-O-P(O)(OH)-O-3'-N 5' de cadena sencilla, en donde N es un nucleótido, p. ej., un 2'-desoxirribonucleótido o ribonucleótido, típicamente un residuo timidina invertido, o una estructura de oligonucleótido invertida, p. ej., que contiene hasta 5 nucleótidos.

Análogos de nucleótidos preferidos se seleccionan de ribonucleótidos modificados con azúcares o en la cadena principal. Cabe señalar, sin embargo, que también son adecuados ribonucleótidos modificados con nucleobases, es decir, ribonucleótidos que contienen una nucleobase de origen no natural en lugar de una nucleobase de origen natural tales como uridinas o citidinas modificados en la posición 5, p. ej., 5-(2-amino)propil-uridina, 5-bromo-uridina; 5-metil-citidina; adenosinas y guanosinas modificadas en la posición 8, p. ej., 8-bromo-guanosina; desaza nucleótidos, p. ej. 7-deaza-adenosina; nucleótidos O- y N-alquilados, p. ej. N6-metil-adenosina. En ribonucleótidos modificados con azúcares preferidas, el grupo OH 2 se reemplaza por un grupo seleccionado de H, OR, R, halo, SH, SR, NH₂, NHR, NR₂ o CN, en donde R es alquilo C₁-C₆, alquenoilo, alquinoilo o metoxietoxi, y halo es F, Cl, Br o I. En ribonucleótidos modificados en la cadena principal preferidos, el grupo fosfoéster que se conecta a ribonucleótidos adyacentes se reemplaza por un grupo modificado, p. ej., un fosforotioato, fosforoditioato, grupo fosforoamidato N3'-O5' y/o N5'-O3'. Cabe señalar que las modificaciones anteriores pueden combinarse. Por ejemplo, nucleótidos complementarios o no complementarios en el extremo 3', particularmente después de al menos 15, más particularmente después de al menos 20 nucleótidos 5'-terminales pueden modificarse sin pérdida significativa de la actividad.

La molécula de ARN de cadena sencilla de la invención se puede preparar mediante síntesis química. Métodos de síntesis de moléculas de ARN son conocidos en la técnica.

Los ARNs de cadena sencilla se pueden preparar también mediante transcripción enzimática a partir de moldes de ADN sintéticos o de plásmidos de ADN aislados de bacterias recombinantes y la subsiguiente modificación 5'-terminal. Típicamente, se utilizan las ARN polimerasas de fagos tal como T7, T3 o SP6 ARN polimerasa.

Un aspecto adicional de la presente invención se refiere a un método según se define en la reivindicación 14.

Preferiblemente, la etapa de contacto (a) comprende introducir la molécula de ARN de cadena sencilla en una célula diana, p. ej., una célula diana aislada, p. ej., en cultivo de células, un microorganismo unicelular o una célula diana o una pluralidad de células diana dentro de un organismo multicelular. Más preferiblemente, la etapa de introducción comprende un suministro mediado por el soporte, p. ej., mediante soportes de liposomas y/o mediante inyección. Sistemas de suministro adicionales adecuados incluyen Oligofectamine (Invitrogen) y reactivo de transfección de ARNsi Transit-TKO (Mirus)

El método de la invención puede ser utilizado para determinar la función de un gen en una célula o un organismo o incluso para modular la función de un gen en una célula o un organismo, que es capaz de mediar en la interferencia de ARN. La célula es preferiblemente una célula eucariótica o una línea celular, p. ej., una célula vegetal o una célula animal, tal como una célula de mamífero, p. ej. una célula embrionaria, una célula madre pluripotente, una célula tumoral, p. ej., una célula de teratocarcinoma o una célula infectada con virus. El organismo es preferiblemente un organismo eucariótico, p. ej., un vegetal o un animal tal como un mamífero, particularmente un ser humano.

El gen diana al que se dirige la molécula de ARN de la invención puede estar asociado con un estado patológico. Por ejemplo, el gen puede ser un gen asociado a patógenos, p. ej., un gen viral, un gen asociado a un tumor o un gen asociado a una enfermedad autoinmune. El gen diana puede ser también un gen heterólogo expresado en una célula recombinante o un organismo alterado genéticamente. Al determinar o modular, en particular, inhibir la función de un gen de este tipo se pueden obtener información valiosa y beneficios terapéuticos en el campo de la agricultura o en el campo de la medicina o la medicina veterinaria.

El ARNss se administra habitualmente en forma de una composición farmacéutica. La administración puede llevarse a cabo por métodos conocidos, en los que un ácido nucleico se introduce en una célula diana deseada in vitro o in vivo. Técnicas de transferencia de genes comúnmente utilizadas incluyen fosfato de calcio, DEAE-dextrano, electroporación y microinyección y métodos virales (Graham, F.L. y van der Eb, A.J. (1973) *Virology* 52, 456; McCutchan, J.H. y Pagano, J.S. (1968), *J. Natl. Cancer Inst.* 41, 351; Chu, G. et al (1987), *Nucl. Acids Res.* 15, 1311; Fraley, R. et al. (1980), *J. Biol. Chem.* 255, 10431; Capecchi, M.R. (1980), *Cell* 22, 479). Una adición reciente a este arsenal de técnicas para la introducción de ácidos nucleicos en células es el uso de liposomas catiónicos (Felgner, P.L. et al. (1987), *Proc. Natl. Acad. Sci USA* 84, 7413). Formulaciones de lípidos catiónicos comercialmente disponibles son, p. ej., Tfx 50 (Promega) o Lipofectamin2000 (Life Technologies). Un método preferido adicional para la introducción de ARN en un organismo diana, en particular en un ratón, es el de la inyección en la vena de la cola a alta presión (Lewis, D.L. et al (2002), *Nat. Genet.* 29, 29; McCaffrey, A.P. et al. (2002), *Nature* 418, 38-39). En esta memoria, una disolución tamponada que comprende el ARN de cadena sencilla (p. ej., aproximadamente 2 ml) se inyecta en la vena de la cola del ratón en el espacio de 10 s.

Por lo tanto, la invención se refiere también a una composición farmacéutica que contiene como agente activo al menos una molécula de ARN de cadena sencilla tal como se describió anteriormente y un soporte farmacéutico. La composición puede utilizarse para el diagnóstico y para aplicaciones terapéuticas en medicina humana o en medicina veterinaria.

Para aplicaciones diagnósticas o terapéuticas, la composición puede estar en forma de una disolución, p. ej., una disolución inyectable, una crema, pomada, comprimido, suspensión o similares. La composición se puede administrar de cualquier forma adecuada, p. ej., por inyección, por aplicación oral, tópica, nasal, rectal, etc. El soporte puede ser cualquier soporte farmacéutico adecuado. Preferiblemente, se utiliza un soporte que sea capaz de aumentar la eficacia de las moléculas de ARN para penetrar en las células diana. Ejemplos adecuados de soportes de este tipo son liposomas, particularmente liposomas catiónicos. Un método de administración adicionalmente preferido es la inyección.

Una aplicación preferida adicional del método de ARNi es un análisis funcional de células eucarióticas, u organismos eucarióticos no humanos, preferiblemente células u organismos de mamíferos y lo más preferiblemente células de seres humanos, p. ej., líneas de células tales como HeLa o 293 o de roedores, p. ej., ratas y ratones. Mediante transfección con moléculas de ARN de cadena sencilla adecuadas, que son homólogas a un gen diana predeterminado o a moléculas de ADN que codifican una molécula de ARN de cadena sencilla adecuada, se puede obtener un fenotipo inactivado específico en una célula diana, p. ej., en cultivo celular o en un organismo diana. La presencia de moléculas de ARN de cadena sencilla cortas no da como resultado una respuesta de interferón de la célula huésped u organismo huésped.

En una realización especialmente preferida, la molécula de ARN se administra asociada con polímeros biodegradables, p. ej. polipéptidos, ácido poli(d,l-láctico-co-glicólico) (PLGA), polilisina o conjugados de polilisina, p. ej., polilisina-injerto-ácido imidazol-acético o poli(éster de beta-amino) o micropartículas tales como microesferas, nanopartículas o nanoesferas. Más preferiblemente, la molécula de ARN se acopla covalentemente al polímero o micropartícula, en donde el acoplamiento covalente se efectúa particularmente a través del extremo 3' de la molécula de ARN.

También se describe una composición farmacéutica para inhibir la expresión de un transcrito diana mediante ARNi que comprende, como agente activo, una molécula de ARN de cadena sencilla que tiene una longitud de 14-50, preferiblemente de 15-29 nucleótidos, en donde al menos los 14-20 nucleótidos más próximos a 5' son sustancialmente complementarios a dicho transcrito diana.

Además, el uso de acuerdo con la reivindicación 1 es preferiblemente para la prevención o el tratamiento de una enfermedad asociada con la sobre-expresión de al menos un gen diana que comprende administrar a un sujeto en necesidad del mismo, una molécula de ARN de cadena sencilla que tiene una longitud de 14-50, preferiblemente 15-29 nucleótidos, en donde al menos los 14-20 nucleótidos más 5' son sustancialmente complementarios a un transcrito diana en una cantidad que es terapéuticamente eficaz para el ARNi.

Se describe, además, una célula eucariótica o un organismo eucariótico no humano que exhibe un fenotipo inactivado específico para el gen diana que comprende una expresión, al menos parcialmente deficiente, de al

menos un gen diana endógeno, en donde dicha célula u organismo es transfectado con al menos una molécula de ARN de cadena sencilla capaz de inhibir la expresión de al menos un gen diana endógeno. Cabe señalar que la presente invención permite el suministro simultáneo de varios ARNs antisentido de diferentes secuencias, que son ya sea cognatas a un gen diferente o al mismo gen diana.

- 5 Fenotipos inactivados específicos para genes de células u organismos no humanos, en particular de células humanas o de mamíferos no humanos pueden ser utilizados en procesos analíticos, p. ej., en el análisis funcional y/o fenotípico de procesos fisiológicos complejos tales como análisis de perfiles de expresión génica y/o proteomas. Por ejemplo, se pueden preparar los fenotipos inactivados de genes humanos en células cultivadas que se supone son los reguladores de los procesos de corte y empalme alternativos. Entre estos genes se encuentran particularmente los miembros de la familia del factor de corte y empalme SR, p. ej., ASF/SF2, SC35, SRp20, SRp40 o SRp55. Además, se puede analizar el efecto de las proteínas SR en los perfiles de ARNm de genes predeterminados cortados y empalmados alternativamente tales como CD44. Preferiblemente, el análisis se lleva a cabo por métodos de alto rendimiento utilizando chips basados en oligonucleótidos.

- 15 Utilizando tecnologías de inactivación basadas en ARNi, la expresión de un gen diana endógeno puede ser inhibida en una célula diana o un organismo diana. El gen endógeno puede ser complementado por un ácido nucleico diana exógeno que codifica la proteína diana o una variante o forma mutada de la proteína diana, p. ej., un gen o un ADNc, que puede estar opcionalmente fusionado a una secuencia de ácido nucleico adicional que codifica un péptido o polipéptido detectable, p. ej., una etiqueta de afinidad, en particular una etiqueta de afinidad múltiple. Variantes o formas mutadas del gen diana difieren del gen diana endógeno en que codifican un producto génico que difiere del producto del gen endógeno en el nivel de aminoácidos por sustituciones, inserciones y/o deleciones de aminoácidos individuales o múltiples. Las variantes o formas mutadas pueden tener la misma actividad biológica que el gen diana endógeno. Por otro lado, el gen diana variante o mutado también pueden tener una actividad biológica que difiere de la actividad biológica del gen diana endógeno, p. ej., una actividad parcialmente suprimida, una actividad completamente suprimida, una actividad mejorada, etc.

- 25 La complementación puede conseguirse co-expresando el polipéptido codificado por el ácido nucleico exógeno, p. ej., una proteína de fusión que comprende la proteína diana y la etiqueta de afinidad y la molécula de ARN de doble cadena para la inactivación del gen endógeno en la célula diana. Esta co-expresión se puede lograr utilizando un vector de expresión adecuado que exprese tanto el polipéptido codificado por el ácido nucleico exógeno, p. ej., la proteína diana modificada con la etiqueta y la molécula de ARN de cadena sencilla o, alternativamente, utilizando una combinación de vectores de expresión. Las proteínas y complejos de proteínas que se sintetizan de novo en la célula diana contendrán el producto génico exógeno, p. ej., la proteína de fusión modificada. Con el fin de evitar la supresión de la expresión del producto génico exógeno por parte de la molécula de ARNi, la secuencia de nucleótidos que codifica el ácido nucleico exógeno puede ser alterada en el nivel de ADN (con o sin provocar mutaciones en el nivel de aminoácidos) en la parte de la secuencia que es homóloga a la molécula de ARN de cadena sencilla. Alternativamente, el gen diana endógeno puede ser complementado por las correspondientes secuencias de nucleótidos de otras especies, p. ej., de ratón.

- Aplicaciones preferidas para la célula u organismo descrito en esta memoria son los análisis de perfiles de expresión de genes y/o proteomas. En una realización especialmente preferida, se lleva a cabo un análisis de una variante o forma mutante de una o varias proteínas diana, en donde dicha variante o forma mutante se reintroduce en la célula u organismo mediante un ácido nucleico diana exógeno tal como se describe anteriormente. La combinación de inactivación de un gen endógeno y el rescate utilizando una diana mutada, p. ej., una diana exógena parcialmente suprimida tiene ventajas en comparación con el uso de una célula inactivada. Además, este método es particularmente adecuado para la identificación de dominios funcionales de la proteína diana. En una realización preferida adicional, se lleva a cabo una comparación, p. ej., de perfiles de expresión de genes y/o proteomas y/o características fenotípicas de al menos dos células u organismos. Estos organismos se seleccionan de:

- (i) una célula de control u organismo de control sin inhibición del gen diana,
- (ii) una célula u organismo con inhibición del gen diana y
- (iii) una célula u organismo con la inhibición del gen diana más la complementación del gen diana por parte de un ácido nucleico diana exógeno.

- 50 El método de la invención también se puede utilizar en un proceso para identificar y/o caracterizar agentes farmacológicos, p. ej., identificar nuevos agentes farmacológicos a partir de una colección de sustancias de ensayo y/o la caracterización de los mecanismos de acción y/o efectos secundarios de agentes farmacológicos conocidos.

También se describe un sistema para la identificación y/o caracterización de agentes farmacológicos que actúan sobre al menos una proteína diana, que comprende:

- (a) una célula eucariótica o un organismo eucariótico no humano capaz de expresar al menos un gen diana endógeno que codifica dicha proteína diana,
 (b) al menos una molécula de ARN de cadena sencilla, capaz de inhibir la expresión de dicho al menos un gen diana endógeno por parte de ARNi y
 (c) una sustancia de ensayo o una colección de sustancias de ensayo, en donde las propiedades farmacológicas de dicha sustancia de ensayo o dicho colección han de ser identificadas y/o caracterizadas.

Además, el sistema, tal como se describe anteriormente, comprende preferiblemente:

- (d) al menos un ácido nucleico exógeno diana que codifica la proteína diana o una variante o forma mutada de la proteína diana, en donde dicho ácido nucleico exógeno diana difiere del gen endógeno diana al nivel de ácido nucleico de tal manera que la expresión del ácido nucleico diana exógeno es sustancialmente menos inhibida por la molécula de ARN de cadena sencilla que la expresión del gen endógeno diana.

Además, el método de complementación de ARN inactivado puede utilizarse para fines preparativos, p. ej. para la purificación por afinidad de proteínas o complejos de proteínas de células eucarióticas, particularmente células de mamíferos y células más particularmente humanas. El ácido nucleico exógeno diana puede codificar una proteína diana, que se fusiona a una etiqueta de afinidad.

El método preparativo se puede emplear para la purificación de complejos de proteínas de alto peso molecular que tienen preferiblemente una masa de ≥ 150 kD y más preferiblemente de ≥ 500 kD y que, opcionalmente, pueden contener ácidos nucleicos tales como ARN. Ejemplos específicos son el complejo de proteínas heterotrimérico que consiste en las proteínas de 20 kD, 60 kD y 90 kD de la partícula snRNP U4/U6, consistiendo el factor de corte y empalme SF3b de 17S U2 snRNP en 5 proteínas que tienen pesos moleculares de 14, 49, 120, 145 y 155 kD y conteniendo la partícula tri-snRNP 25S U4/U6/U5 las moléculas U4, U5 y U6 ARNs y aproximadamente 30 proteínas, que tiene un peso molecular de aproximadamente 1,7 MD.

Este método es adecuado para el análisis de proteoma funcional en células de mamífero, particularmente células de seres humanos.

Se describe un complejo de silenciamiento inducido por ARN de mamífero, particularmente humano (RISC) purificado y aislado, que tiene un peso molecular aparente de menos de aproximadamente 150-160 kDa, p. ej., de aproximadamente 120 a 150-160 kDa. El RISC comprende componentes de polipéptido y, opcionalmente, de ácido nucleico, en particular moléculas de ARN de cadena sencilla tal como se ha descrito anteriormente. El RISC se puede utilizar como una diana para el diagnóstico y/o terapia, como un agente de diagnóstico y/o terapéutico propiamente dicho, como un reactivo de biología molecular o como componente en un proceso de cribado para la identificación y/o caracterización de agentes farmacéuticos.

Componentes polipeptídicos de RISC pueden comprender miembros de la familia Argonaute de proteínas y contienen eIF2C1 y/o eIF2C2, y posiblemente al menos otro miembro de la familia eIF2C expresado, en particular seleccionado de eIF2C3, eIF2C4, HILI y HIWI.

La expresión o sobre-expresión de una o varias proteínas presentes en el RISC en células huésped adecuadas, p. ej., células eucarióticas, particularmente células de mamíferos, es útil para ayudar a una respuesta de ARNi. Estas proteínas también se pueden expresar o sobre-expresar en animales transgénicos, p. ej., vertebrados, particularmente mamíferos, para producir animales particularmente sensibles a ARNs de cadena sencilla o de doble cadena inyectados. Además, los genes que codifican las proteínas pueden ser administrados con fines terapéuticos, p. ej., por medio de vectores de suministro de genes virales o no virales.

También es concebible administrar un complejo de ARNs/eIF2C1 ó 2 directamente mediante la asistencia de reactivos de transfección de proteína (p. ej., Protein Transfection Reagents, proteína ProVectin (Imgenex), o productos similares) en lugar de la transfección de ARN/ADN. Esto puede tener ventajas técnicas sobre la transfección de ARNs que se limita a la transfección de ácido nucleico.

Además, la presente invención se explica con mayor detalle en las siguientes figuras y ejemplos.

Leyendas de las figuras

Figura 1. Extractos citoplasmáticos HeLa S100 muestran una escisión de ARN diana dependiente de ARNs.

(A) Representación del ARN diana de 177 nt marcado con remate ^{32}P con el ARNs dúplex diana. También se muestra de sitio de escisión del ARN diana y la longitud de los productos de escisión esperados. La línea

gruesa en negro colocada debajo del ARNsi antisentido se utiliza en las siguientes figuras como símbolo para indicar la región del ARN diana, que es complementaria a la secuencia de ARNsi antisentido. (B) Comparación del sitio de escisión del ARN diana mediada por ARNsi utilizando el sistema de embrión de *D. melanogaster* in vitro previamente establecido y extracto citoplasmático de células HeLa S100. ARN diana 10 nM marcado en el remate se incubó con ARNsi 100 nM tal como se describe en materiales. Los productos de reacción se resolvieron en un gel de secuenciación al 6%. Marcadores de posición se generaron mediante digestión parcial con RNasa T1 (T1) e hidrólisis alcalina parcial (OH) del RNA diana marcado en el remate. La flecha indica el producto de escisión 5', el fragmento 3' no está marcado y, por lo tanto, es invisible.

Figura 2. La modificación química del extremo 5' del ARNsi antisentido pero no del ARNsi sentido impide la escisión del ARN diana sentido en extractos HeLa S100. (A) Ilustración de las posibles modificaciones 5' y 3' del aminoenzalizador y de las cadenas sentido y antisentido de un ARNsi dúplex. L5 representa un aminoenzalizador de la cadena de 6 carbonos conectado a través de un enlace 5'-fosfodiéster, L3 representa un aminoenzalizador de 7 carbonos conectado a través de un enlace fosfodiéster al 3'-fosfato terminal. s, sentido; as, antisentido. (B) Escisión de ARN diana ensayando diversas combinaciones de ARNsi dúplex modificado en el aminoenzalizador en 5' y 3'. NC (control negativo) muestra una reacción de incubación del ARN diana en ausencia de ARNsi dúplex. T1, escalera de RNasa T1; OH, escalera de hidrólisis alcalina parcial.

Figura 3. ARNsi que contiene fosfatos 3'-terminales se someten a reacciones de ligamiento, así como de desfosforilación.

(A) Secuencia del ARNsi dúplex radiomarcado. El nucleótido marcado se unió al ARNsi sintético de 20 nt, antisentido, mediante ligamiento con T4 ARN de 32pCp. Las diversas combinaciones de hidroxilo/fosfato 5' y 3' se prepararon como se describe en materiales. X e Y indican modificaciones 5' y 3' del ARNsi antisentido. (B) Destino del ARNsi antisentido durante la incubación de los ARNsi dúplex modificados en extracto HeLa S100 en presencia de ARN diana no radiomarcado. Las diferentes formas fosforiladas del ARNsi antisentido se distinguen sobre la base de su movilidad en gel. Se obtuvieron resultados idénticos cuando se utiliza ARNsi sentido 5' fosforilado o cuando se deja fuera el ARN diana durante la incubación. Los productos de ligamiento sólo se observan cuando 3' fosfatos estaban presentes en el ARNsi antisentido marcado.

Figura 4. RISC es un complejo estable que no intercambia rápidamente ARNsi unido.

Concentraciones crecientes de ARNsi no específico compiten con la formación de RISC específico de la diana cuando se añade simultáneamente a los extractos HeLa S100 (pistas 4 a 7). Sin embargo, cuando el ARNsi dúplex inespecífico se añade 15 min después de la pre-incubación con el ARNsi dúplex específico, no se observó una mayor competencia (3 pistas a la derecha). T1, escalera de RNasa T1.

Figura 5. Purificación parcial de RISC humano.

(A) Representación gráfica de la estructura del ARNsi dúplex biotinilado utilizado para la purificación por afinidad de los factores asociados a ARNsi. L3 indica un aminoenzalizador de C7 que se conjugó con un éster N-hidroxisuccinimidílico de biotina foto-escindible; UV indica la fotoescisión del enlace sensible a UV para liberar complejos seleccionados de afinidad en condiciones nativas. (B) Análisis de filtración en gel Superdex-200 de complejos de ARNsi-proteína (siRNPs) recuperados mediante tratamiento/elución UV (UV elu) de la columna de afinidad de estreptavidina. Las fracciones se ensayaron en cuanto a su capacidad para escindir específicamente el ARN diana marcado en el remate. Se indica el número de las fracciones recogidas y las posiciones relativas de los marcadores de tamaño de aldolasa (158 kDa) y BSA (66 kDa). (C) Sedimentación con gradiente de glicerol (5%-20%) de los siRNPs recuperados mediante tratamiento/elución UV de la columna de afinidad de estreptavidina. Para la leyenda, véase (B). Al supervisar el tamaño preciso de los fragmentos de escisión de ARN diana utilizando ARNm marcado internamente con ³²P-UTP, rematado, la suma es igual al transcrito de longitud completa, lo que indica que el ARN diana de hecho sólo se escinde una vez en el centro de la región abarcada por el ARNsi.

Figura 6. RISC contiene un ARNsi de cadena sencilla.

siRNPs fueron sometidos a selección por afinidad después de la incubación utilizando ARNsi dúplex con una o las dos cadenas biotiniladas. El producto eluido recuperado después del tratamiento UV o la fracción no unida después de selección por afinidad de estreptavidina (flujo no retenido) se ensayó en cuanto a la degradación de ARN diana. Si se biotiniló la cadena antisentido, todo el RISC de escisión de ARN diana sentido se unió a las perlas de estreptavidina, mientras que la biotinilación de ARNsi resultó en la actividad de RISC del flujo no retenido. La

reacción de escisión en la fracción de flujo no retenido fue menos eficiente que en el producto eluido de UV, porque su RISC seleccionado por afinidad era más concentrado.

Figura 7. ARNsis antisentido de cadena sencilla reconstituyen el RISC en extractos HeLa S100.

Análisis de la reconstitución de RISC mediante ARNsi de cadena sencilla o dúplex que compara extractos HeLa S100 (A) y el lisado de embrión de *D. melanogaster* anteriormente descrito (B). Diferentes concentraciones de ARNsi de cadena sencilla (s, sentido; as, antisentido) y ARNsi dúplex (ds) se sometieron a ensayo en cuanto a la fijación específica como objetivo de ARN sustrato marcado en el remate. Concentraciones 100 nM del ARNsi antisentido reconstituyeron RISC en el extracto HeLa S100, aunque a niveles reducidos en comparación con el ARNsi dúplex. Reconstitución con ARNsi de cadena sencilla era casi indetectable en lisado de *D. melanogaster*, presumiblemente debido a la mayor actividad de nucleasa en este lisado provocando una rápida degradación de los ARN de cadena sencilla no rematados.

Figura 8. ARNsi de cadena sencilla antisentido median en el silenciamiento de genes en células HeLa.

(A) Silenciamiento de la proteína lamina A/C de la envoltura nuclear. Tinción de fluorescencia de células transfectadas con ARNsi específicos para lamina A/C y GL2 luciferasa (control). Fila superior, tinción con anticuerpo específico para lamina A/C; fila central, tinción con Hoechst de la cromatina nuclear; fila inferior, las imágenes de contraste de fase de células fijadas. (B) Cuantificación del descenso de lamina A/C después de un análisis de transferencia Western. La mancha fue separada después del sondeo con lamina A/C y se volvió a sondear con anticuerpo vimentina. La cuantificación se realizó utilizando un aparato Lumi-Imager (Roche) y el software LumiAnalyst para cuantificar las señales ECL (Amersham Biosciences), las diferencias en la carga de gel fueron corregidas respecto a los niveles de proteína vimentina no fijada como objetivo. Los niveles de la proteína lamina A/C se normalizaron al dúplex de ARNsi GL2 no específico.

Figura 9. ARNsi antisentido de diferente longitud dirigen la escisión directa de ARN diana en extractos HeLa S100.

(A) Representación gráfica del experimento. ARNsi antisentido se extendieron hacia el lado 5' (serie 1, 20 a 25 nt) o el lado 3' (serie 2, 20 a 23 nt). (B) Escisión de ARN diana utilizando los ARNsi antisentido descritos en (A). Extracto HeLa S100 se incubó con ARN diana marcado en el remate 10 nM y ARNsi antisentido 100 nM a 30°C durante 2,5 h. Los productos de reacción se resolvieron en un gel de secuenciación al 6%. Marcadores de posición se generaron por digestión parcial con RNasa T1 (T1) e hidrólisis alcalina parcial (OH) del ARN diana marcado en el remate. Las flechas indican la posición de los productos de escisión 5' generados mediante los diferentes ARNsi antisentido. Las líneas negras gruesas a la izquierda (serie 1) y a la derecha (serie 2) indican la región del ARN diana, que es complementaria a las secuencias de ARNsi antisentido.

Figura 10. Dependencia de la longitud de ARNsi antisentido y efecto de las modificaciones terminales para fijar como objetivo la escisión del ARN en extractos HeLa S100.

Extracto HeLa S100 se incubó con ARN diana marcado en el remate 10 nM y ARNsi antisentido a 30°C durante 2,5 h. Los productos de reacción se resolvieron en un gel de secuenciación al 6%. Marcadores de posición se generaron por digestión parcial con RNasa T1 (T1) del ARN diana marcado en el remate. La línea negra gruesa a la izquierda indica la región del ARN diana, que es complementaria a la secuencia de ARNsi antisentido de 21 nt. Las secuencias de ARNsi utilizadas en cada uno de los experimentos se enumeran a continuación (los ARNsi sentido y antisentido se enumeran juntos, fueron pre-reasociados para formar ARNsi dúplex). p, fosfato; t, 2'-desoxitimidina, c, 2'-desoxicitidina, g, 2'-desoxicitidina, g, 2'-desoxiguanosina; L, aminoenzalizador, B, biotina fotoescindible; A, C, G, U, ribonucleótidos.

Pista	ARNsi sentido (5'-3')	ARNsi antisentido (5'-3')
1		pUCGAAGUAUUC CG CG
2		pUCGAAGUAUUC CG CGUACGUG
3		pUCGAAGUAUUC CG CGUACGUGAUGU
4		pUCGAAGUAUUC CG CGUACGUGAUGUUC
5		pUCGAAGUAUUC CG CGUACGUGAUGUUC AC
6		pUCGAAGUAUUC CG CG
7		pUCGAAGUAUUC CG CGUACGUG
8		pUCGAAGUAUUC CG CGUACGUGAUGU
9		pUCGAAGUAUUC CG CGUACGUGAUGUUC
10		pUCGAAGUAUUC CG CGUACGUGAUGUUC AC
11		pUCGAAGUAUUC CG CGUACGUG
12		pUCGAAGUAUUC CG

(continúa)

		CGUACGtg
13		pUCGAAGUAUUCCG CGUACGUU
14		pUCGAAGUAUUCCG CGUACGtt
15		pUCGAAGUAUUCCG CGUACGUG
16		pUCGAAGUAUUCCG CGUACGtg
17		pUCGAAGUAUUCCG CGUACGUU
18		pUCGAAGUAUUCCG CGUACGtt
19	CGUACGCGGAAUACUUCG AAA	pUCGAAGUAUUCCG CGUACGUG
20	CGUACGCGGAAUACUUCG AAA	pUCGAAGUAUUCCG CGUACGtg
21	CGUACGCGGAAUACUUCG AAA	pUCGAAGUAUUCCG CGUACGUU
22	CGUACGCGGAAUACUUCG AAA	pUCGAAGUAUUCCG CGUACGtt
23		tCGAAGUAUUCCGC GUACGUULB
24	cGUACGCGGAAUACUUCG AUULB	tCGAAGUAUUCCGC GUACGUULB
25		ptCGAAGUAUUCCGC GUACGttLB
26	cGUACGCGGAAUACUUCG	ptCGAAGUAUUCCGC

(continúa)

	AttLB	GUACGttLB
27		ptCGAAGUAUCCGC GUACGttL

Figura 11: ARNsi antisentido de cadena sencilla median en el silenciamiento de genes en células HeLa.

Cuantificación del descenso de lamina A/C después de análisis de transferencia Western. La mancha fue retirada después del sondeo con lamina A/C y se volvió a sondear con anticuerpo vimentina. La cuantificación se realizó con un aparato Lumi-Imager (Roche) y el software LumiAnalyst para cuantificar las señales ECL (Amersham Biosciences), las diferencias en la carga de gel fueron corregidas con respecto a los niveles de proteína vimentina no fijados como objetivo. Los niveles de la proteína lamina A/C se normalizaron al ARNsi dúplex GL2 no específico.

Figura 12. Composición de proteínas del RISC purificado por afinidad.

(A) Gel de SDS-PAGE teñido con plata de complejos de ribonucleoproteína seleccionados por afinidad después de sedimentación con gradiente de glicerol (5% -20%). La flecha indica la banda que contiene eIF2C1 y eIF2C2. A la izquierda se indican los marcadores de tamaño molecular. El asterisco indica una fracción para la cual el sedimento de proteína se perdió después de la precipitación. (B) Ensayo de escisión de ARN diana de las fracciones recogidas. La actividad del RISC alcanzó su punto máximo en la fracción 7 y 8; bu, tampón.

Figura 13. Caracterización por espectroscopía de masas de eIF2C1 y eIF2C2. La banda de 100 kDa se analizó mediante espectrometría de masas. Espectro de masas que indica los picos de péptidos correspondientes a eIF2C2 (A) y eIF2C1 (B). (C) Alineación de las secuencias de aminoácidos eIF2C2 y eIF2C1 que indican la posición de los péptidos identificados. Las diferencias de secuencia se indican mediante recuadros amarillos.

Figura 14. Predicción de las secuencias de aminoácidos de los seis miembros de la familia de proteínas Argonaute humanas.

Figura 15. Alineamiento de las secuencias de los seis miembros de la familia de proteínas Argonaute humanas. Secuencias predichas de eIF2C1-4 humana, HILI y HIWI han sido alineados utilizando el programa ClustalX.

Figura 16. Secuencias de ADNc previstas de los seis miembros de la familia de proteínas Argonaute humanas.

Figura 17. Todos los miembros de la familia Argonaute, a excepción de HIWI se expresan en células HeLa.

Análisis RT-PCR en ARN poliA de células HeLa. Los cebadores (A) (directo e inverso) utilizados para la amplificación por PCR anidada y semi-anidada de los diferentes Argonautes y la longitud esperada de los productos de PCR. (B) Electroforesis en gel de agarosa de los productos de PCR obtenidos, confirmando la longitud esperada. Pistas de la izquierda, escalera de ADN de 100 pb.

Ejemplo

1. Material y métodos

1.1 Síntesis de ARNsi y conjugación de biotina

ARNsi se sintetizaron químicamente utilizando fosforoamiditas de ARN (Proligo, Hamburg, Alemania) y se desprotegeron y se purificaron en gel tal como se describió anteriormente. Aminoenzaladores 5' se introdujeron mediante acoplamiento de MMT-aminoenzalador C6 fosforoamidita (Proligo, Hamburg). Aminoenzaladores C7 3' se introdujeron mediante el ensamblaje de la cadena de oligorribonucleótidos en aminomodificador 3' de soporte de vidrio de poro de control (TFA) C7 ICAA (Chemgenes, MA, EE.UU). Las secuencias de ARNsi de luciferasa GL2 eran como se ha descrito (Elbashir et al., 2001 a, supra). Si 5'-fosfatos hubieran de ser introducidos, 50 a 100 nmol de ARNsi sintéticos fueron tratados con polinucleótido T4 quinasa (reacción de 300 µl, ATP 2,5 mM, Tris-HCl 70 mM, pH 7,6, MgCl₂ 10 mM, DTT 5 mM, 30 U T4 PNK, New England Biolabs, 45 min, 37°C), seguido de precipitación con etanol.

El marcaje con 32pCp 3'-terminal (Figura 3) se realizó en una reacción de 30 µl (ARNsi 17 µM, 32pCp 0,5 µM (110 TBq/mmol), DMSO al 15%, 20 U de T4 ARN ligasa, NEB y tampón de reacción suministrado por 1x NEB) durante 1,5 h a 37°C, y se purificó en gel. La mitad del ARN marcado con pCp se desfosforiló (reacción de 25 µl, 500 U de fosfatasa alcalina, Roche, tampón suministrado por Roche, 30 min, 50°C), seguido por extracción con fenol/cloroformo y precipitación con etanol. La mitad de esta reacción se fosforiló en 5' (reacción de 20 µl, 2 unidades de T4 polinucleótido quinasa, NEB, ATP 10 mM, tampón suministrado por NEB, 60 min, 37°C). Una cuarta parte del ARNsi marcado con pCp inicial fue también fosforilado en 5' (reacción de 10 µl, 10 unidades de T4 polinucleótido quinasa 3' libre de fosfatasa, Roche, ATP 10 mM, tampón suministrado por Roche, 3 min, 37°C).

Para la conjugación de biotina, 20 a 65 nmol de ARNsi modificado con aminoenzalizador totalmente desprotegido se disolvieron en 100 µl de tampón borato de sodio 100 mM (pH 8,5) y se mezclaron con una disolución de 1 mg de EZ-Link NHS-LC-PC-Biotina (Pierce, IL, EE.UU.) en 100 µl de dimetilformamida anhidra. La disolución se incubó durante 17 h a 25°C en la oscuridad. Subsiguientemente, ARNsi se precipitaron mediante la adición de 60 µl de acetato de sodio 2 M (pH 6,0) y 1 ml de etanol. El sedimento de ARN se recogió mediante centrifugación y ARNsi conjugado con biotina fue separado de ARNsi que no había reaccionado en un gel de acrilamida al 18% desnaturizante preparativo (40 cm de longitud) en la oscuridad. Las bandas de ARN se visualizaron mediante sombreado a UV de 254 nm y se minimizó el tiempo de exposición. Las bandas fueron cortadas y el ARN se eluyó durante la noche en NaCl 0,3 M a 4°C y se recuperaron mediante precipitación en etanol. ARNsi dúplex se formaron tal como se describió previamente (Elbashir et al., Methods 26 (2002), 199-213).

1.2 Preparación de extractos S100 de células HeLa

El citoplasma de células HeLa adaptado para crecer a alta densidad se preparó siguiendo el protocolo de Dignam para el aislamiento de núcleos de células HeLa (Dignam et al., Nucleic Acids Res. 11 (1983), 1475-1489). La fracción citoplásmica se suplementó con KCl, MgCl₂ y glicerol a concentraciones finales de 100 mM, 2 mM y 10%, respectivamente. En esta fase, los extractos se pueden almacenar congelados a -70°C después de una congelación rápida en nitrógeno líquido sin pérdida de actividad. Extractos S100 se prepararon mediante ultracentrifugación a 31.500 rpm durante 60 minutos a 4°C utilizando un rotor Sorvall T-865. La concentración de proteínas de extracto HeLa S100 varió entre 4 y 5 mg/ml tal como se determina mediante el ensayo de Bradford.

1.3 Purificación por afinidad de RISC con ARNsi dúplex 3' biotinilados

Para la purificación por afinidad complejos de proteínas asociados a ARNsi a partir de extractos HeLa S100, 10 nM de un ARNsi dúplex 3' doble biotinilado se incubaron en ATP 0,2 mM, GTP 0,04 mM, 10 U/ml de RNasin, 6 µg/ml de creatina quinasa y fosfato de creatina 5 mM en 60% de extracto S100 a 30°C durante 30 a 60 min y una rotación suave. Después de esto, se añadió 1 ml de suspensión de Proteína de Unión a Biotina Neutravidina Inmovilizada (Pierce, IL, EE.UU.) por cada 50 ml de disolución de reacción y se continuó la incubación durante otros 60 a 120 min a 30°C con rotación suave. Las perlas de Neutravidina se recogieron después a 2000 rpm durante 2 minutos a 4°C en una centrífuga Heraeus Megafuge 1.0 R utilizando un rotor de cubeta oscilante tipo 2704. La captura eficaz de los componentes del RISC después de la selección por afinidad se confirmó ensayando el sobrenadante en cuanto a la actividad de RISC residual con y sin suplementar ARNsi dúplex recientes. Las perlas de Neutravidina recogidas se lavaron con 10 volúmenes de tampón A con relación al volumen de las perlas (HEPES 30 mM, pH 7,4, KCl 100 mM, MgCl₂ 2 mM, DTT 0,5 mM, glicerol al 10%) seguido de lavado con 5 volúmenes de tampón B (igual que el tampón A con sólo el 3% de contenido de glicerol). Las perlas se transfirieron a una columna de cromatografía Poly-Prep de 0,8 x 4 cm (BioRad; CA, EE.UU.) mediante resuspensión en 3 volúmenes de tampón B a 4°C, seguido por 10 volúmenes de lavado con tampón B. El lavado de las perlas se continuó mediante 10 volúmenes de tampón B, aumentados a KCl 300 mM. La columna se volvió a equilibrar después con tampón B regular. Para recuperar complejos asociados con ARNsi nativos, la columna se irradió en el cuarto frío, colocándola a una distancia de 2 cm rodeada por cuatro lámparas UV de 312 nm (tubo UV-B, 8 W, Herolab, Alemania) durante 30 minutos. Para recuperar la disolución siRNP foto-escindida, la columna se colocó en un tubo Falcon de 50 ml y se centrifugó a 2000 rpm durante 1 minuto a 4°C, utilizando de nuevo el rotor 2704. Para la recuperación completa de los siRNPs, las perlas se resuspendieron de nuevo en tampón B seguido de una segunda ronda de tratamiento UV durante 15 minutos. Los dos productos eluidos se reunieron y se ensayaron en cuanto a la degradación del ARN diana.

1.4 Ensayos de escisión de ARN diana

ARN diana de 177 nt, marcado en el remate, se generó como se ha descrito (Elbashir et al., EMBO J. 20 (2001 c), 6877-6888), excepto que su guanilil transferasa etiquetada con his se expresó en *E. coli* de un plásmido generosamente proporcionado por J. Wilusz y se purificó hasta la homogeneidad. Si no se indica lo contrario, ARNsi 5'-fosforilado o ARNsi dúplex fue pre-incubado en extracto HeLa S100 suplementado a 30°C durante 15 min antes de la adición de ARN diana marcado en el remate. Después de la adición de todos los componentes, las concentraciones finales fueron ARNsi 100 nM, ARN diana 10 nM, ATP 1 mM, GTP 0,2 mM, 10 U/ml de RNasin, 30 µg/ml de creatina quinasa, fosfato de creatina 25 mM, 50% de extracto S100. La incubación se continuó durante 2,5

h. La escisión de ARN diana mediada por ARNsi en lisado de embrión de *D. melanogaster* se realizó tal como se ha descrito (Zamore et al., Cell 101 (2000), 25-33). Se ensayó el RISC purificado por afinidad en tampón B en cuanto a la escisión del ARN diana sin preincubación ni adición de ARNsi extra (ARN diana 10 nM, ATP 1 mM, GTP 0,2 mM, 10 U/ml de RNasin, 30 µg/ml de creatina quinasa, fosfato de creatina 25 mM, 50% de RISC en tampón B). Las reacciones de escisión fueron detenidas mediante la adición de 8 vol. de tampón proteinasa K (Tris-HCl 200 mM, pH 7,5, EDTA 25 mM, NaCl 300 mM, SDS al 2% p/v). Proteinasa K, disuelta en Tris-HCl 50 mM pH 8,0, CaCl₂ 5 mM, glicerol al 50%, se añadió a una concentración final de 0,6 mg/ml y se procesó como se ha descrito (Zamore et al. (2000), supra). Las muestras se separaron en geles de secuenciación al 6%.

1.5 Filtración en Gel Analítico

Los productos eluidos UV en tampón B se fraccionaron mediante filtración en gel utilizando una columna Superdex 200 PC 3.2/30 (Amersham Biosciences) equilibrada con tampón A en un sistema SMART (Amersham Biosciences). El fraccionamiento se realizó utilizando un caudal de 40 µl/minuto y recogiendo fracciones de 100 µl. Las fracciones se ensayaron para determinar la escisión específica de ARN diana. La calibración del tamaño se realizó utilizando los marcadores de tamaño molecular tiroglobulina (669 kDa), ferritina (440 kDa), catalasa (232 kDa), aldolasa (158 kDa) y BSA (66 kDa) (Amersham Biosciences).

1.6 Sedimentación en gradiente de glicerol

Los productos eluidos UV se dispusieron en capas en la parte superior de 4 ml de un gradiente lineal de glicerol de 5% a 20% (p/p) ajustado a HEPES 30 mM, pH 7,4, KCl 100 mM, MgCl₂ 2 mM, DTT 0,5 mM. La centrifugación se realizó a 35000 rpm durante 14,5 h a 4°C utilizando un rotor Sorvall SW 60. Veinte fracciones de un volumen de 0,2 ml fueron separadas secuencialmente de la parte superior y se utilizaron partes alícuotas de 15 µl para el ensayo de escisión de ARN diana.

2. Resultados

2.1 Un sistema bioquímico humano para el análisis funcional de ARNsi

Los autores de la invención estaban interesados en ensayar la degradación del ARN diana mediada por ARNsi en extractos de células humanas, porque los ARNsi son reactivos de gran alcance para hacer descender la expresión de genes en células humanas, pero la acción de ARNsi en células humanas era incierto. Para investigar si los ARNsi guían la degradación de ARN diana en células humanas con un mecanismo similar al observado en *D. melanogaster* (p. ej., Elbashir et al. (2001 b), supra), prepararon sustratos para la degradación de ARNm fijado como objetivo tal como se ha descrito previamente (Elbashir et al. (2001 c), supra). Un transcrito de ARN de 177 nt, marcado en el remate con 5'-³²P, derivada de un segmento del gen de luciferasa de luciérnaga, se incubó en células HeLa S100 o extractos de embriones de *D. melanogaster* con un ARNsi dúplex de 21 nt en presencia de un sistema de regeneración de ATP (Figura 1A, B). Ensayos de escisión de ARNsi se realizaron a 25°C en lisado de *D. melanogaster* y a 30°C en extractos HeLa S100 durante 2,5 h. Después de la desproteinización utilizando proteinasa K, los productos de reacción se separaron en un gel de secuenciación al 6%.

De manera similar a la observación anterior en lisado de *D. melanogaster*, los autores de la invención observaron la aparición de un producto de escisión en el extracto HeLa S100 exactamente en la misma posición, lo que indica que el ARNsi dúplex guía la escisión de ARN diana en el sistema humano con la misma especificidad y mecanismo. La reacción de escisión apareció menos eficiente cuando se compara con el sistema de *D. melanogaster*, pero esto podría explicarse por la concentración 5 veces menor de proteína total de los extractos HeLa S100 (25 mg/ml frente a 5 mg/ml). De manera similar a lisados de *D. melanogaster*, ARNsi dúplex sin 5' fosfato fueron rápidamente 5' fosforilados en extractos HeLa S100 (véase más adelante) y la capacidad de escindir el ARN diana fue independiente de la presencia de un 5' fosfato en los ARNsi dúplex sintéticos.

Análisis comparativos de la eficiencia de los ARNsi dúplex de diferente longitud en lisado de *D. melanogaster* y en células de mamífero transfectadas indicó que las diferencias en las eficiencias de silenciamiento entre ARNsi dúplex de 20 a 25 nt eran menos pronunciadas en células de mamífero que en *D. melanogaster* (Elbashir et al. (2002), supra).

Dúplex de ARNsi de 24 y 25 nt eran inactivos en lisados de *D. melanogaster*, mientras que los mismos dúplex eran bastante eficaces para silenciar cuando se introducían por transfección en células HeLa. Por lo tanto, los autores de la invención se preguntaron si ARNsi dúplex de 20 a 25 nt son capaces de reconstituir RISC también con aproximadamente la misma eficiencia. De hecho, no se observaron grandes diferencias en el ensayo bioquímico, y la posición de la escisión del ARN diana era como se predijo de acuerdo con la escisión de reglas de guía establecidas en el lisado de *D. melanogaster* (datos no mostrados). Por lo tanto, los resultados bioquímicos respaldan las observaciones in vivo.

2.2 La modificación 5' del ARNsi guía inhibe la actividad del RISC

Modificación de ARNsi en sus extremos es importante para el desarrollo de esquemas de purificación de afinidad basados en ARNsi o para conjugar etiquetas de informador para mediciones biofísicas. El método más común para la introducción de cadenas laterales reactivas en los ácidos nucleicos es mediante síntesis utilizando derivados de aminoenzalizador (Eckstein (1991), *Oligonucleotides and analogues*, 2ª Ed., Oxford Reino Unido, Oxford University Press). Después de la desprotección completa del oligonucleótido, la amina primaria se hace reaccionar típicamente con el éster de N-hidroxisuccinimidilo del compuesto deseado. Los autores de la invención han introducido aminoenzaladores 5' y 3' con seis y siete grupos metileno como separadores, respectivamente. Los ARNsi dúplex modificados con enlazador se ensayaron para mediar la degradación del ARN diana en el extracto HeLa S100 (Figura 2A, B). La modificación del extremo 5' del ARNsi guía antisentido abolió la escisión de ARN, mientras que la modificación de sentido ni el extremo 5' ni de ambos extremos 3' mostró efecto inhibitorio alguno. En un experimento idéntico utilizando lisado de embrión de *D. melanogaster*, observaron un patrón similar de actividad de RISC aunque el dúplex que porta el ARNsi antisentido modificado en el aminoenzalizador 5' mostró alguna actividad residual (datos no mostrados). Presumiblemente, la introducción de átomos adicionales o el cambio en el fosfato terminal de carga eléctrica en el extremo 5' de la ARNsi antisentido interfieren con su capacidad de funcionar como ARNsi guía. La función crítica del extremo 5' del ARNsi guía fue previamente documentada (Elbashir et al. (2001 c), supra).

La capacidad de modificar ARNsi en su extremo 3' sugiere que los ARNsi no juegan un papel importante para cebar la síntesis de ARNs y no actúan como cebadores para la PCR degenerativa. El destino de un ARNsi en extractos HeLa S100 era seguido directamente por la incubación de un ARNsi dúplex radiomarcados internamente con ³²pCp. La cadena de ARNsi antisentido radiomarcada se preparó también con diferentes modificaciones 5' y 3' fosfato (Figura 3A). Todas las combinaciones descritas de ARNsi dúplex eran plenamente competentes para la degradación de ARN diana RISC-dependiente (datos no mostrados). Como se observó previamente para lisados de *D. melanogaster* (Nykänen et al. (2001), supra), era evidente la rápida 5'-fosforilación de ARNsi dúplex con terminales 5' hidroxilo libres. Para sorpresa de los autores de la invención, observaron que una pequeña fracción del ARNsi antisentido 3'-fosforilada podría ser ligada al 5'-hidroxilo opuesto del ARNsi sentido, produciendo una banda de menor movilidad. El ligamiento entre cadenas fue confirmado al cambiar la longitud del ARNsi sentido no marcado, lo que resultó en cambios de movilidad esperados del producto de ligamiento (datos no mostrados). La actividad de ARN ligasa se observó previamente en los extractos HeLa S100 y está mediada por dos actividades enzimáticas (p. ej., Vicente y Filipowicz, Eur. J. Biochem., 176 (1988), 431- 439). El fosfato 3' terminal se convierte primero en un fosfato 2',3'-cíclico que requiere ATP y adenilato fosfato 3' terminal. A continuación, el hidroxilo 5' opuesto se liga al extremo de fosfato cíclico mediante una ligasa de ARN todavía no caracterizada. Los autores de la invención sintetizaron químicamente el producto de ligamiento de 42 nt 5'-fosforilado predicho y encontraron que es incapaz de mediar en la escisión del ARN diana, presumiblemente debido a que no puede formar RISC activado. La mayoría de los ARNsi dúplex 3'-fosforilados se desfosforilaron gradualmente en su extremo 3' y emergieron de una forma químicamente similar a ARNsi generado de forma natural. En conjunto, estas observaciones indican que la célula tiene un mecanismo para preservar la integridad de ARNsi. Los autores de la invención fueron incapaces de detectar un producto de polimerización cebado con ARNsi propuesto (Figura 3B), lo que sugiere que los ARNsi no funcionan como cebadores para la síntesis de ARNs dependiente del molde en su sistema. Sin embargo, reconocieron que una actividad polimerasa dependiente de ARN propuesta puede haber sido inactivada durante la preparación de sus extractos.

2.3 ARNsis incorporados en RISC no compiten con una agrupación de ARNsi libres

Con el fin de analizar el ensamblaje y la estabilidad del RISC, los autores de la invención sometieron a ensayo si ARNsi dúplex inespecíficos para la diana fueron capaces de competir con ARNsi dúplex específicos para la diana. Cuando ARNsi dúplex específicos y no específicos se co-incubaron en extractos HeLa S100, concentraciones crecientes de ARNsi dúplex inespecífico compitieron con la formación de RISC específicos para la diana (Figura 4, pistas de la izquierda). Sin embargo, cuando ARNsis específicos para la diana se pre-incubaron en extracto HeLa S100 durante 15 min en ausencia de ARNsi dúplex competidor, el ARNsi ensamblado en el RISC específico para la diana ya no podía competir con el ARNsi dúplex no específico para la diana (Figura 4, pistas de la derecha). Este resultado sugiere que RISC se forma durante los primeros 15 minutos de incubación y que ARNsis se asocian irreversiblemente con los componentes proteicos del RISC durante la ventana de tiempo de 2,5 h del experimento.

2.4 Purificación del RISC humano

Después de tener los extremos 3' de ARNsi definidos como la posición más adecuada para la modificación química, un derivado de biotina foto-escindible se conjugó a los ARNsis modificados con 3'-aminoenzalizador. Un derivado de biotina foto-escindible se seleccionó debido a la ventaja de recuperar RISC bajo condiciones no desnaturalizantes después de capturar complejos en soportes de afinidad revestidos con estreptavidina. La conjugación 3' de biotina a las cadenas sentido, antisentido o ambas no afectó de escisión del ARN diana en comparación con ARNsi no biotinilados (datos no mostrados). Por lo tanto, se utilizaron ARNsi dúplex con residuos de biotina en ambos extremos 3' para la purificación por afinidad (Figura 5A). El ARNsi dúplex biotinilado se incubó en extractos HeLa S100 en presencia de ATP, GTP, fosfato de creatina y creatina quinasa para la regeneración de ATP. Después de

ello, se añadieron perlas de agarosa conjugadas con estreptavidina para capturar los complejos de ribonucleoproteína ARNsi biotinilados (siRNPs) incluyendo RISC. Después de un prolongado lavado de las perlas recogidas, los siRNPs se liberaron mediante irradiación UV a 312 nm. El producto eluido escindió el ARN diana de forma específica para la secuencia, lo que indica que RISC se recuperó en su estado nativo de la resina (Figura 5B, C, pista UV elu). El flujo a través de la selección por afinidad no mostró actividad detectable alguna, indicando una unión completa del RISC por parte de las perlas (Figura 6). El producto eluido por afinidad se analizó adicionalmente aplicándolo a una columna de filtración en gel Superdex 200 (Figura 5B), así como una ultra-centrifugación con un gradiente de glicerol de 5%-20% (Figura 5C). Las fracciones individuales se recogieron y se ensayaron en cuanto a la escisión de ARN diana sin la adición de ARNsi adicional. Actividad del RISC apareció entre la marcadores de tamaño molecular aldolasa (158 kDa) y BSA (66 kDa) después de la filtración en gel o centrifugación en gradiente de glicerol (Figura 5B, C). Por lo tanto, el tamaño molecular del RISC humano se estima entre 90 y 160 kDa, significativamente menor que el complejo previamente analizado en lisados de *D. melanogaster* (Hammond et al (2000), supra; Nykänen et al (2001), supra). El pequeño tamaño del RISC sugiere que Dicer (210 kDa) no está contenida en RISC y que la formación de RISC a partir de ARNsi sintéticos puede producirse independientemente de Dicer. Si bien estos resultados no descartan un papel para Dicer durante el ensamblaje del RISC, destacan la ausencia de Dicer en RISC.

2.5 RISC contiene un ARNsi de cadena sencilla y puede ser reconstituido mediante ARNsi de cadena sencilla

Se discuten actualmente dos modelos sobre la composición de la cadena de ARNsi de RISC. El primer modelo sugiere que las dos cadenas del ARNsi dúplex añadido inicialmente están presentes físicamente en el RISC, pero en una conformación desenrollada. El segundo modelo propone que RISC porta solamente una sola cadena de ARNsi, lo que implica la pérdida de una de las cadenas de ARNsi durante el ensamblaje. Este último modelo ha sido favorecido basado en la analogía con el procesamiento del precursor de ARNm, en donde sólo se acumuló una cadena de 21 nt de un precursor de horquilla ARNs. La base molecular para la asimetría de la reacción de procesamiento del precursor de ARNm todavía no se entiende. Debido a que ARNsi tiene colgantes 3' de 2 nt simétricos, se supone que ARNsi dúplex penetra en el RISC con igual probabilidad para ambas orientaciones, dando así origen a los distintos RSICs sentido y antisentido de fijación como objetivo.

Para hacer frente a la constitución de ARNsi en el RISC, los autores de la invención seleccionaron por afinidad los complejos ensamblados con ARNsi dúplex que se biotinilaron en sólo una de las dos cadenas constituyentes o en ambas (Figura 6). Si las dos cadenas estaban presentes juntas en el RISC, la actividad de escisión se debe seleccionar por afinidad en Neutravidina independientemente de la posición del residuo biotina. En contraposición, los autores de la invención observaron una escisión del ARN diana de productos eluidos de UV después de la selección de estreptavidina sólo para dúplex de ARNsi con biotina conjugado a la cadena antisentido, pero no a la cadena sentido (Figura 6). La actividad de RISC, ensamblado en ARNsi dúplex con sólo el ARNsi sentido biotinilado, permaneció en el flujo no detenido. Estos datos sugieren que el RISC contiene sólo una molécula de ARN de cadena sencilla.

Para evaluar si ARNsi de cadena sencilla pueden ser capaces de reconstituir RISC, ARNsi de cadena sencilla 5'-fosforilados así como el ARNsi dúplex se incubaron a concentraciones entre 1 y 100 nM con ARN diana marcado en el remate en extracto HeLa S100 (Figura 7A). A ARNsi antisentido de cadena sencilla 100 nM, los autores de la invención detectaron una escisión de ARN diana específica para RISC, lo que confirma que los ARNsi de cadena sencilla están presentes en el RISC. A concentraciones más bajas de ARNsi de cadena sencilla, la formación de RISC permaneció indetectable mientras que ARNsi dúplex estaban formando efectivamente el RISC incluso a una concentración de 1 nM. Por lo tanto, existe una vía específica que convierte ARNsi de doble cadena en ARNsi de cadena sencilla que contiene RISC. Utilizando lisado de embrión de *D. melanogaster*, los autores de la invención no han podido detectar la actividad de RISC de ARNsi antisentido (Figura 7B), presumiblemente debido a la alta carga de ribonucleasas específicas de cadena sencilla (Elbashir et al. (2001 b), supra). Además, si ARNsi antisentido de 20 a 25 nt 5'-fosforilado eran capaces de mediar en la degradación de ARN diana específico para RISC en extracto HeLa S100, produciendo los mismos sitios de escisión de ARN diana que los ARNsi dúplex de esta longitud (datos no mostrados).

Finalmente, los autores de la invención sometieron a ensayo ARNsi de cadena sencilla y dúplex para fijar como objetivo un gen endógeno en células HeLa siguiendo su protocolo estándar previamente establecido para el silenciamiento de lamina A/C. Concentraciones 200 nM de ARNsi de cadena sencilla con y sin 5'-fosfato y concentraciones 100 nM de ARNsi dúplex se transfectaron a células HeLa. Los niveles de lamina A/C fueron monitorizados 48 h más tarde utilizando inmunofluorescencia (Figura 8A) y análisis de transferencia Western cuantitativo basado en la luminiscencia (Figura 8B). ARNsi antisentido no fosforilado provocó un descenso sustancial de lamina A/C a aproximadamente al 25% de su nivel normal, mientras que ARNsi 5'-fosforilado redujo el contenido de lamina A/C a menos de 5%, similar a la reducción observada con la lamina A/C 5'-fosforilada (datos no mostrados) o ARNsi dúplex no fosforilado (Figura 8). ARNsi sentido y ARNsi inespecífico para GL2 no afectaron a los niveles de lamina A/C. Los niveles de proteína vimentina no específica se monitorizaron y se utilizaron para la normalización de la carga de las pistas de transferencia Western de lamina A/C.

El silenciamiento génico también se observó con ARNsi antisentido fosforilados así como no fosforilados que varían en tamaño entre 19 y 29 nt. Los ARNsis antisentido fosforilados se comportaban consistentemente mejor que los antisentido no fosforilados, y sus eficiencias de silenciamiento eran equiparables a la del ARNsi dúplex convencional (Figura 11).

5 2.6 Composición de proteínas del RISC

Con el fin de identificar los componentes de las proteínas del complejo de silenciamiento inducido por ARN (RISC) en el extracto HeLa S100, se utilizó la selección por afinidad específica previamente descrita. Productos eluidos de UV se fraccionaron en un gradiente de glicerol de 5-20%, las fracciones se recuperaron del gradiente y se analizaron en cuanto a la composición de proteínas y la actividad endonucleolítica del ARN diana.

- 10 Dos proteínas de aproximadamente 100 kDa se identificaron por espectrometría de masas en la fracción pico de la actividad endonucleolítica (Figura 12, fracciones 7 y 8), correspondiente a eIF2C1 y eIF2C2/GERp95 (Figura 13A y B). Estas proteínas son un 82% similares y son miembros de la familia Argonaute (Figura 13C). La primera evidencia de que las proteínas Argonaute son parte del RISC fue proporcionada por los estudios de fraccionamiento bioquímicos clásicos utilizando células S2 de *D. melanogaster* transfectadas con ARNs (Hammond et al., 2001, supra). Se demostró recientemente que el pariente más cercano a *D. melanogaster* Argonaute2, *D. melanogaster* Argonaute1, era necesarios para el ARNi (Williams y Rubin, PNAS USA 99 (2002), 6889-6894).

El análisis por espectrometría de masas también reveló la presencia de tres péptidos que pertenecen exclusivamente al miembro HILI de la familia Argonaute de proteínas. Las secuencias de esos péptidos son: NKQDFMDLSICTR, correspondiente a las posiciones 17-29 de la proteína; TEYVAESFLNCLRR, correspondiente a las posiciones 436-449 de la proteína, y; YNHDLPARIIYR, correspondiente a las posiciones 591-603 de la proteína. Este hallazgo sugiere que la proteína HILI también puede ser parte de RISC.

En seres humanos, la familia Argonaute está compuesta por 6 miembros, eIF2C1, eIF2C2, eIF2C3, eIF2C4, HILI y HIWI (Figura 14). El alineamiento de las seis secuencias de aminoácidos predichas muestra una alta conservación, en particular entre los miembros eIF2C y HILI y HIWI (Figura 15). También se muestran secuencias de ADNc predichas que codifican las proteínas Argonaute (Figura 16).

La expresión de las proteínas Argonaute humanas también se investigó en células HeLa mediante análisis de RT-PCR utilizando (A) ARN total y poli seleccionado. Se detectaron todos los miembros de la familia, excepto HIWI (Figura 17).

3. Discusión

- 30 El desarrollo de un sistema bioquímico humano para el análisis del mecanismo de ARNi es importante dado el reciente éxito de ARNsi dúplex para silenciar genes expresados en células cultivadas humanas y el potencial de convertirse en un agente terapéutico específico para la secuencia. Sistemas bioquímicos son útiles para definir las etapas individuales del proceso de ARNi y para evaluar los requisitos de constitución y moleculares de los complejos macromoleculares que participan. Para el análisis de ARNi, se han desarrollado varios sistemas, siendo los sistemas de *D. melanogaster* los más completos, ya que permiten reconstituir el procesamiento de ARNs, así como la fijación como objetivo del ARNm. Para los sistemas de mamíferos, todavía no se ha logrado la reconstitución de la reacción de fijación como objetivo del ARNm. Aquí, se describe el desarrollo y la aplicación de un sistema bioquímico preparado a partir de la fracción citoplasmática de células HeLa humanas, que es capaz de reconstituir el complejo de silenciamiento inducido por ARN (RISC) que fija como objetivo ARNm humano. La formación de RISC se logró utilizando ARNsi dúplex 5'-fosforilados o no fosforilados; así como ARNsi antisentido de cadena sencilla; ARNsi dúplex no fosforilados y presumiblemente también ARNsi antisentido de cadena sencilla son rápidamente 5'-fosforilados en extractos de células HeLa (Figura 3).

Caracterización bioquímica de la función de ARNsi

- 45 La reconstitución de la actividad de RISC sólo se observó utilizando extractos citoplasmáticos HeLa. Extractos nucleares HeLa ensayados bajo las mismas condiciones no sustentaron la escisión de ARN diana específica para ARNsi, lo que sugiere que los componentes de RISC se encuentran predominantemente en el citoplasma (datos no mostrados).

- 50 Las modificaciones de los extremos 5' y 3' de ARNsi se sometieron a ensayo con el fin de evaluar la importancia de los extremos de ARNsi para la etapa de fijación de objetivo. Se encontró que la modificación en el extremo 5' del ARNsi guía era más inhibidora de la escisión de ARN diana que la modificación en el extremo 3'. La introducción por afinidad de la etiqueta tag de biotina 3' en el ARNsi permitió a los autores de la invención seleccionar por afinidad el RISC que fija como objetivo ARN sentido, mientras la biotinilación 3' de la cadena de ARNsi sentido dio como resultado la actividad RISC en el flujo retenido. Además, la composición de la cadena de ARN sencilla de RISC se confirmó mediante la reconstitución del complejo de endonucleasa específica para la secuencia utilizando ARNsi guía de cadena sencilla 5'-fosforilado. La reconstitución de RISC a partir de ARNsi de cadena sencilla era, sin

embargo, menos eficaz y requería concentraciones 10 a 100 veces mayores en comparación con ARNsi de doble cadena. La Reconstitución de RISC a partir de ARNsi de cadena sencilla era indetectable utilizando lisado de embrión de *D. melanogaster*, lo que se explica más probablemente por el alto contenido de 5' a 3' exonucleasas en el lisado de embrión.

- 5 El tamaño de RISC en el lisado HeLa se determinó por filtración en gel, así como ultracentrifugación en gradiente de glicerol después de purificación por afinidad con estreptavidina con ARNsi dúplex 3' biotinilado. Tamaños de RISC en sistemas de *D. melanogaster* han sido reseñados dentro de un intervalo de menos de 230 a 500 kDa, sin embargo se llevaron a cabo determinaciones de tamaño sin tener un RISC purificado por afinidad. Los sedimentos de RISC purificados por afinidad de los autores de la invención en un intervalo estrecho entre los marcadores de tamaño de 66 y 158 kDa. Las diferencias a los tamaños reseñados para RISC no son específicas para la especie, ya que los autores de la invención observaron un tamaño similar para RISC en extractos citoplasmáticos de células S2 después de la purificación por afinidad (datos no mostrados).

- 15 Previamente se ha propuesto que los ARNsi actúan como cebadores para la síntesis de ARNs con molde de ARN diana (Lipardi et al., Cell 107 (2001), 297-307), aunque los homólogos para tales ARN polimerasas dependientes de ARN conocidos por participar en el silenciamiento de genes en otros sistemas no se identifican en genomas de *D. melanogaster* o de mamíferos. El análisis del destino de ARNsi dúplex en el sistema de células HeLa no proporcionó evidencia de una actividad cebada de ARNsi de este tipo (Figura 3), pero indica que la vía predominante para el silenciamiento de los genes mediado por ARNsi es la degradación del ARN diana endonucleolítico específico para la secuencia.

- 20 ARNsi antisentido 5'-fosforilado de cadena sencilla como desencadenantes del silenciamiento de genes de mamíferos

- Previamente se señaló que la introducción de ARNs sentido y antisentido de varios cientos de nucleótidos de longitud en *C. elegans* fue capaz de silenciar de forma específica para la secuencia genes homólogos (Guo y Kemphues, Celular 81 (1995), 611-620). Más tarde, se sugirió que la preparación de ARN sentido y antisentido estaba contaminada con una pequeña cantidad de ARNs, que era responsable del efecto de silenciamiento y es un inductor mucho más potente del silenciamiento de genes (Fire et al. (1998), supra). Sin embargo, es concebible que el ARN antisentido contribuya directamente en el inicio del silenciamiento. De hecho, recientemente se demostró que ARNs antisentido entre 22 y 40 nt, pero no ARNs sentido eran capaces de activar el silenciamiento de genes en *C. elegans* (Tijsterman et al., Science 295 (2002), 694-697). Los autores, sin embargo, están a favor de la hipótesis de la síntesis de ARNs cebada con ARNsi.

- Los autores de la invención han demostrado que la modificación de los extremos 3' de ARNsi antisentido no interfería en la reconstitución de RISC en el sistema humano. En conjunto, estas observaciones sugieren que las fuerzas motrices para el silenciamiento de genes en *C. elegans* pueden ser predominantemente la síntesis de ARNs, seguida por la escisión de Dicer, mientras que en el ser humano y, posiblemente, también en *D. melanogaster* predomina la degradación de ARNm diana específica para RISC.

- La fijación como objetivo de lamina A/C expresada de forma endógena por transfección de ARNsi dúplex en células HeLa fue el primer ejemplo reseñado del silenciamiento de genes inducido por ARNsi. La proteína lamina A/C se redujo drásticamente mediante ARNsi dúplex específico para lamina A/C en el espacio de dos días después de la transfección, mientras que el ARNsi dúplex inespecífico no mostró efecto alguno. En el momento, la transfección de ARNsi sentido o antisentido no fosforilado no reveló un efecto sustancial sobre los niveles de lamina A/C, aunque más recientemente se observó una pequeña reducción tras la transfección de ARNsi antisentido cuando se suministraron concentraciones de ARNsi antisentido similares tal como se describe en este estudio. Sin embargo, el efecto no se interpretó como un efecto específico de RISC. El ensayo de ARNsi antisentido 5'-fosforilado reveló un incremento sustancial en el silenciamiento de lamina A/C. Probablemente los ARNsi 5'-fosforilados son más estables o penetran más rápidamente en el RISC. Alternativamente, el extremo 5' de ARNsi de cadena sencilla transfectado puede ser fosforilado menos rápidamente en la célula en comparación con ARNsi dúplex.

- Finalmente, debe señalarse que las células HeLa son generalmente deficientes en nucleasas y representan uno de los sistemas preferidos de mamíferos para estudiar el procesamiento de ARN o reacciones de transcripción in vivo e in vitro. Sin embargo, se puede esperar que los ARNsi antisentido de cadena sencilla 5'-fosforilados sean adecuados para hacer descender la expresión de los genes en otros tipos de células o tejidos con un contenido diferente de nucleasas, ya que están disponibles estrategias químicas para mejorar la resistencia a las nucleasas de ARN de cadena sencilla. La facilidad de silenciamiento general de diversos tipos de células también puede depender de los niveles relativos de proteínas eIF2C1 e eIF2C2 libres de ARNsi/ARNmi, capaces de asociarse con ARNsi suministrado de forma exógena.

- 55 En resumen, los ARNsi antisentido 5'-fosforilados de cadena sencilla de un tamaño de 19 a 29 nt amplían el uso de moléculas de ARN para el silenciamiento de los genes, ya que pueden penetrar en la vía del ARNi de mamíferos in vitro así como in vivo a través de la reconstitución de RISC. eIF2C1 y/o eIF2C2 humanos parecen jugar un papel crítico en el proceso. Considerando la posibilidad de modular la estabilidad y las propiedades de captación del ARN

de cadena sencilla, ARNs antisentido de cadena sencilla, 5'-fosforilados, pueden expandir adicionalmente la utilidad de la tecnología del silenciamiento de genes basada en ARNi como herramienta para la genómica funcional así como aplicaciones terapéuticas.

- 5 Proteínas Argonaute son una clase distinta de proteínas que contiene un dominio PAZ y Piwi (Cerutti et al., 2000, supra) y han estado implicadas en muchos procesos previamente ligados al silenciamiento post-transcripcional, pero sólo hay disponible una información bioquímica limitada.

- 10 eIF2C2 humana es el ortólogo de GERp95 de rata, que se identificó como un componente del complejo de Golgi o del retículo endoplásmico y se co-purificó con membranas intracelulares (Cikaluk et al., Mol. Biol. Cell 10 (1999), 3357-3722). Más recientemente, eIF2C2 de células HeLa demostró estar asociada con microARNs y componentes del complejo SMN, un regulador del ensamblaje de ribonucleoproteínas, sugiriendo que eIF2C2 juega un papel en el procesamiento del precursor ARNm o en la función de ARNm (Mourelatos et al., Genes & Dev. 16 (2002), 720-728). Una hipótesis más provocativa es que los ARNm están también en un complejo de tipo RISC, que potencialmente podría mediar en la degradación de ARN diana, si sólo existieran ARNm diana de ARNm perfectamente emparejados. Sin embargo, el análisis de la secuencia utilizando seres humanos y ratones clonados, no reveló la presencia de secuencias perfectamente complementarias entre los genomas (Lagos-Quintana et al., Science 294 (2001), 853-858). Por lo tanto, miRNPs pueden sólo funcionar como reguladores de la traducción del ARNm diana parcialmente emparejados de forma errónea, probablemente al reclutar factores adicionales que evitan la disociación de ARNm diana erróneamente emparejados.

- 20 eIF2C1 humana no ha sido ligada previamente al silenciamiento de genes, pero es más de un 80% similar en la secuencia a eIF2C2. (Koesters et al., Genomics 61 (1999), 210-218). Esta similitud puede indicar una redundancia funcional pero también es concebible que el RISC funcional pueda contener heterodímeros de eIF2C1 y eIF2C2. El peso molecular predicho de este complejo heterodimérico sería ligeramente mayor que el tamaño observado de 90-160 kDa, pero dado que el fraccionamiento del tamaño se basa en una forma globular, los autores de la invención no pueden desechar esta posibilidad en este momento.

- 25 Debido a la elevada conservación entre los miembros de la familia Argonaute, es posible que péptidos que se derivan de regiones 100% conservadas en las 6 proteínas predichas pueden pertenecer a miembros distintos de eIF2C1 y eIF2C2. A este respecto, se identificaron tres péptidos con masas que corresponden a HILI, lo que significa que esta proteína también podría ser un componente de RISC.

- 30 Para evaluar con precisión la composición de la proteína de RISC, se debe conseguir una reconstitución de la escisión de ARN diana mediada por ARNs utilizando proteínas recombinantes que pueden obtenerse mediante la clonación y expresión en sistemas bacterianos o eucarióticos adecuados.

- 35 Los autores de la invención esperan que la caracterización bioquímica o el proceso de degradación de ARN diana mediada por ARNs tendrán aplicaciones inmediatas tales como el desarrollo de líneas celulares o animales transgénicos que sobre-expresan componentes del RISC. La eficacia en fijar como objetivo genes endógenos en esas líneas u organismos será potenciada. Además de ello, un sistema in vitro reconstituido para el ARNi permitirá el diseño de ARNs más potente y específico para conseguir el silenciamiento de genes.

REIVINDICACIONES

1. Una molécula de ARN de cadena sencilla que tiene una longitud de 14-50 nucleótidos, en donde al menos los 14-20 nucleótidos más próximos a 5' son sustancialmente complementarios a un transcrito diana para uso en terapia, en donde la molécula de ARN tiene un análogo de fosfato en el extremo 5' que se selecciona del 5'-difosfato, 5'-trifosfato, remate de 5'-guanosina 7-metilado o no metilado, remate de 5'-adenosina, cualquier estructura de remate de nucleótidos no modificado N-O-5'-(HO)(O)P-O-(HO)(O)P-OP(HO)(O)-O-5', 5'-monotiofosfato, 5'-monoditiofosfato, 5'-fósforotiolato, cualquier combinación adicional de monofosfato, difosfato y trifosfatos reemplazados con oxígeno/azufre, 5'-fósforoamidatos, 5'-alquilfosfonatos y 5'-alquiléterfosfonatos, en donde la molécula del ARN de cadena sencilla es capaz de inhibir la expresión de transcritos diana mediante ARN-interferencia.
2. La molécula de ARN de cadena sencilla para uso de la reivindicación 1, en donde dicha molécula de ARN tiene una longitud de 15-29 nucleótidos.
3. La molécula de ARN de cadena sencilla para uso de una cualquiera de las reivindicaciones 1 a 2, en donde dicha molécula de ARN es completamente complementaria a dicho transcrito diana opcionalmente con excepción de nucleótidos que se extienden más allá de la posición 20 contada desde el extremo 5'.
4. La molécula de ARN de cadena sencilla para uso de una cualquiera de las reivindicaciones 1 a 3, en donde dicha molécula de ARN comprende al menos un análogo de nucleótidos modificado.
5. La molécula de ARN de cadena sencilla para uso de la reivindicación 4, en donde los análogos de nucleótidos modificados se seleccionan de ribonucleótidos modificados en la cadena principal de azúcares y modificados en nucleobases y combinaciones de los mismos.
6. La molécula de ARN de cadena sencilla para uso de una cualquiera de las reivindicaciones 1 a 5, en donde la inhibición de la expresión del gen diana es *in vitro*.
7. La molécula de ARN de cadena sencilla para uso de una cualquiera de las reivindicaciones 1 a 5, en donde la inhibición de la expresión del gen diana es *in vivo*.
8. La molécula de ARN de cadena sencilla para uso de una cualquiera de las reivindicaciones 1 a 7, en donde dicha molécula de ARN está formulada como una composición farmacéutica que contiene un soporte farmacéuticamente aceptable.
9. La molécula de ARN de cadena sencilla para uso de la reivindicación 8, en donde dicho soporte se selecciona de liposomas catiónicos y formulaciones de lípidos catiónicas.
10. La molécula de ARN de cadena sencilla para uso de una cualquiera de las reivindicaciones 1 a 9, en donde dicha molécula de ARN está asociada con polímeros o micropartículas biodegradables.
11. La molécula de ARN de cadena sencilla para uso de la reivindicación 10, en donde dicha asociación comprende un acoplamiento covalente.
12. La molécula de ARN de cadena sencilla para uso de la reivindicación 11, en donde dicho acoplamiento covalente se produce a través del extremo 3' de la molécula de ARN.
13. La molécula de ARN de cadena sencilla para uso de una cualquiera de las reivindicaciones 1 a 12, en donde las enfermedades se seleccionan de enfermedades tumorales, enfermedades inflamatorias, enfermedades infecciosas, enfermedades degenerativas y enfermedades autoinmunes.
14. Un método *in vitro* para inhibir la expresión de un transcrito diana en células de mamífero mediante ARN-interferencia, que comprende las etapas de:
 - (a) poner en contacto dichas células con una molécula de ARN de cadena sencilla que tiene una longitud de 14-50 nucleótidos, en donde al menos los 14-20 nucleótidos más próximos a 5' son sustancialmente complementarios a dicho transcrito diana, en donde la molécula de ARN tiene un análogo de fosfato en el extremo 5' que se selecciona del 5'-difosfato, 5'-trifosfato, remate de 5'-guanosina 7-metilado o no metilado, remate de 5'-adenosina, cualquier estructura de remate de nucleótidos modificado o no modificado N-O-5'-(HO)(O)P-O-(HO)(O)P-OP(HO)(O)-O-5', 5'-monotiofosfato, 5'-monoditiofosfato, 5'-fósforotiolato, cualquier combinación adicional de monofosfato, difosfato y trifosfatos reemplazados con oxígeno/azufre, 5'-fósforoamidatos, 5'-alquilfosfonatos y 5'-alquiléterfosfonatos, y
 - (b) inhibir la expresión de un transcrito diana mediante ARN-interferencia efectuado por el ARN de cadena sencilla hacia un ácido nucleico diana que tiene una porción de secuencia sustancialmente complementaria al ARN de cadena sencilla.

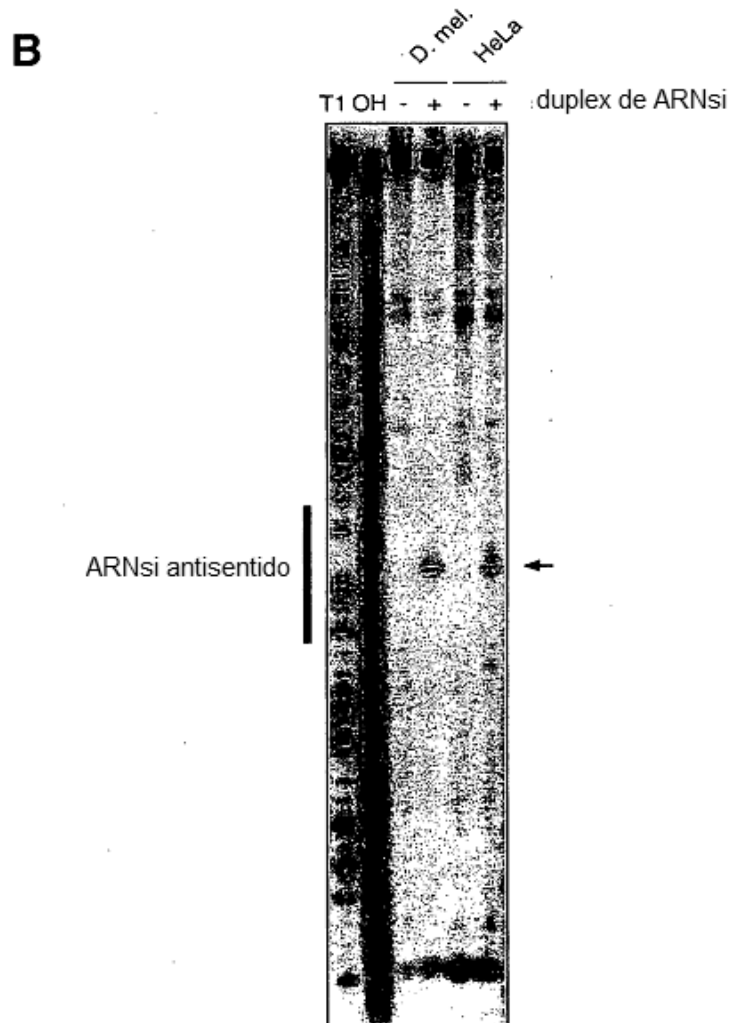
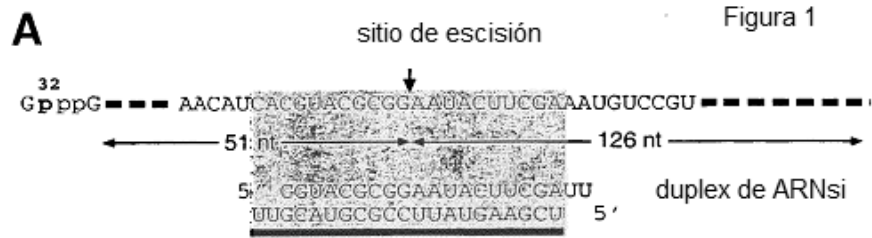


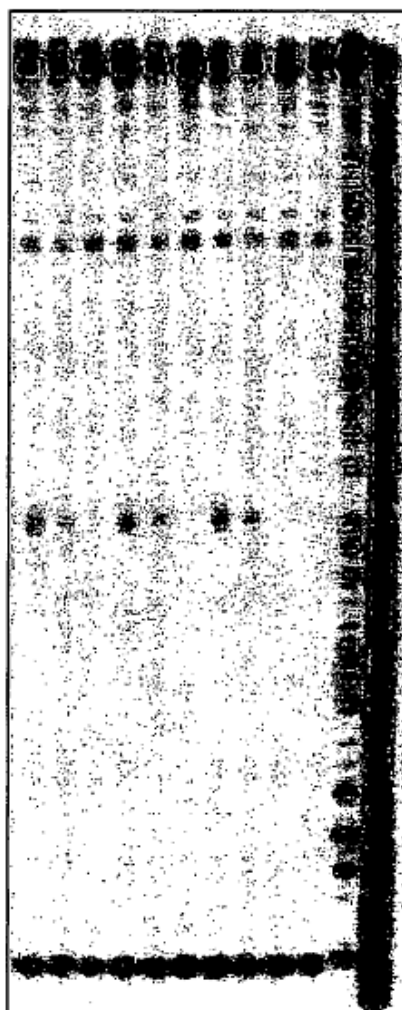
Figura 2

A



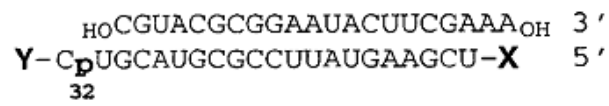
B

s 3' 3' 3' - - - 5' 5' 5'
as 3' - 5' 3' - 5' 3' - 5' NCT1OH



A

Figura 3



B

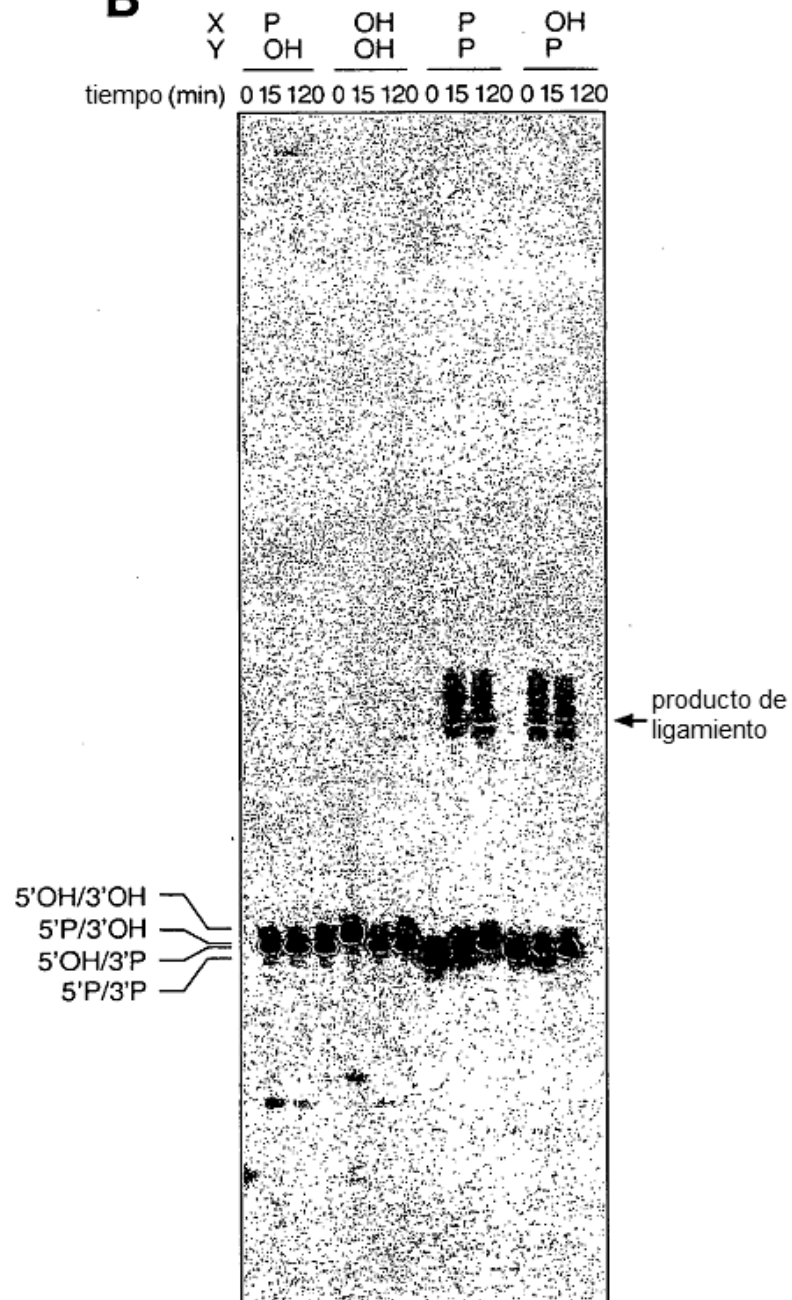
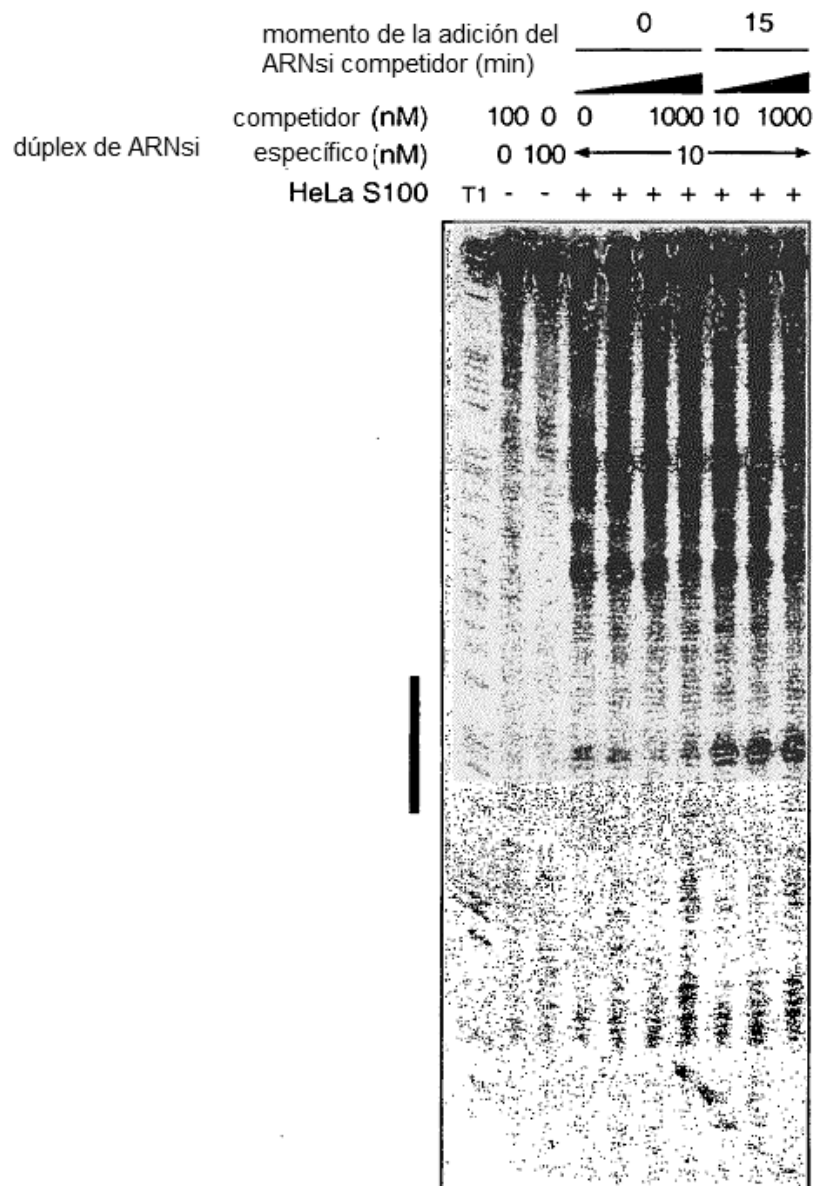


Figura 4



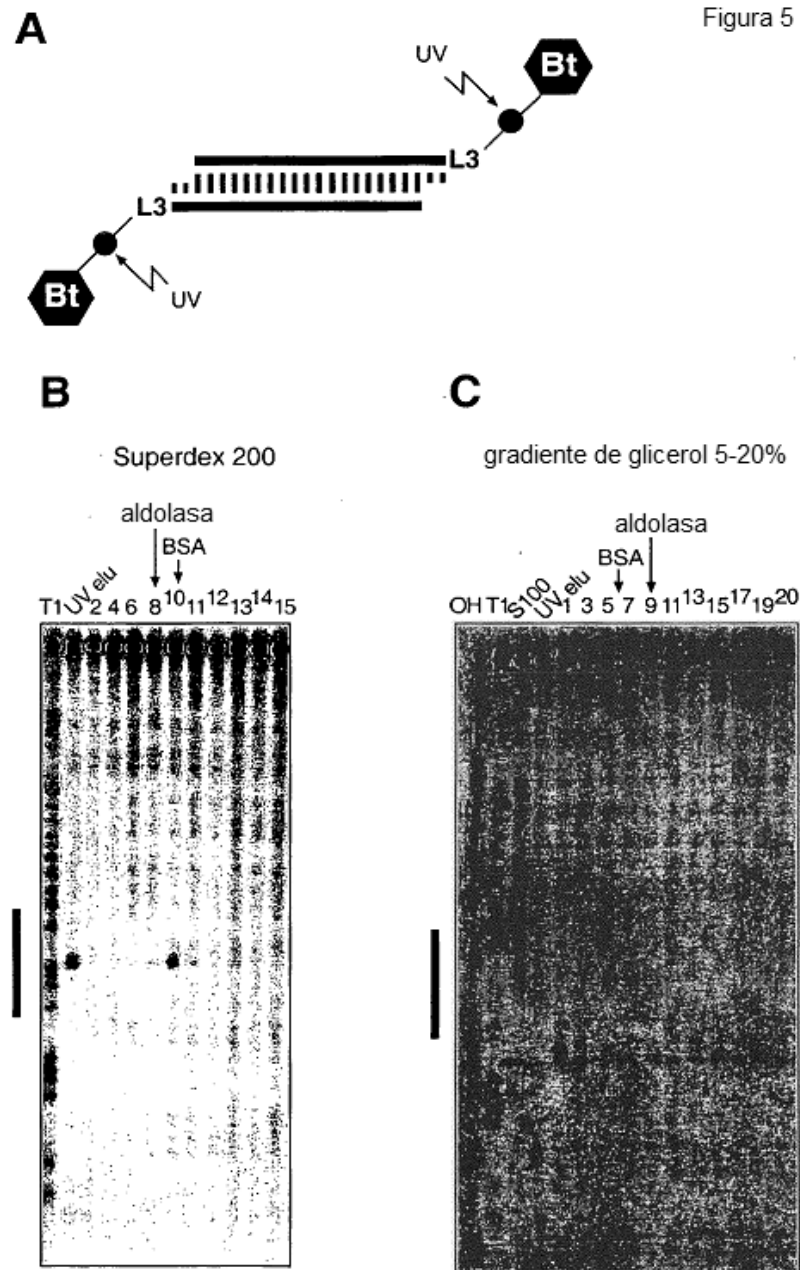


Figura 6

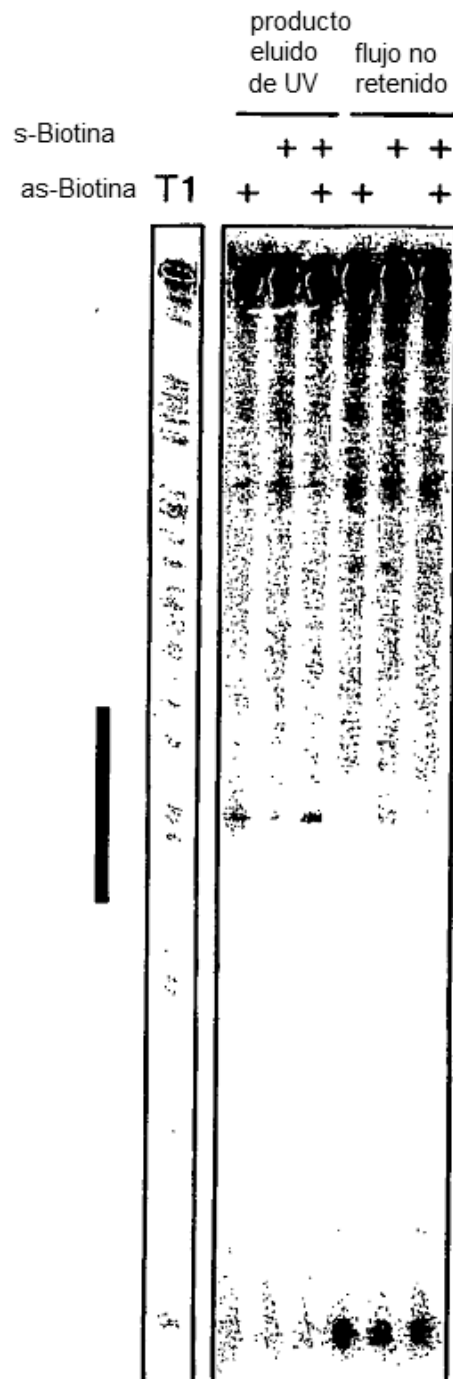
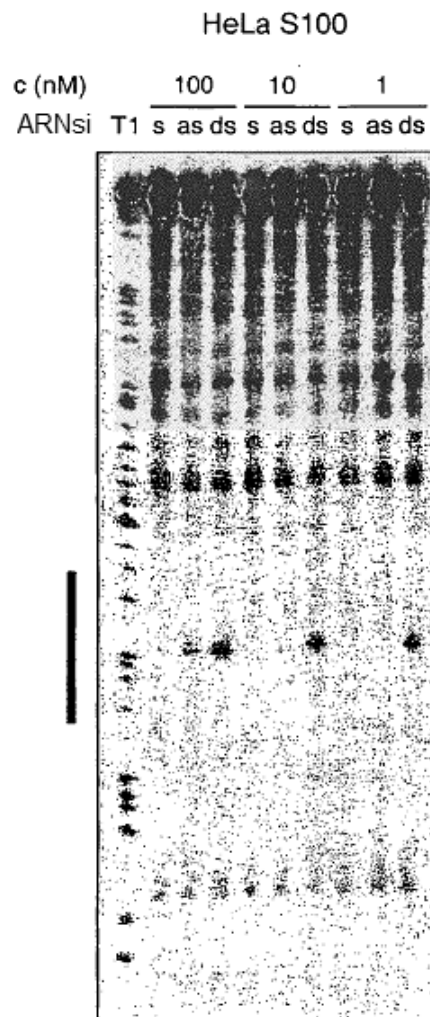


Figura 7

A



B

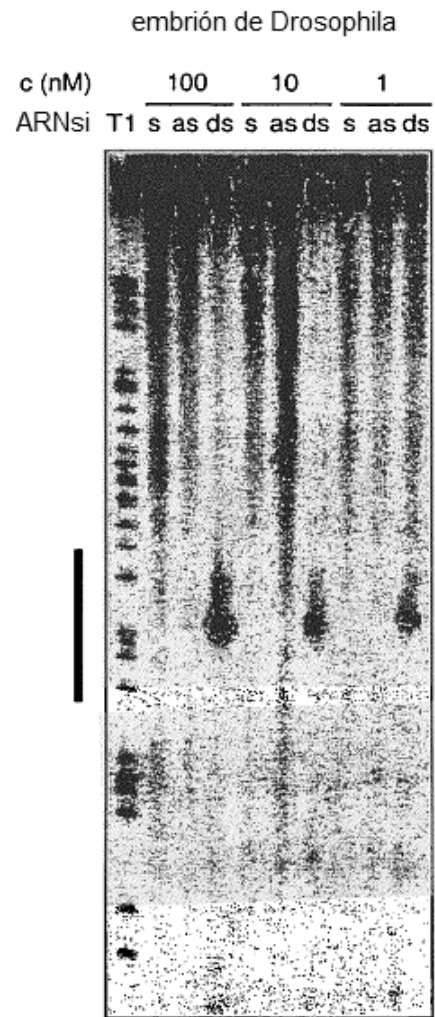


Figura 8

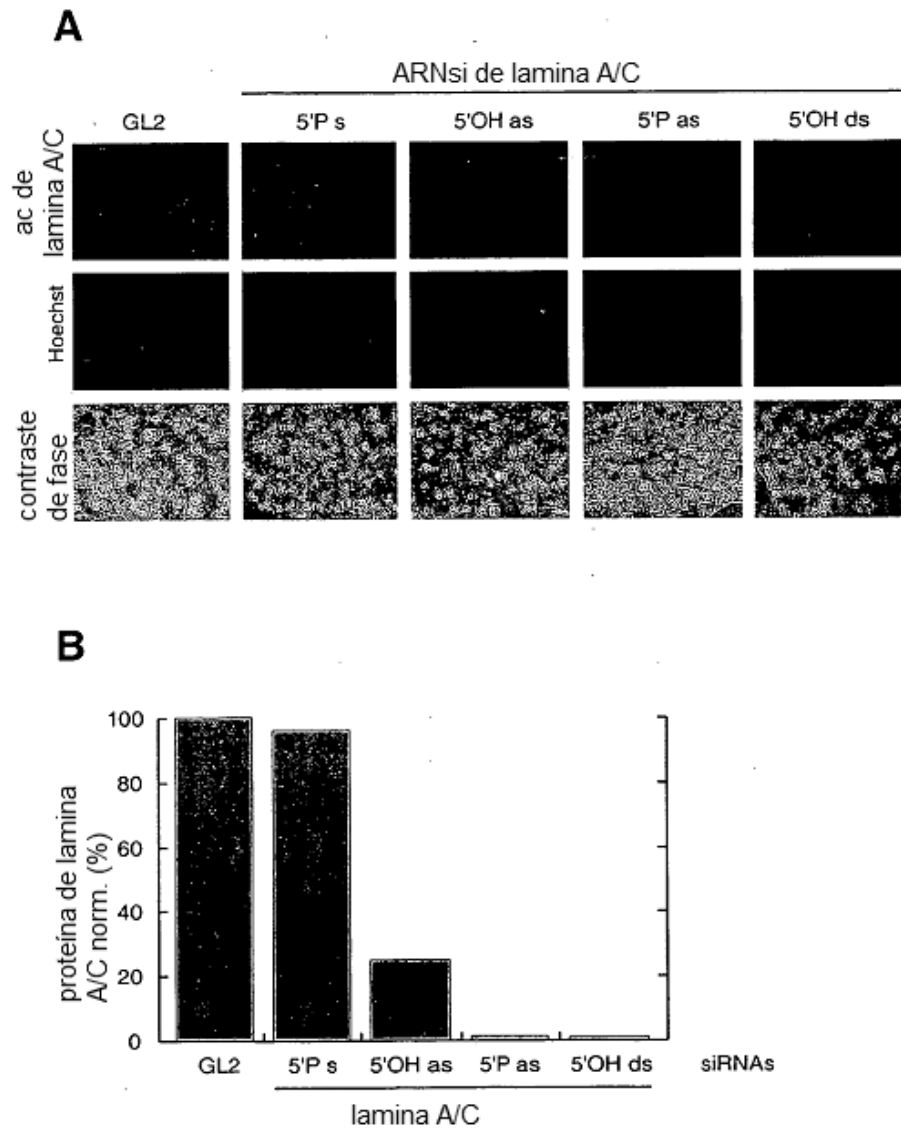
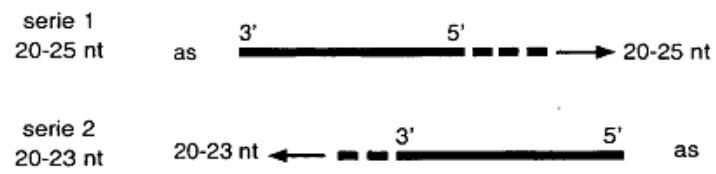


Figura 9

A



B

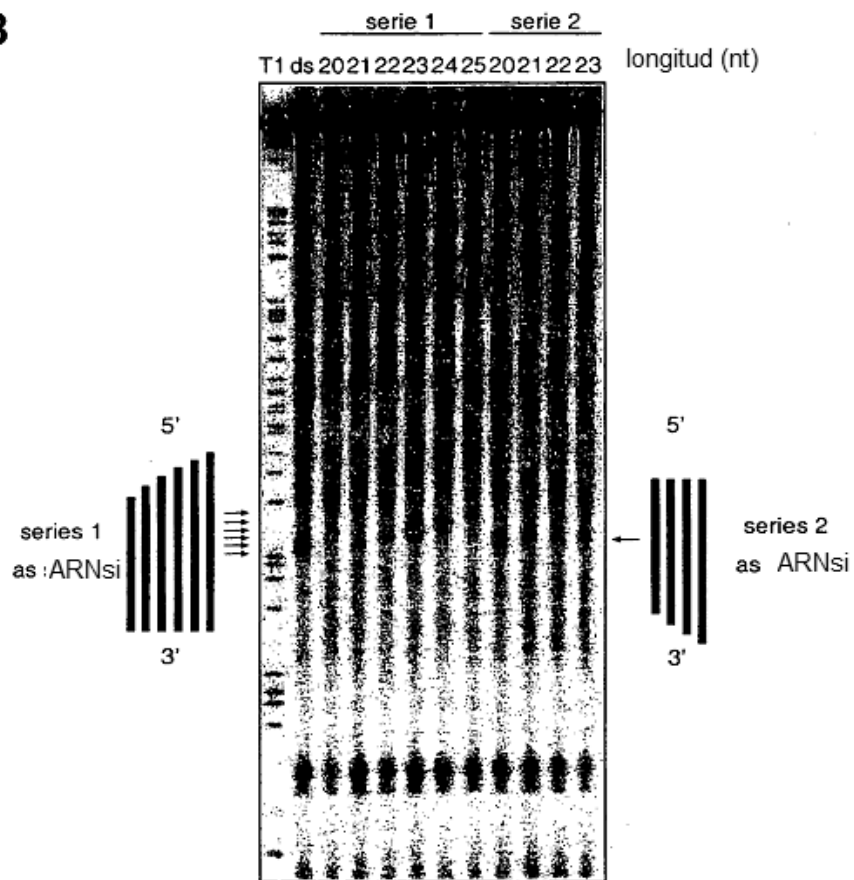


Figura 10

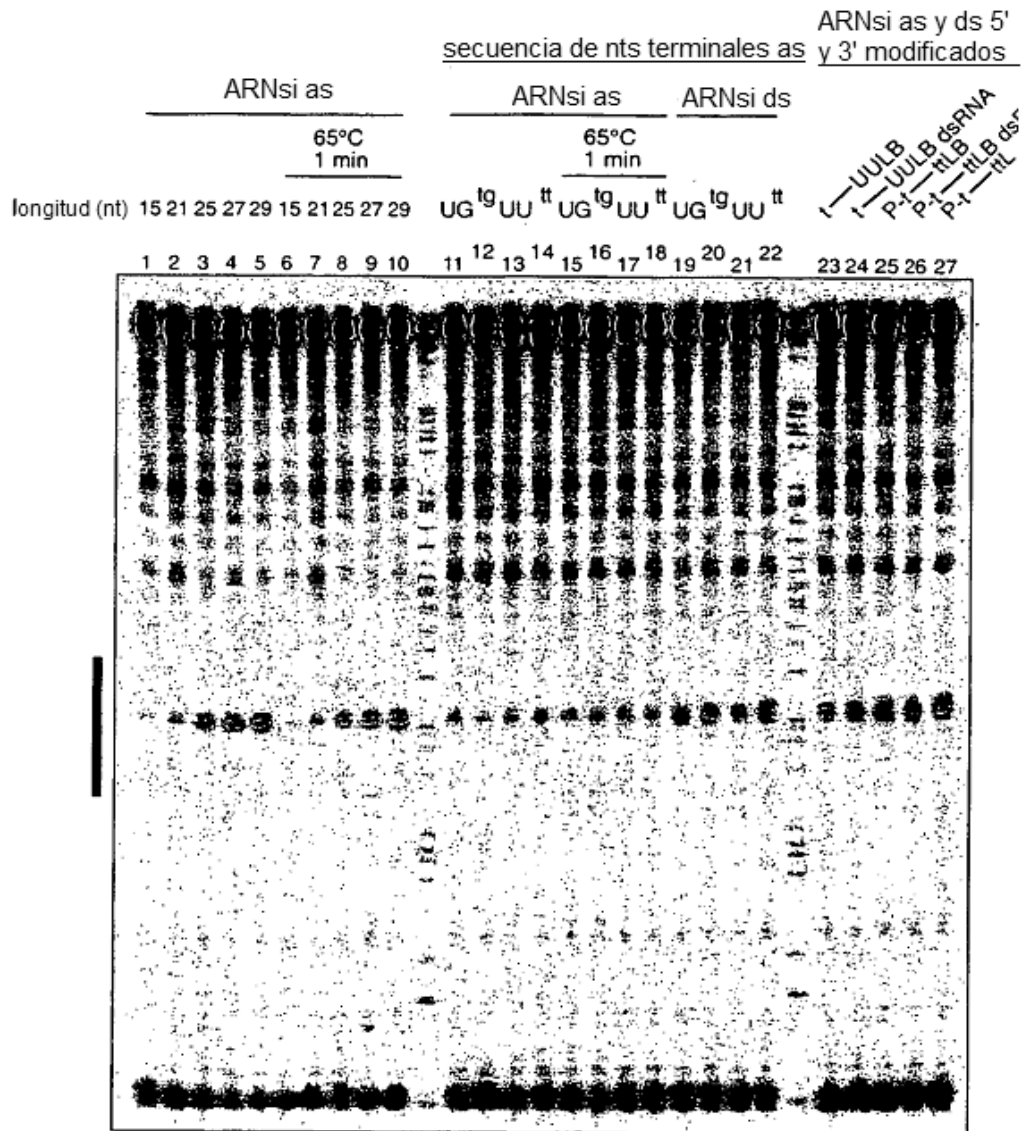


Figura 11

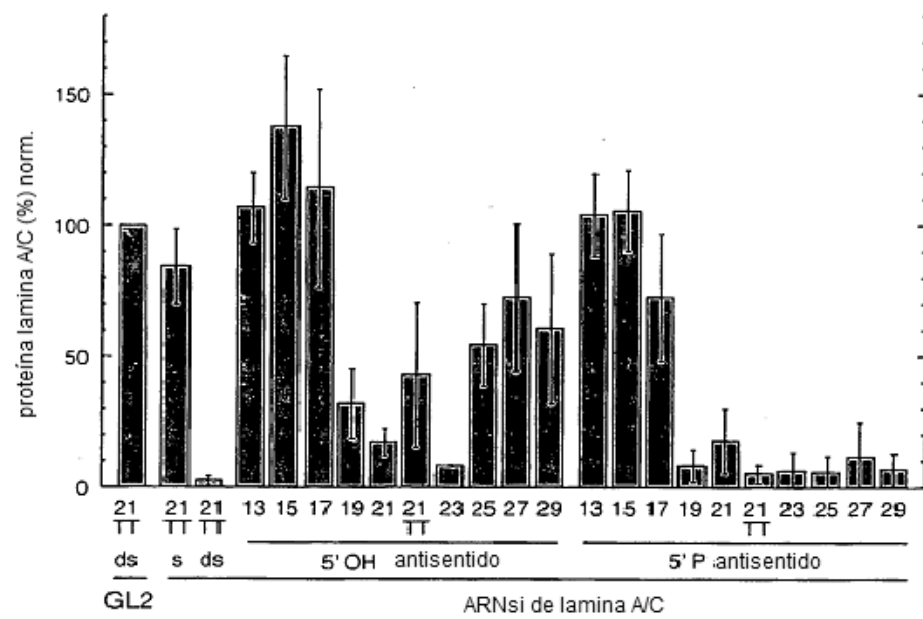


Figura 12

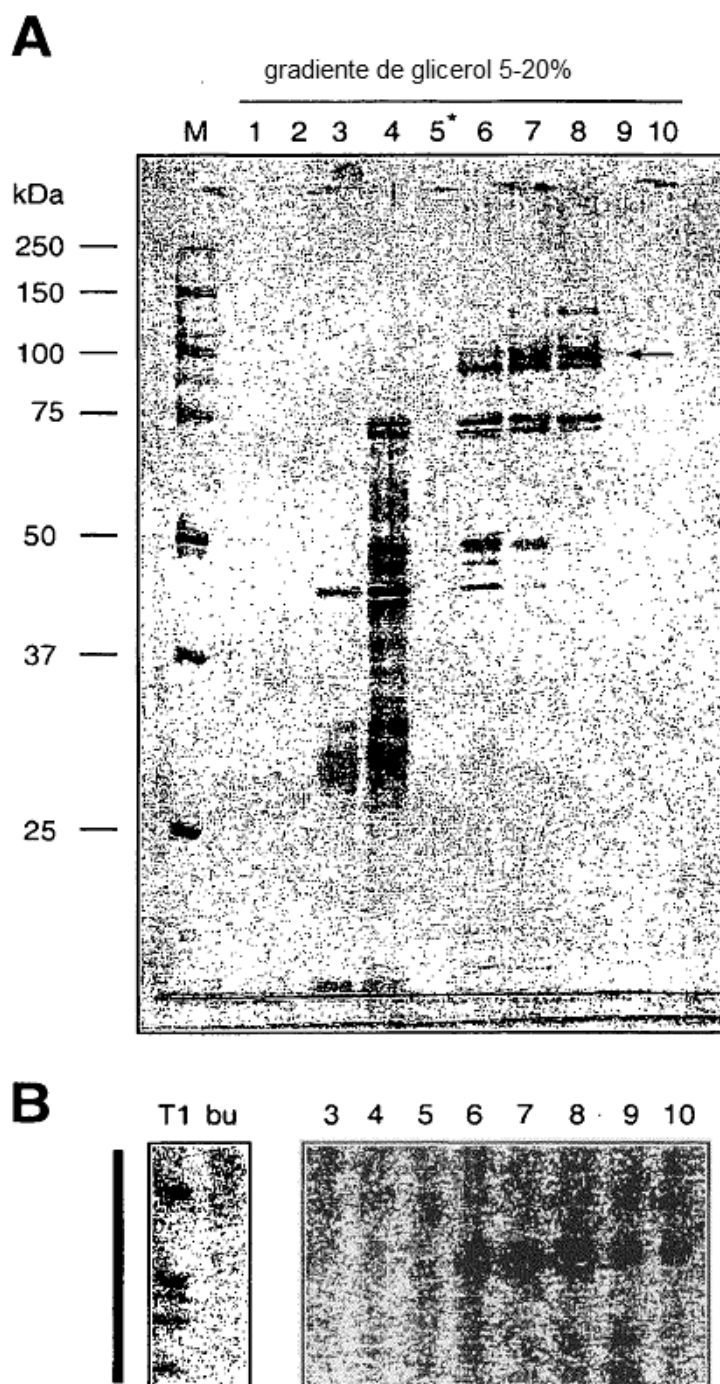


Figura 13 A

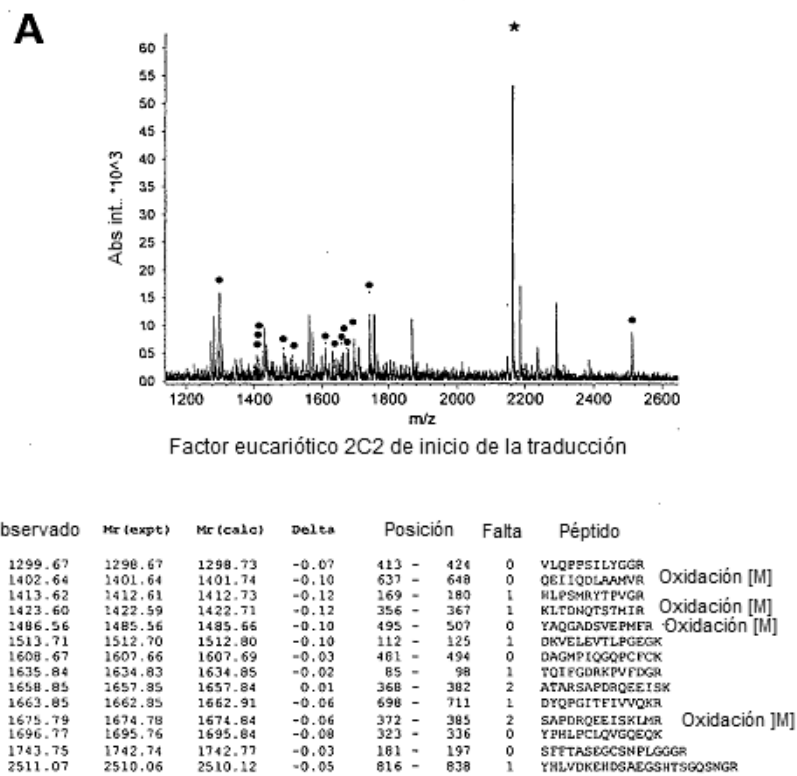
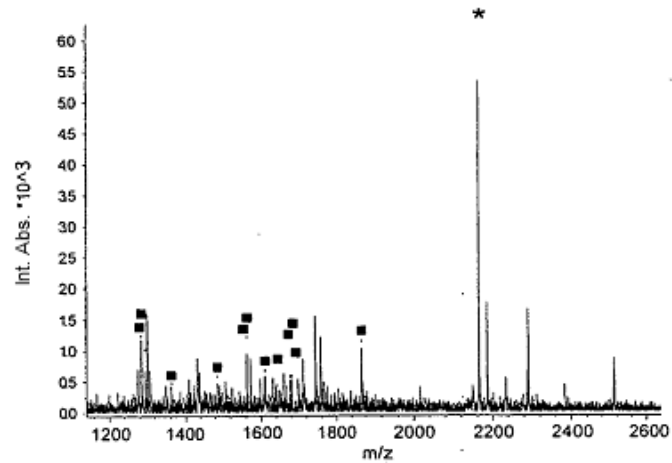


Figura 13B

B

Factor eucariótico 2C1 de inicio de la traducción

Observado	Mr (expt)	Mr (calc)	Delta	Posición	Falta	Péptido
1283.66	1282.65	1282.74	-0.09	410 -	421	0 VLPAPILQYGGR
1294.65	1293.64	1293.67	-0.03	794 -	805	0 SVSIPAPAYYAR
1361.61	1360.60	1360.70	-0.10	553 -	564	0 TSPQTLNLCLX
1486.56	1485.56	1485.66	-0.10	492 -	504	0 YAQCADSVEPMER
1560.76	1559.75	1559.83	-0.08	97 -	110	0 NIYTVTALPIQNER
1561.76	1560.75	1560.78	-0.02	111 -	124	1 VDPEVTIPGEGKDR
1608.67	1607.66	1607.69	-0.03	478 -	491	0 DAGMPTIQGPCECK
1640.74	1639.73	1639.82	-0.08	240 -	253	0 NIDEQPKPLTDSQR
1675.79	1674.78	1674.84	-0.06	369 -	382	2 SAPDRQEEISRLMK
1679.86	1678.85	1678.90	-0.05	695 -	708	1 DYQPGITYIVVQKR
1696.77	1695.76	1695.84	-0.08	320 -	333	0 YPHLPCLQVQGEQK
1867.85	1866.85	1866.87	-0.02	178 -	194	0 SFFSPPEGYYHPLGGGR

Oxidación [M]

Oxidación [M]

C

Figura 13C

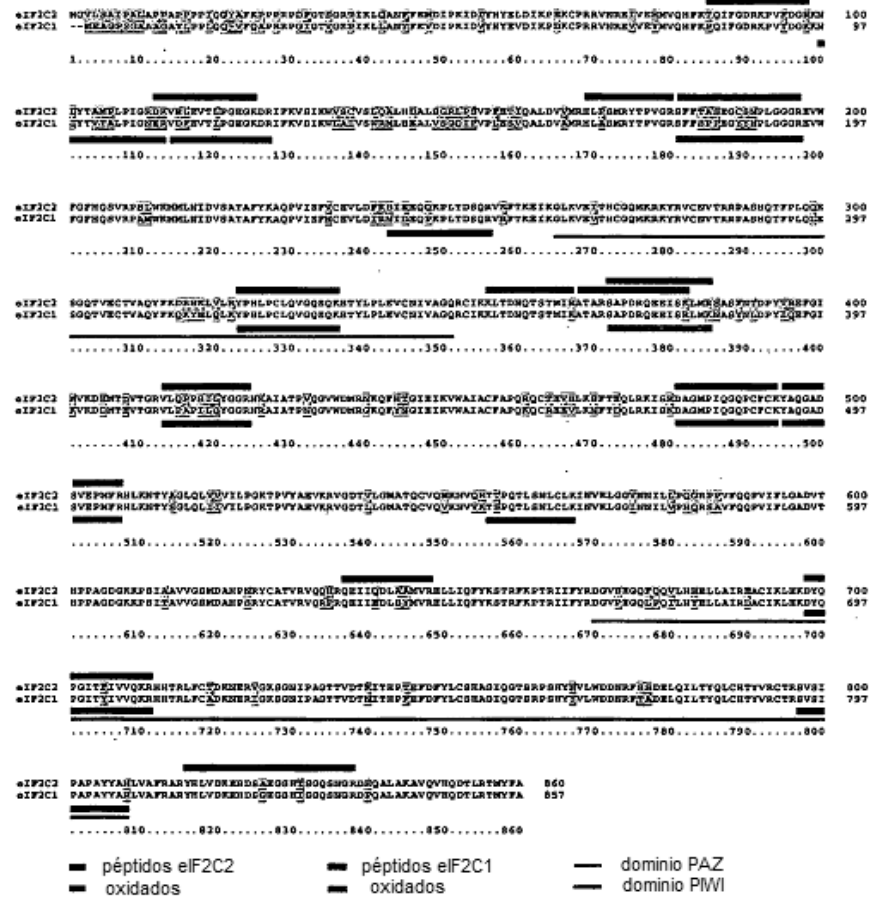


Figura 14

>eIF2C1, secuencia de proteínas predicha

MEAGPSGAAAGAYLPPLQQVFQAPRRPGIGTVGKPIKLLANYFEVDIPKIDVYHYEVDIKPD
KCPRRVNREVVEYMQHFKPQIFGDRKPVYDGGKNIYTVTALPIGNERVDFEVTIPGEGKDR
IFKVSIIKWLAIIVSWRMLHEALVSGQIPVPLESVQALDVAMRHLASMRYPVGRSFFSPPEGY
YHPLGGGREVWFGFHQSVRPAMWKMMLNIDVSATAFYKAQPVIEFMCEVLDIRNIDEQPKPL
TDSQRVRFTKEIKGLKVEVTHCGQMKRKYRVCNVTRRPASHQTFPLQLESGQTVECTVAQYF
KQKYNLQLKYPHLPCLQVGQEQKHTYLPLEVCNIVAGQRCIKKLTNDQSTMIKATARSAPD
RQEEISRLMKNASYNLDPIYIEFGIKVKDDMTTEVTGRVLPAPILQYGGNRRAIATPNQGVWD
MRGKQFYNGIEIKVWAIACFAPQKQCREEVLNFTDQLRKISKDAGMPIQGQPCFCKYAQGA
DSVEPMFRHLKNTYSGLQLIIVILPGKTPVYAIEVKRVGDTLLGMATQCVQVKNVVKTSPTL
SNLCLKINVKLGGINNVLVPHQRSVVFQQPVIFLGADVTHPPAGDGKKPSITAVVGSMDAHP
SRYCATVRVQRPRQEIIEDLSYMVRELLIQFYKSTRFKPTRIIFYRDGVPEGQLPQILHYEL
LAIRDACIKLEKDYQPGITYIVVQKRHHTRLFCADKNERIGKSGNI PAGTTVDTNITHPEF
DFYLCSHAGIQGTSRPSHYVVLWDDNRFADDELQILTYQLCHTYVRCTRVSIPAPAYYARL
VAFRARYHLVDKEHDSGEGSHISGQSNGRDPQALAKAVQVHQDTLRTMYFA

>eIF2C2, secuencia de proteínas predicha

MGVLSAIPALAPPAPPPPIQGYAFKPPRPDFGTSGRTIKLQANFFEMDIPKIDIIHYELDI
KPEKCPRRVNREIVEHMQHFKTQIFGDRKPVFDGRKNLYTAMPLPIGRDKVELEVTLPGEG
KDRIFKVSIIKWSVSLQALHDALSGRLPSVPFETIQALDVVMRHLPSMRYPVGRSFFTAS
EGCSNPLGGGREVWFGFHQSVRPSLWKMMLNIDVSATAFYKAQPVIEFVCEVLDFKSIIEQQ
KPLTDSQRVKFTKEIKGLKVEITHCGQMKRKYRVCNVTRRPASHQTFPLQLESGQTVECTVA
QYFKDRHKLVLRYPHLPCLQVGQEQKHTYLPLEVCNIVAGQRCIKKLTNDQSTMIKATARS
APDRQEEISKLMSASFNTDPYVREFGIMVKDEMTDVTGRVLQPPSILYGGNRKAIATPVQ
VWDMRNKQFHTGIEIKVWAIACFAPQRCQTEVHLKSFTQELRKISR DAGMPIQGQPCFCKYA
QGADSVPEPMFRHLKNTYAGLQLVVVILPGKTPVYAIEVKRVGDTVLGMATQCVQMKNVQRTTP
QTLNCLKINVKLGGVNNILLPQGRPPVVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSMD
AHPNRYCATVRVQQRQEI IQDLAAMVRELLIQFYKSTRFKPTRIIFYRDGVSEGOQFQVLH
HELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDKNERVKGSGNI PAGTTVDTKITHP
TEFDYLCSHAGIQGTSRPSHYVVLWDDNRFSSDELQILTYQLCHTYVRCTRVSIPAPAYY
AHLVAFRARYHLVDKEHDSAEGSHTSGQSNGRDHQALAKAVQVHQDTLRTMYFA

>eIF2C3, secuencia de proteínas predicha

SRSRVPVPGPGAAAAAPCAPASPRRHPSANIPEIKRYAAAAAGPGAGGAGDRGEAAPAA
AMEALGPGPASLFQPPRRPGLGTGKPIRLLANHFQVQIPKIDVYHYDVIDKPEKPRRRVN
REVVDTMVRHFQMIFGDRQPGYDGKRNMYTAHPLPIGRDRVDMEVTLPGEGKDQTFKVSQ
WVSVSLQLLLEALAGHLNEVPDDSVQALDVITRHLPSMRYPVGRSFFSPPEGYYHPLGGG
REVWFGFHQSVRPAMWNNMLNIDVSATAFYRAQPIIEFMCEVLDIQNINEQTKPLTDSQRVK
FTKEIRGLKVEVTHCGQMKRKYRVCNVTRRPASHQTFPLQLENGQAMECTVAQYFKQKYSQ
LKYPHLPCLQVGQEQKHTYLPLEVCNIVAGQRCIKKLTNDQSTMIKATARSAPDRQEEISR
LVKSNSMVGGPDPYKKEFGIVVHNEMTELTGRVLPAPMLQYGGNRKTVATPNQGVWDMRGKQ
FYAGIEIKVWAVACFAPQKQCREDLLKSFTDQLRKISKDAGMPIQGQPCFCKYAQGADSVPE
MPKHLKMTYVGLQLIIVVILPGKTPVYAIEVKRVGDTLLGMATQCVQVKNVVKTSPTLNLCL
KINAKLGGINNVLVPHQRSVVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSMDGHPSRYCA
TVRVQTSRQEISQELLYSQEVIQDLTNMVRELLIQFYKSTRFKPTRIIYYRGGVSEGQMKQV
AWPELIAIRKACISLEEDYRPGITYIVVQKRHHTRLFCADKTERVKGSGNVPAGTTVDSTIT
HPSEFDYLCSHAGIQGTSRPSHYQVLWDDNCFADDELQLLTYQLCHTYVRCTRVSIPAPA
YYARLVAFRARYHLVDKDHDSAEGSHVSGQSNGRDPQALAKAVQIHHDQTQHTMYFA

Figura 14

>eIF2C4, secuencia de proteínas predicha

AGPAGAQLLMVPRRPGYGTMGKPIKLLANCFQVEIPKIDVYLYEVDIKPKDKCPRRVNREVV
DSMVQHFKVTIFGDRRPVYDGRSLYTANPLPVATTGVDLDVTLPGEGGKDRPFKVSIFKVS
RVSWHLLHEVLTGRTLPEPLELDKPISTNPVHAVDVVLRHLPSMKYTPVGRSFFSAPEGYDH
PLGGGREVWFQGHQSVRPAMWKMLNIDVSATAFYKAQPVQFMCEVLDIHNIDEQFRPLTD
SHRVKFTKEIKGLKVEVTHCGTMRRKYRVCNVTRRPASHQTFPLQLENGQTVERTVAQYFRE
KYTLQLKYPHLPCLQVQGEQKHTYLPLEVCNIVAGQRCIKKLTNDQTSTMIKATARSAPDRQ
EEISRLVRSANYETDPFVQEFQFKVRDEMAHVTGRVLPAPMLQYGGNRNRTVATPSHGVDMMR
GKQFHTGVEIKMWAIACFATQRQCREEILKGFTDQLRKISKDAGMPIQGGQPCFCYKQAQADS
VEPMFRHLKNTYSGQLIIVILPGKTPVYAEVKRVGDTLLGMATQCVQVKVNIKTSPQTLN
LCLKINVKLGGINNILVPHQRPSVFFQPVIFLGADVTHPPAGDGKKPSIAAVVGSMDAHPSR
YCATVRVQRPRQEI IQDLASMVRELLIQFYKSTRFKPTRIIFYRDGVSEGGFRQVLYYELLA
IREACISLEKDYQPGITYIVVQKRHHTRLFCADRTERVGRSGNI PAGTTVDTDITHPYEFD
YLCSHAGIQGTSRPSHYHVLWDDNCF TADELQLLTYQLCHTYVRCRTRSVSIPAPAYYAHLLVA
FRARYHLVDKEHDSAEGSHVSGQSNGRDPQALAKAVQIHQDTRLRTMYFA

>HILI, secuencia de proteínas predicha

ISSGDAGSTFMERGVKNKQDFMDLSICTREKLAHVRNCKTGSSGIPVKLVNTNLFNLDFFQDW
QLYQYHVITYIPDLASRRRLRIALLYSHSELSNKAFAFDGAILFLSQKLEEKVTELSSETQRGE
TIKMTITLRELPSPPVCIQVFNIIFRKILKKLSMYQIGRNFYNPSEPMEIPQHKLSLWPG
FAISVSYFERKLLFSADVSYKVLNRNETVLEFMTALCQRTGLSCFTQTCEKQLIGLIVLTRYN
NRTYSIDDDIDWSVKPTHFQKRGDTEITYVDYKQYDITVSDLNQPMVLVSLKKKRNNDNSE
AQLAHLIPELCFLTGLTDQATSDFLMKAVAETRLSPSGRQQLARLVNDNIQRNTNARFEL
ETWGLHFGSQISLTGRIVPSEKILMQDHICQPVSAADWSKDIRTCKILNAQSLNTWLILCSD
RTEYVAESFLNCLRRVAGSMGFNVMCILPSNQKTYYSIKKYLSSDCPVPSQCVLARTLNKQ
GMMMSIATKIAMQMTCKLGGELWAVEIPLKSLMVVGIDVCKDALSKDVMVVGCVASVNPRIT
RWF SRCILQRTMTDVADCLKVFMTGALNKWKYNHDL PARIIVYRAGVGDGQLKTLIEYEV
QLSSVAESSNTSSRLSVIVVRKKCMRPFTEMNRTVQNPPLGTVDSEATRNEWQYDFYL
ISQVACRGTVSPTYYNVIYDDNGLKPDHMQRLTFKLCHLYYNWPGIVSVAPCQYAHKLTLFL
VAQSIHKEPSLELANHLFYL

>HIWI, secuencia de proteínas predicha

MTGRARARARGRARGQETAQLVGSTASQQPGYIQPRPQPPPAEGELFGRGRQRGTAGGTAKS
QGLQISAGFQELSLAERGGRRRDFHDLGVNTRQNL DHVKESKTGSSGIIVRLSTNHFRLTSR
PQWALYQYHIDYNPLMEARRLRSALLFQHEDLIGKCHAFDGTILFLPKRLQKQVTEVFSKTR
NGEDVRITITLTNELPPTSPTCLQFYNIIFRRLKIMNLQQIGRNYNPNNDPIDIPSHRLVI
WPGFTTSILQYENSIMLCTDVSHKVLRSETVLD MFNFYHQTEEHKFQEQVSKELIGLVVLT
KYNNKTYRVDDIDWDQNPKSTFKKADGSEVSFLEYRKYQYNQEITDLKQPVLSQPKRRRGP
GGTLPGPAMLIPELCYLTGLTDKMRNDFNVMKDLAVHTRLTPEQRQREVGRLLIDYIHKNDNV
QREL RDWGLSFDSNLLSFSGRILQTEKIHQGGKTFDYNPQFADWSKETRGAPLISVKPLDNW
LLIYTRRNYEAANS LIQNLFKVTPAMGMQMRKAIMIEVDDRTEAYLRVLQKQVTADTQIVVC
LLSSNRKDKYDAIKKYLCTDCPTPSQCVVARTLGKQQTVMATKIALQMNCKMGGELWRVD
IPLKLVMIVGIDCYHDMTAGRRSIAGFVASINEGMTRWFSRCIFQDRGQELVDGLKVCQLQAA
LRAWNSCNEYMPSRIIVYRDGVGDGQLKTLVNYEVPQFLDCLKSIGRGYNPRLTVIVVKRVR
NTRFFAQSGGRLQNPLPGTVIDVEVTRPEWYDFFI VSAVRSGSVSPHYNYIYDNSGLKPD
HIQRLTYKLCHIIYNNWPGVIRVPAPCQYAHKLAFVLVGQSIHREP NLSLSNRLYYL

Figura 15

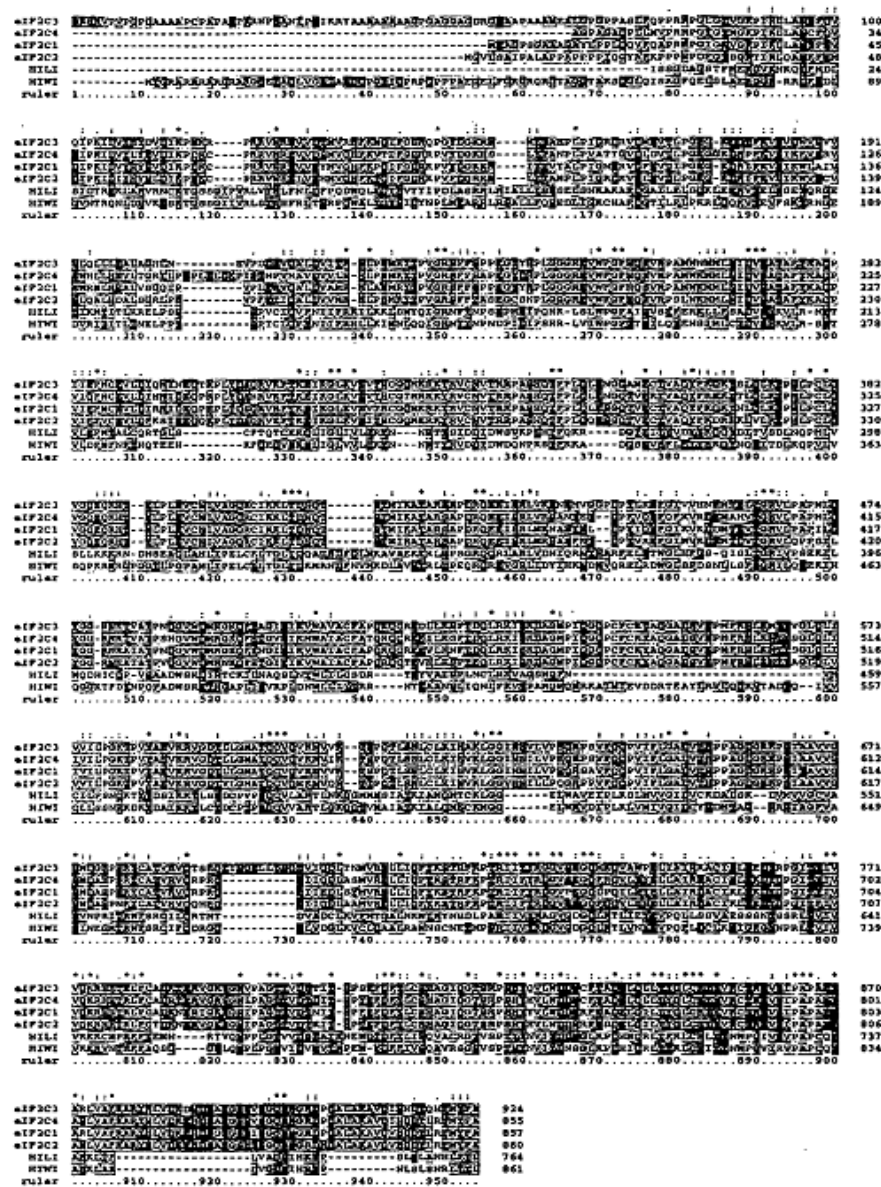


Figura 16

>eIF2C1, secuencia de ADNc de ORF predicho

```

ATGGAAGCGGGACCCTCGGGAGCAGCTGCGGGCGCTTACCTGCCCCCCTGCAGCAGGTGTT
CCAGGCACCTCGCCGGCCTGGCATTGGCACTGTGGGGAAACCAATCAAGCTCCTGGCCAATT
ACTTTGAGGTGGACATCCCTAAGATCGACGTGTACCCTACGAGGTGGACATCAAGCCGGAT
AAGTGTCCCCGTAGAGTCAACCGGGAAGTGGTGGAAATACATGGTCCAGCATTTCAGCCTCA
GATCTTTGGTGTATCGCAAGCCTGTGTATGATGGAAAGAAGAACATTTACACTGTACAGCAC
TGCCCATTTGGCAACGAACGGGTCGACTTTGAGGTGACAATCCCTGGGGAAGGGAAGGATCGA
ATCTTTAAGGTCTCCATCAAGTGGCTAGCCATTGTGAGCTGGCGAATGCTGCATGAGCCCT
GGTCAGCGGCCAGATCCCTGTTCCCTTGGAGTCTGTGCAAGCCCTGGATGTGGCCATGAGGC
ACCTGGCATCCATGAGGTACACCCCTGTGGGCGCGCTCCTTCTTCTCACCAGCTGAGGGCTAC
TACCACCCGCTGGGGGGTGGGCGCGAGGTCTGGTTCGGCTTTTACCAGTCTGTGCGCCCTGC
CATGTGGAAGATGATGCTCAACATTGATGTCTCAGCCACTGCCTTTTATAAGGCACAGCCAG
TGATTGAGTTTCATGTGTGAGGTGCTGGACATCAGGAACATAGATGAGCAGCCCAAGCCCTC
ACGGACTCTCAGCGCGTTTCGCTTCACCAAGGAGATCAAGGGCCTGAAGGTGGAAGTCACCCA
CTGTGGACAGATGAAGAGGAAGTACCGCGTGTGTAATGTTACCCGTCGCCCTGCTAGCCATC
AGACATTCCCCTTACAGCTGGAGAGTGGACAGACTGTGGAGTGCACAGTGGCACAGTATTTT
AAGCAGAAATATAACCTTCAGCTCAAGTATCCCCATCTGCCCTGCCCTACAAGTTGGCCAGGA
ACAAAAGCATACCTACCTTCCCCCTAGAGGTCTGTAACATTGTGGCTGGGCAGCGCTGTATTA
AAAAGCTGACCGACAACAGACCTCGACCATGATAAAGGCCACAGCTAGATCCGCTCCAGAC
AGACAGGAGGAGATCAGTCGCCTGATGAAGAATGCCAGCTACAACCTAGATCCCTACATCCA
GGAATTTGGGATCAAAGTGAAGGATGACATGACGGAGGTGACAGGGCGAGTGCTGCCGGCGC
CCATCTTGACAGTACGGCGGCCCGGAACCGGGCCATTGCCACACCCAATCAGGGTGTCTGGGAC
ATGCGGGGGAAACAGTTCTACAATGGGATTGAGATCAAAGTCTGGGCCATCGCCTGCTTCGC
ACCCCAAAAACAGTGTGCGAGAAGAGGTGCTCAAGAACTTCACAGACCAGCTGCGGAAGATTT
CCAAGGATGCGGGGATGCCTATCCAGGGTCAACCTTGTTTCTGCAAATATGCACAGGGGGCA
GACAGCGTGGAGCCTATGTTCCGGCATCTCAAGAACACCTACTCAGGGCTGCAGCTCATTAT
TGTCATCCTGCCAGGGAAGACGCCGGTGTATGCTGAGGTGAAACGTGTGCGGAGATACACTCT
TGGGAATGGCTACGCAGTGTGTGCAGGTGAAGAACGTGGTCAAGACCTCACCTCAGACTCTG
TCCAACCTCTGCCCTCAAGATCAATGTCAAACCTTGGTGGCATTAACAACATCCTAGTCCCACA
CCAGCGCTCTGCCGTTTTTCAACAGCCAGTGATATTCCTGGGAGCAGATGTTACACACCCCC
CAGCAGGGGATGGGAAAAAACCTTCTATCACAGCAGTGGTAGGCAGTATGGATGCCACCCCC
AGCCGATACTGTGCTACTGTGCGGGTACAGCGACCACGGCAAGAGATCATTTGAAGACTTGTC
CTACATGGTGCCTGAGCTCCTCATCCAATTCTACAAGTCCACCCGTTTCAAGCCTACCCGCA
TCATCTTCTACCGAGATGGGGTGCCTGAAGGCCAGCTACCCAGATACTCCACTATGAGCTA
CTGGCCATTCTGTATGCCTGCATCAAACCTGGAAAAGGACTACCAGCCTGGGATCACTTATAT
TGTGGTGCAGAAACGCCATCACACCCGCTTTTCTGTGCTGACAAGAATGAGCGAATTGGGA
AGAGTGGTAACATCCCAGCTGGGACCACAGTGGACACCAACATCACCCACCCATTTGAGTTT
GACTTCTATCTGTGACGCCACGCAGGCATCCAGGGCACCAGCCGACCATCCCATTACTATGT
TCTTTGGGATGACAACCGTTTTACAGCAGATGAGTCCAGATCCTGACGTACCAGCTGTGCC
ACACTTACGTACGATGCACACGCTCTGTCTCTATCCCAGCACCTGCCTACTATGCCCGCCTG
GTGGCTTTCCGGGCACGATAACACCTGGTGGACAAGGAGCATGACAGTGGAGAGGGGAGCCA
CATATCGGGGCAGAGCAATGGGCGGGACCCCAAGGCCCTGGCCAAAGCCGTGCAGGTTTACC
AGGATACTCTGCGCACCATGTACTTCGCT

```

Figura 1 de 6

>eIF2C2, secuencia ADNc de ORF predicho

Figura 16

ATGGGTGTTCTCTCTGCCATTCCCGCACTTGACACCTCCTGCGCCGCCGCCGCCCATCCAAGG
ATATGCCTTCAAGCCTCCACCTAGACCCGACTTTGGGACCTCCGGGAGAACAATCAAAATTAC
AGGCCAATTTCTTCGAAATGGACATCCCCAAAATTGACATCTATCATTTATGAATTGGATATC
AAGCCAGAGAAGTGCCCGAGGAGAGTTAACAGGGAAATCGTGGAACACATGGTCCAGCACTT
TAAACACAGATCTTTGGGGATCGGAAGCCCGTGTGTTGACGGCAGGAAGAATCTATACACAG
CCATGCCCCTTCCGATTTGGGAGGGACAAGGTGGAGCTGGAGGTCACGCTGCCAGGAGAAGGC
AAGGATCGCATCTTCAAGGTGTCCATCAAGTGGGTGTCTGCGTGAGCTTGCAGGCGTTACA
CGATGCACTTTCAGGGCGGCTGCCACGCTCCCTTTTGAGACGATCCAGGCCCCGGACGTGG
TCATGAGGCACTTGCCATCCATGAGGTACACCCCCGTGGGCCGCTCCTTCTTCACCGCGTCC
GAAGGCTGCTCTAACCTCTTGCGGGGGGCCGAGAAGTGTGGTTTGGCTTCCATCAGTCCGT
CCGGCCTTCTCTCTGGAATGATGCTGAATATTGATGTGTGTCAGCAACAGCGTTTACAAGG
CACAGCCAGTAATCGAGTTTGTGTTGTGAAGTTTGGATTTTAAAAGTATTGAAGAACAACAA
AAACCTCTGACAGATTCCCAAAGGGTAAAGTTTACCAAAGAAATTAAAGGTCTAAAGGTGGA
GATAACGCACCTGTGGGCAGATGAAGAGGAAGTACCGTGTCTGCAATGTGACCCGGCGGCCG
CCAGTCAACAAACATTTCCCGCTGCAGCAGGAGAGCGGGCAGACGGTGGAGTGCACGGTGGCC
CAGTATTTCAAGGACAGGCACAAGTTGGTTCTGCGCTACCCCCACCTCCCATGTTTACAAGT
CGGACAGGAGCAGAAACACACCTACCTTCCCTGGAGGTCTGTAACATTGTGGCAGGACAAA
GATGTATTAATAAATTAACGGACAATCAGACCTCAACCATGATCAGAGCAACTGCTAGGTGCG
GCGCCCGATCGGCAAGAGAGATTAGCAAATTGATGCGAAGTGCAAGTTTCAACACAGATCC
ATACGTCCGTGAATTTGGAATCATGGTCAAAGATGAGATGACAGACGTGACTGGGCGGGTGC
TGCAGCCGCCCTCCATCCTCTACGGGGCAGGAATAAAGCTATTGCGACCCCTGTCCAGGGC
GTCTGGGACATGCGGAACAAGCAGTTCCACACGGGCATCGAGATCAAGGTGTGGGCCATTGC
GTGCTTCCGCCCCCAGCGCCAGTGCACGGAAGTCCATCTGAAGTCTTACAGAGCAGCTCA
GAAAGATCTCGAGAGACGCTGGCATGCCATCCAGGGCCAGCCGTGCTTCTGCAAATACGCG
CAGGGGGCGGACAGCGTGGAGCCCATGTTCCGGCACCTGAAGAACACGTATGCGGGCCTGCA
GCTGGTGGTGGTCATCCTGCCCGGCAAGACGCCCGTGTACGCCGAGGTCAAGCGCGTGGGAG
ACACGGTGTGCGGGATGGCCACGCAGTGCCTGCAGATGAAGAACGTGCAGAGGACCACGCCA
CAGACCCTGTCCAACCTTTGCTTGAAGATCAACGTCAAGCTGGGAGGCGTGAACAACATCCT
GCTGCCCCAGGGCAGGCCGCCGGTGTTCAGCAGCCCGTCATCTTCTGGGAGCAGACGTCA
CTCACCCCCCGCCGGGGATGGGAAGAAGCCCTCCATTGCCCGCGTGGTGGGCAGCATGGAC
GCCCCACCCAATCGCTACTGCGCCACCGTGCCTGTCAGCAGCACCAGGAGATCATACA
AGACCTGGCCGCCATGGTCCGCGAGCTCCTCATCCAGTTCTACAAGTCCACGCGCTTCAAGC
CCACCCGCATCATCTTCTACCGCGACGGTGTCTCTGAAGGCCAGTTCCAGCAGGTCTCCAC
CACGAGTTGCTGGCCATCCGTGAGGCCTGTATCAAGCTAGAAAAAGACTACCAGCCCGGGAT
CACCTTCATCGTGGTGCAGAAGAGGCACCACACCCGGCTCTTCTGCACTGACAAGAACGAGC
GGGTGGGAAAAGTGGAAACATTCCAGCAGGCACGACTGTGGACACGAAAATCACCCACCCC
ACCGAGTTGCACTTCTACCTGTGTAGTCACGCTGGCATCCAGGGGACAAGCAGGCCTTCGCA
CTATCACGTCTCTGGGACGACAATCGTTTCTCTCTGATGAGCTGCAGATCCTAACCTACC
AGCTGTGTACACCTACGTGCGCTGCACACGCTCCGTGTCCATCCCAGCGCCAGCATACTAC
GCTCACCTGGTGGCCTTCCGGGCCAGGTACCACCTGGTGGATAAGGAACATGACAGTGCTGA
AGGAAGCCATACCTCTGGGCAGAGTAACGGGCGAGACCACCAAGCACTGGCCAAGCGGTCC
AGGTTACCAAGACACTCTGCGCACCATGTACTTTGCT

Figura 2 de 6

Figura 16

>eIF2C3, secuencia de ADNc de ORF predicho

```

AGCCGGAGCCGGGTCCCTGTCCCCGGGCGGGCGCCGCCGCCCGCCCTGCCAGCGCCCGC
GTCTCCGCGGGCGCCACCCAGCGCCAATATTCCGGAGATCAAGCGTTACGCGGGCGGGCGG
CGGCGGGCGGGGGCCCGGAGCGGGAGGCGCCGGGGACCGGGGCGAGGCGGGCCCCGCCGCC
GCCATGGAGGCGCTGGGACCCGGACCTCCGGCTAGCCTGTTTCAGCCACCTCGTCGTCCTGG
CCTTGGAAGTGTGGAAAACCAATTCGACTGTTAGCCAATCATTTTCAGGTTTCAGATTCCCTA
AAATAGATGTGTATCACTATGATGTGGATATTAAGCCTGAAAAACGGCCTCGTAGAGTCAAC
AGGGAGGTAGTAGATACAATGGTGGCGCACTTCAAGATGCAAATATTTGGTGATCGGCAGCC
TGGGTATGATGGCAAAAGAAACATGTACACAGCACATCCACTACCAAATTGGACGGGATAGGG
TTGATATGGAGGTGACTCTTCCAGGCGAGGGTAAAGACCAAACATTTAAAGTGTCTGTTTCAG
TGGGTGTCAGTTGTGAGCCTTCAGTTGCTTTTAGAAGCTTTGGCTGGGCACTTGAATGAAGT
CCCAGATGACTCAGTACAAGCACTTGATGTTATACAAGACACCTTCCCTCCATGAGGTACA
CCCCAGTGGGCGCTTCTTTTCTCACCCCGGAAGGTTACTACCACCCTCTGGGAGGGGGC
AGGGAGGTCTGGTTTGGTTTTCATCAGTCTGTGAGACCTGCCATGTGGAATATGATGCTCAA
CATTGATGTATCTGCAACTGCTTTCTACCGGGCTCAGCCTATCATTGAGTTCATGTGTGAGG
TTTTAGACATTGAGAACATCAATGAACAGACCAAACCTCTAACAGACTCCCAGCGTGTCAAA
TTTACCAAAGAAATCAGAGGTCTCAAAGTTGAGGTGACCCACTGTGGACAGATGAAACGAAA
ATACCGAGTTTGTAAATGTGACTAGACGGCCAGCCAGTCATCAAACCTTTTCTTTGACAGTAG
AAAACGGTCAAGCTATGGAATGTACAGTAGCTCAATATTTTAAGCAAAAGTATAGTCTGCAA
CTGAAATACCCCATCTTCCCTGTCTCCAAGTGGGACAAGAACAAAAGCATACATACTTGCC
ACTCGAGGTCTGTAATATAGTGGCAGGACAGCGATGTATCAAGAAGCTCACAGACAATCAGA
CTTCCACAATGATCAAAGCTACAGCAAGATCTGCTCCTGACAGACAGGAAGAGATCAGTAGA
CTGGTGAAGAGCAACAGTATGGTGGGTGGACCTGATCCATACCTTAAAGAATTTGGTATTGT
TGTCACAATGAAATGACAGAGCTCACAGGCAGGGTACTTCCAGCACCAATGCTGCAATATG
GAGGCCGGAATAAAACAGTAGCCACACCCAACAGGGTGTCTGGGACATGCGAGGAAAGCAG
TTTTATGCTGGCATTGAAATTAAGTTTGGGCAGTTGCTTTGTTTGCACCTCAGAAACAATG
TAGGGAAGATTTACTAAAGAGTTTCACTGACCAGCTGCGTAAATCTCTAAGGATGCAGGAA
TGCCCATCCAGGGTCAGCCATGTTTCTGCAAGTATGCACAAGGTGCAGACAGTGTGGAGCCT
ATGTTTAAACATCTGAAAATGACTTATGTGGGCCTACAGCTAATAGTGGTTATCCTGCCTGG
AAAGACACCAGTATATGCGGAGGTGAAACGTGTTGGAGATACCCTTCTAGGTATGGCCACAC
AGTGTGTCCAGGTAAAAAATGTAGTGAAGACCTCACCTCAAACCTTTTCCAATCTTTGCCTG
AAGATAAATGCAAACTTGGAGGAATTAACAATGTGCTTGTGCCTCATCAAAGGCCCTCGGT
GTTCCAGCAGCCTGTCATCTTCTTGGGAGCGGATGTCACACACCCCCAGCAGGGGATGGGA
AGAAACCTTCCATTGCTGCTGTGGTTGGCAGTATGGATGGCCACCCAGCCGGTACTGTGCC
ACCGTTCGGGTGCAGACTTCCCGGCAGGAGATCTCCCAAGAGCTCCTCTACAGTCAAGAGGT
CATCCAGGACCTGACTAACATGGTTTCGAGAGCTGCTGATTCAAGTTCTACAAATCCACACGCT
TCAAACCCACTCGGATCATCTATTACCGTGGAGGGGTATCTGAGGGACAAATGAAACAGGTA
GCTTGGCCAGAACTAATAGCAATTCGAAAGGCATGTATTAGCTTGGAAAGAAGATTACCGGCC
AGGAATAACTTATATTGTGGTGCAAAAAAGACATCACACACGACTCTTCTGTGCAGATAAAA
CAGAAAGGGTAGGGAAAAGTGGCAATGTACCAGCAGGCACTACAGTGGATAGTACCATCACA
CATCCATCTGAGTTTGACTTTTACCTCTGTAGTCATGCAGGAATTCAGGGAACCAGCCGTCC
CTCACATTACCAGGTCTTGTGGGATGACAACCTGCTTCACTGCAGATGAACTCCAGCTACTGA
CTTACCAGCTGTGTACACCTATGTGAGGTGCACTCGCTCAGTCTCTATTCCAGCCCCTGCA
TATTATGCCCCGGCTGTAGCATTTAGGGCAAGGTATCATCTGGTGGATAAAGATCATGACAG
TGCGGAAGGCAGTCATGTGTGAGGACAGAGCAACGGCCGGGATCCTCAGGCCTTGGCTAAGG
CTGTGCAAATCCACCATGATACCCAGCACACGATGTATTTTGCC

```

Figura 3 de 6

Figura 16

>eIF2C4, secuencia ADNc de ORF predicho

```
GCAGGACCCGCTGGGGCCCAGCCCCTACTCATGGTGCCAGAACCTGGCTATGGCACCAT
GGGCAAACCCATTAAACTGCTGGCTAACTGTTTTCAAGTTGAAATCCCAAAGATTGATGTCT
ACCTCTATGAGGTAGATATTAAACCAGACAAGTGTCTTAGGAGAGTGAACAGGGAGGTGGTT
GACTCAATGGTTCAGCATTTTAAAGTAACATATTTGGAGACCGTAGACCAGTTTATGATGG
AAAAAGAAGTCTTTACACCGCCAATCCACTTCCTGTGGCAACTACAGGGGTAGATTTAGACG
TTACTTTACCTGGGGAAGGTGGAAAAGATCGACCTTTCAAGGTGTCAATCAAATTTGTCTCT
CGGGTGAGTTGGCACCTACTGCATGAAGTACTGACAGGACGGACCTTGCTGAGCCACTGGA
ATTAGACAAGCCAATCAGCACTAACCCGTGTCATGCCGTTGATGTGGTGCTACGACATCTGC
CCTCCATGAAATACACACCTGTGGGGCGTTCATTTTTCTCCGCTCCAGAAGGATATGACCAC
CCTCTGGGAGGGGGCAGGGAAGTGTGGTTTGGATTCCATCAGTCTGTTCCGGCTGCCATGTG
GAAATGATGCTTAATATCGATGTTTCTGCCACTGCCTTCTACAAAGCACAACTGTAAATTC
AGTTCATGTGTGAAGTTCTTGATATTCATAATATTGATGAGCAACCAAGACCTCTGACTGAT
TCTCATCGGGTAAAATTCACCAAAGAGATAAAAGGTTTGAAGGTGAAGTGACTCATTGTGG
AACAATGAGACGGAAATACCGTGTGTGTAATGTAACAAGGAGGCCTGCCAGTCATCAAACCT
TTCCTTTACAGTTAGAAAACGGCCAACTGTGGAGAGAACAGTAGCGCAGTATTTACAGAGAA
AAGTATACTCTTCAGCTGAAGTACCCGCACCTTCCCTGTCTGCAAGTCGGGCAGGAACAGAA
ACACACCTACCTGCCACTAGAAGTCTGTAATATTGTGGCAGGGCAACGATGTATCAAGAAGC
TAACAGACAATCAGACTTCCACTATGATCAAGGCAACAGCAAGATCTGCACCAGATAGACAA
GAGGAAATTAGCAGATTGGTAAGAAGTGCAAATTATGAAACAGATCCATTTGTTTCAGGAGTT
TCAATTTAAAGTTCGGGATGAAATGGCTCATGTAAGTGGACGCGTACTTCCAGCACCTATGC
TCCAGTATGGAGGACGGAATCGGACAGTAGCAACACCGAGCCATGGAGTATGGGACATGCCA
GGGAAACAATTCACACAGGAGTTGAAATCAAATGTGGGCTATCGCTTGTTTGCCACACA
GAGGCAGTGCAGAGAAGAAATATTGAAGGGTTTCACAGACCAGCTGCGTAAGATTTCTAAGG
ATGCAGGGATGCCCATCCAGGGCCAGCCATGCTTCTGCAAATATGCACAGGGGGCAGACAGC
GTAGAGCCCATGTTCCGGCATCTCAAGAACACATATTCTGGCCTACAGCTTATTATCGTCAT
CCTGCCGGGGAAGACACCAGTGTATGCGGAAGTGAAACGTGTAGGAGACACACTTTTGGGTA
TGGCTACACAATGTGTTCAAGTCAAGAATGTAATAAAAAACATCTCCTCAAACCTGTCAAAC
TTGTGCCTAAAGATAAATGTTAACTCGGAGGGATCAATAATATTCTTGTACCTCATCAAAG
ACCTTCTGTGTTCCAGCAACCAGTGATCTTTTTGGGAGCCGATGTCACTCATCCACCTGCTG
GTGATGGAAAGAAGCCTTCTATTGCTGCTGTTGTAGGTAGTATGGATGCACACCCCAAGCAGA
TACTGTGCCACAGTAAGAGTTCAGAGACCCCGACAGGAGATCATCCAGGACTTGGCCTCCAT
GGTCCGGGAACCTTCTTATTCAATTTTATAAGTCAACTCGGTTCAAGCCTACTCGTATCATCT
TTTATCGGGATGGTGTTCAGAGGGGCAGTTTAGGCAGGTATTATATTATGAACTACTAGCA
ATTCGAGAAGCCTGCATCAGTTTGGAGAAAGACTATCAACCTGGAATAACCTACATTGTAGT
TCAGAAGAGACATCACACTCGATTATTTTGTGCTGATAGGACAGAAAGGGTTGGAAGAAGTG
GCAATATCCCAGCTGGAACAACAGTTGATACAGACATTACACACCCATATGAGTTCGATTTT
TACCTCTGTAGCCATGCTGGAATACAGGGTACCAGTCGTCCTTACACTATCATGTTTTATG
GGATGATAACTGCTTTACTGCAGATGAACCTCAGCTGCTAACTTACCAGCTCTGCCACACTT
ACGTACGCTGTACACGATCTGTTTCTATACCTGCACCAGCGTATTATGCTCACCTGGTAGCA
TTTAGAGCCAGATATCATCTTGTGGACAAAGAACATGACAGTGTGAAGGAAGTCACGTTTC
AGGACAAAGCAATGGGCGAGATCCACAAGCTCTTGCCAAGGCTGTACAGATTACCAAGATA
CCTTACGCACAATGTACTTCGCTTAA
```

Figura 4 de 6

Figura 16

>HILI, secuencia ADNc de ORF predicho

ATATCTTCTGGTGATGCTGGAAGTACCTTCATGGAAAGAGGTGTGAAAAACAAACAGGACTT
TATGGATTTGAGTATCTGTACCAGAGAAAAATTGGCACATGTGAGAAATTGTAAAAACAGGTT
CCAGTGGAATACCTGTGAAACTGGTTACAAACCTCTTTAACTTAGATTTTCCCCAAGACTGG
CAGCTATACCAGTACCATGTGACATATATTCAGATTTAGCATCTAGAAGGCTGAGAATTGC
TTTACTTTATAGTCATAGTGAACCTTTCCAACAAAAGCAAAAGCATTCGACGGTGCCATCCTTT
TTCTGTCACAAAAGCTAGAAGAAAAGGTCACAGAGTTGTCAAGTGAACTCAAAGAGGTGAG
ACTATAAAGATGACTATCACCCCTGAAGAGGGAGCTGCCATCAAGTTCTCCCGTGTGCATCCA
GGTCTTCAATATCATCTTCAGAAAGATCCTCAAAAAGTTGTCCATGTACCAAATTGGACGGA
ACTTCTATAATCCTTCAGAGCCAATGGAAATTCCCCAGCACAAATTATCCCTTTGGCCTGGG
TTTGCCATTTCTGTGTATATTTTGAAAGGAAGCTCCTGTTTAGTGCTGATGTGAGTTACAA
AGTCCTCCGGAATGAGACGGTTCTGGAATTCATGACTGCTCTCTGTCAAAGAAGTGGCTTGT
CCTGTTTCACCCAGACGTGTGAGAAGCAGCTAATAGGGCTCATTGTCCTTACAAGATACAAT
AACAGAACCTACTCCATTGATGACATTGACTGGTCAGTGAAGCCACACACACCTTTCAGAA
GCGGGATGGCACCAGATCACCTATGTGGATTACTACAAGCAGCAGTATGATATTACTGTAT
CGGACCTGAATCAGCCCATGCTTGTAGTCTGTAAAGAAGAAGAGAAATGACAACAGTGAG
GCTCAGCTCGCCACCTGATACCTGAGCTCTGCTTTCTAACAGGGCTGACTGACCAGGCAAC
ATCTGATTTCCAGCTGATGAAGGCTGTGGCTGAAAAGACACGTCTCAGTCCTTCAGGCCGGC
AGCAGCGCCTGGCCAGGCTTGTGGACAACATCCAGAGGAATACCAATGCTCGCTTTGAACTA
GAGACCTGGGGACTGCATTTTGAAGCCAGATATCTCTGACTGGCCGGATTGTGCCTTCAGA
AAAAATATTAATGCAAGACCACATATGTCAACCTGTGTCTGCTGCTGACTGGTCCAAGGATA
TTCGAACTTGCAAGATTTTAAATGCACAGTCTTTGAATACCTGGTTGATTTTATGTAGCGAC
AGAAGTGAATATGTTGCCGAGAGCTTTCTGAACTGCTTGAGAAGAGTTGCAGGTTCCATGGG
ATTTAATGTAATGTGCATTCTGCCTTCTAATCAGAAGACCTATTATGATTCATTAAAAAAT
ATTTGAGCTCAGACTGCCCAGTCCCAAGCCAATGTGTGCTTGCTCGGACCTTGAATAAACAG
GGCATGATGATGAGTATCGCCACCAAGATCGCTATGCAGATGACTTGCAAGCTCGGAGGCGA
GCTGTGGGCTGTGGAAATACCTTTAAAGTCCCTGATGGTGGTCGGTATTGATGTCTGTAAAG
ATGCACTCAGCAAGGACGTGATGGTTGTTGGATGCGTGGCCAGTGTTAACCCCAAGATCACCC
AGGTGGTTTTTCCCGCTGTATCCTTCAGAGAACAATGACTGATGTTGCAGATTGCTTGAAAGT
TTTCATGACTGGAGCACTCAACAAATGGTACAAGTACAATCATGATTTGCCAGCACGGATAA
TTGTGTACCGTGCTGGTGTAGGGGATGGTCAGCTGAAAACACTTATTGAATATGAAGTCCCA
CAGCTGCTGAGCAGTGTGGCAGAATCCAGCTCAAATACCAGCTCAAGACTGTCGGTGATTGT
GGTCAGGAAGAAGTGCATGCCACGATTCTTTACCGAAATGAACCGCACTGTACAGAACCCCC
CACTTGGCACTGTTGTGGATTCAAGAAGCAACACGTAACGAATGGCAGTATGACTTTTATCTG
ATCAGCCAGGTGGCCTGCCGGGAACTGTTAGTCTACCTACTATAATGTCATCTATGATGA
CAACGGCTTGAAGCCCGACCATATGCAGAGACTTACATTCAAATTGTGCCACCTGTACTACA
ACTGGCCGGGCATAGTCAGTGTCCAGCACCATGTGAGTATGCTCACAAGCTGACCTTTCTG
GTGGCACAAAGCATTCATAAAGAACCCAGTCTGGAATTAGCCAACCATCTCTTCTACCTG

Figura 5 de 6

Figura 16

>HIWI, secuencia ADNc de ORF predicho
 ATGACTGGGAGAGCCCGAGCCAGAGCCAGAGGAAGGGCCCGCGGTCAGGAGACAGCGCAGCT
 GGTGGGCTCCACTGCCAGTCAGCAACCTGGTTATATTCAGCCTAGGCCTCAGCCGCCACCAG
 CAGAGGGGGAATTATTTGGCCGTGGACGGCAGAGAGGAACAGCAGGAGGAACAGCCAAGTCA
 CAAGGACTCCAGATATCTGCTGGATTTCAGGAGTTATCGTTAGCAGAGAGAGGAGGTCGTCTG
 TAGAGATTTTCATGATCTTGGTGTGAATACAAGGCAGAACCTAGACCATGTTAAAGAATCAA
 AAACAGGTTCTTCAGGCATTATAGTAAGGTTAAGCACTAACCATTTCGCGCTGACATCCCGT
 CCCCAGTGGGCCTTATATCAGTATCACATTGACTATAACCCACTGATGGAAGCCAGAAGACT
 CCGTTCAGCTCTTCTTTTCAACACGAAGATCTAATTGGAAAGTGCCATGCTTTTGATGGAA
 CGATATTATTTTACCTAAAAGACTACAGCAAAAGGTTACTGAAGTTTGTAGTAAGACCCGG
 AATGGAGAGGATGTGAGGATAACGATCACTTTAACAAATGAACTCCACCTACATCACCAC
 TTGTTTGCAGTTCTATAATATTATTTTCAGGAGGCTTTTGAAAATCATGAATTTGCAACAA
 TTGGACGAAATTATTATAACCCAAATGACCAATGATATTCCAAGTCAAGGTTGGTGATT
 TGGCCTGGCTTCACTACTTCCATCCTTCAGTATGAAAACAGCATCATGCTCTGCACTGACGT
 TAGCCATAAAGTCCTTCGAAGTGAGACTGTTTTGGATTTCATGTTCAACTTTTATCATCAGA
 CAGAAGAACATAAATTTCAAGAACAAGTTTCCAAAGAACTAATAGGTTTAGTTGTTCTTACC
 AAGTATAACAATAAGACATACAGAGTGGATGATATTGACTGGGACCAGAATCCCAAGAGCAC
 CTTTAAGAAAGCCGACGGCTCTGAAGTCAGCTTCTTAGAATACTACAGGAAGCAATACAACC
 AAGAGATCACCGACTTGAAGCAGCCTGTCTTGGTCAGCCAGCCCAAGAGAAGGCGGGGCCCT
 GGGGGGACACTGCCAGGGCCTGCCATGCTCATTCCTGAGCTCTGCTATCTTACAGGTCTAAC
 TGATAAAATGCGTAATGATTTTAACGTGATGAAAGACTTAGCCGTTTCATACAAGACTAACTC
 CAGAGCAAAGGCAGCGTGAAAGTGGGACGACTCATTGATTACATTCATAAAAACGATAATGTT
 CAAAGGGAGCTTCGAGACTGGGGTTTGAGCTTTGATTCCAACCTACTGTCTTCTCAGGAAG
 AATTTTGCAACAGAAAAGATTCACCAAGGTGGAAAAACATTTGATTACAATCCACAATTTG
 CAGATTGGTCCAAAGAAACAAGAGGTGCACCATTAATTAGTGTTAAGCCACTAGATAACTGG
 CTGTTGATCTATACGCGAAGAAATTATGAAGCAGCCAATTCATTGATACAAAATCTATTAA
 AGTTACACCAGCCATGGGCATGCAAAATGAGAAAAGCAATAATGATTGAAGTGGATGACAGAA
 CTGAAGCCTACTTAAGAGTCTTACAGCAAAAAGGTCACAGCAGACACCCAGATAGTTGTCTGT
 CTGTTGTCAAGTAATCGGAAGGACAAATACGATGCTATTAAAAAATACCTGTGTACAGATTG
 CCCTACCCCAAGTCAGTGTGTGGTGGCCCGAACCTTAGGCAAACAGCAAACCTGTCATGGCCA
 TTGCTACAAAGATTGCCCTACAGATGAACTGCAAGATGGGAGGAGAGCTCTGGAGGGTGGAC
 ATCCCCCTGAAGCTCGTGATGATCGTTGGCATCGATTGTTACCATGACATGACAGCTGGGCG
 GAGGTCAATCGCAGGATTTGTTGCCAGCATCAATGAAGGGATGACCCGCTGGTTCTCACGCT
 GCATATTTTCAGGATAGAGGACAGGAGCTGGTAGATGGGCTCAAAGCTCTGCCTGCAAGCGGCT
 CTGAGGGCTTGGAATAGCTGCAATGAGTACATGCCAGCCGGATCATCGTGTACCGCGATGG
 CGTAGGAGACGGCCAGCTGAAAACACTGGTGAACCTACGAAGTGCCACAGTTTTTGGATTGTC
 TAAAATCCATTGGTAGAGGTTACAACCCTAGACTAACGGTAATTGTGGTGAAGAAAAGAGTG
 AACACCAGATTTTTTGCTCAGTCTGGAGGAAGACTTCAGAATCCACTTCCTGGAACAGTTAT
 TGATGTAGAGGTTACCAGACCAGAATGGTATGACTTTTTTATCGTGAGCCAGGCTGTGAGAA
 GTGGTAGTGTCTTCCACACATTACAATGTCATCTATGACAACAGCGGCCTGAAGCCAGAC
 CACATACAGCGCTTGACCTACAAGCTGTGCCACATCTATTACAACCTGGCCAGGTGTCAATCG
 TGTTCCTGCTCCTTGCCAGTACGCCCACAAGCTGGCTTTTCTTGTGGCCAGAGTATTCACA
 GAGAGCCAAATCTGTCACTGTCAAACCGCCTTTACTACCTC

Figura 6 de 6

Figura 17

A

Nombre del gen	1º par de cebadores (5' - 3')	2º par de cebadores (5' - 3')	longitud esperada (bp)
eIF2C1	GAGGTCTGTAACATTGTGGC* CGGTAGAAGATGATGCGGGT	GAGGTCTGTAACATTGTGGC* AAGTTCTTGAGCACCTCTTCTCGA	287
	GAGGTCTGTAACATTGTGGC CGGTAGAAGATGATGCGGGT	CCACACCAGCGCTCTGCC CTCACGCACCATGTAGGA	207
eIF2C2	GAGGTCTGTAACATTGTGGC CGGTAGAAGATGATGCGGGT	ATCCTGCTGCCCCAAGGG GATCTCCTGCCGGTGCTG	186
	GAGGTCTGTAACATTGTGGC* CGGTAGAAGATGATGCGGGT	GAGGTCTGTAACATTGTGGC* GATCTCCTGCCGGTGCTG	891
eIF2C3	AGAGCAACAGTATGGTGGGTGGAC TGGATGTGTGATGGTACT*	CCTCTACAGTCAAGAGGT TGGATGTGTGATGGTACT*	334
	CACCTGAATGAAGTCCCA TCCTGGATGACCTCTTGACTGTAG*	AGAGCAACAGTATGGTGGGTGGAC TCCTGGATGACCTCTTGACTGTAG*	808
eIF2C4	TCCGGCATCTCAAGAACACATATTCT GAAGTCATATGGGTGTGTAATGTCTG*	ATCCAGGACTTGGCCTCC GAAGTCATATGGGTGTGTAATGTCTG*	324
HILI	CAGCACAAATTATCCCTT* CGGCCTGAAGGACTGAGACGTGT	CAGCACAAATTATCCCTT* GTGTGTGGGCTTCACTGA	264
	TCTCTGTCAAAGAACTGGCTTGTCTT* CTGTACAGTCCGGTTCAT	TCTCTGTCAAAGAACTGGCTTGTCTT* CGGCCTGAAGGACTGAGACGTGT	393

* cebadores utilizados en las dos reacciones (PCR semi-anidada)

B

Nombre del gen	eIF2C1		eIF2C2		eIF2C3		eIF2C4	HILI	
longitud esperada (bp)	287	207	186	891	808	334	324	264	393
Productos de PCR									