



**CENTRO DE INVESTIGACIÓN Y DE  
ESTUDIOS AVANZADOS DEL  
INSTITUTO POLITÉCNICO NACIONAL**

DEPARTAMENTO DE GENÉTICA Y BIOLOGÍA MOLECULAR

**“Caracterización estructural de los precursores de *let-7*  
y su relevancia en la interacción con LIN28B”**

**TESIS**

que presenta

**M. en C. Sergio Israel Rangel Guerrero**

Para obtener el grado de

**Doctor en Ciencias**

En la especialidad de

**Genética y Biología Molecular**

**Director de Tesis**

Dr. Luis Marat Alvarez Salas

Ciudad de México

Marzo, 2021

El presente trabajo se realizó bajo la dirección del Dr. Luis Marat Alvarez Salas en el laboratorio de Terapia Génica del Departamento de Genética y Biología Molecular del Centro de Investigación y de Estudios Avanzados del IPN.

El autor de este trabajo expresa su agradecimiento a los ciudadanos mexicanos que permiten la generación de becas y estímulos a todos aquellos que soñamos con hacer ciencia en nuestro país.

Puntualmente, el autor contó con una beca nacional otorgada por el Consejo Nacional de Ciencia y Tecnología (CONACYT) con el número de registro **455381** y un apoyo adicional para realizar una estancia en el extranjero, publicado en la convocatoria **291250**.

## Agradecimientos

Un profundo agradecimiento a mi comité asesor conformado por la Dra Thompson Bonilla, la Dra. Bermúdez Cruz, el Dr. Durán Figueroa, el Dr. Valdés Flores y al Dr. Kameyama Kawabe por su gran atención, pasión y comentarios sobre este proyecto.

A mi tutor y gran amigo el Dr. Luis Marat Alvarez Salas, su guía y comprensión fueron un faro a través de este camino.

Un agradecimiento especial a la Dra. Encarnación Martínez Salas del CBMSO por la gran oportunidad de aprender de ella como ser humano y gran científica. Así mismo, a su grupo de trabajo: Rosario (Charo), Jorge y Azmane.

A mis antecesoras: Dra. Aida, Dra. Natalia, Dr. Alejandro, Dr. Giovanni y Dra. Diana; gracias por enseñarme como ser mejor cada día.

A todos mis compañeros del InDRE, por su gran compromiso y su increíble calidez humana.

A mis grandes amigos: Rocío, Ana, Saraí, Marisol, Octavio, Jonathan y Aldo.

A Diana por su gran amor a través de los tropiezos, los errores y los éxitos. Sin ti esto nunca hubiera sido posible.

A mis padres, Silvana Guerrero y Sergio Rangel. Nunca podré terminar de agradecer todo su apoyo y cariño durante estos años.

...cuando imaginamos una ciencia regida por una mezcla de objetividad y de rigor frío, que vuela en el limbo de las ideas puras, protegida de las pasiones y de las querellas, nos olvidamos que ha tenido un seno de notorias divergencias, y que, de no tenerlas, sería un pozo de tedio inacabable. No es un proceso continuo, no es una simple acumulación, no es el receptáculo de nuestras certidumbres, no es serena.

Su historia no es un largo río tranquilo: es palpitación y convulsión.

Es vida.”

***Ikram Antaki***

# Contenido

<b>Índice de Figuras .....</b>	<b>7</b>
<b>Índice de tablas.....</b>	<b>8</b>
<b>Índice de graficas .....</b>	<b>9</b>
<b>Resumen .....</b>	<b>10</b>
<b>Abstract.....</b>	<b>12</b>
<b>Introducción.....</b>	<b>14</b>
Caracterización estructural de RNA.....	14
miRNAs: características y biogénesis .....	17
La familia <i>let-7</i> : generalidades. ....	20
Regulación de la familia <i>let-7</i> .....	20
Regulación de <i>let-7</i> por LIN28 .....	22
<b>Antecedentes directos .....</b>	<b>24</b>
<b>Justificación.....</b>	<b>27</b>
<b>Hipótesis .....</b>	<b>27</b>
<b>Objetivo general.....</b>	<b>27</b>
<b>Objetivos específicos.....</b>	<b>28</b>
<b>Materiales y métodos .....</b>	<b>29</b>
Síntesis del reactivo NAI.....	29
Verificación del reactivo NAI .....	29
Evaluación <i>in cellulo</i> de la acilación del reactivo NAI .....	29

Determinación estructural del aptámero Sc5-c3 <sup>54</sup> .....	30
Análisis densitométrico de SHAPE y modelado estructural del aptámero Sc5-c3. ....	30
Síntesis <i>in vitro</i> de los precursores de <i>let-7</i> y sus mutantes.....	31
Clonación de los precursores de <i>let-7</i> .....	33
Transcripción <i>in vitro</i> de los precursores de <i>let-7</i> partiendo de las construcciones pCAP- <i>let-7c</i> , pCAP- <i>let-7f-1</i> , pCAP-miR-98 y pCAP- <i>let-7g</i> . ....	35
Modificación del RNA con el reactivo NMIA.....	36
Reverso transcripción del RNA modificado.....	36
Análisis de electroferogramas y modelado estructural. ....	37
Análisis estructural de los precursores de <i>let-7</i> .....	37
Inducción de proteína recombinante LIN28B .....	38
Ensayos de retardamiento electroforético (EMSA) .....	39
Análisis de los ensayos EMSA. ....	40
Estandarización de la transfección de pre- <i>let-7f-1</i> en células HEK-293TT..	40
Evaluación de la actividad de Luciferasa en células 293TT transfectadas con pCAP- <i>let-7f-1</i> .....	41
Evaluación del transcrito de fusión <i>let-7f-1</i> —luciferasa en células transfectadas.....	42
<b>Resultados .....</b>	<b>45</b>
Síntesis <i>in vitro</i> de pre- <i>let-7</i> .....	45
Caractrización estructural <i>in vitro</i> de pre- <i>let-7</i> .....	46
Ensayos de interacción LIN28B:pre- <i>let-7</i> .....	50

Evaluación de mutantes estructurales en la interacción pre-let-7:LIN28B...	51
Síntesis y verificación del reactivo NAI .....	53
Validación de la metodología SHAPE .....	55
Transfección de la construcción pCAP-let-7f-1 en células HEK-293TT.....	58
<b>Discusión .....</b>	<b>63</b>
<b>Conclusiones .....</b>	<b>70</b>
<b>Perspectivas .....</b>	<b>70</b>
<b>Referencias .....</b>	<b>71</b>
<b>Anexo .....</b>	<b>80</b>
Datos reactividad SHAPE: .....	80
Secuenciaciones construcciones pCAP-let7 .....	84
Modelo tridimensional pre-let-7f-1 .....	87
GEMIN5, dominio RBS1 y el RNA H12.....	87
<b>Artículos publicados .....</b>	<b>89</b>

## Índice de Figuras

Figura 1. Determinación estructural de RNAs mediante SHAPE.....	17
Figura 2. Biogénesis de miRNAs .....	19
Figura 3. Regulación de la familia <i>let-7</i> por LIN28.....	22
Figura 4. Características de LIN28A y LIN28B.....	22
Figura 5. Sitios de unión de LIN28 con <i>let-7</i> .....	23
Figura 6. <i>Let-7a-3</i> escapa a la regulación de LIN28A.. .....	24

Figura 7. Estructuras de pre-let-7 .....	26
Figura 8. SDS-PAGE y Western Blot de LIN28B recombinante.. .....	39
Figura 9. Construcción pre-let-7f-1 y digestión con SphI. ....	45
Figura 10. Transcripción <i>in vitro</i> de pre-let-7f-1.....	46
Figura 11. Estructuras secundarias de pre-let-7 .....	49
Figura 12. Árbol guía generado con LocARNA .....	50
Figura 13. Ensayos EMSA entre LIN28B y pre-let-7.. .....	51
Figura 15. Reacción de acilación con el reactivo NAI .....	54
Figura 16. RNA total modificado <i>in cellulo</i> con NAI .....	55
Figura 17. Patrón de reactividad del aptámero Sc5-c3 determinada por SHAPE. .....	56
Figura 18. Estructura Secundaria del aptámero SC5-c3 .....	57
Figura 19. RNA total de células HEK-293TT transfectadas.....	61
Figura 21. Estructuras ideales y alternas para la interacción con LIN28B .....	66
Figura 22. Estructura mínima de unión con LIN28. ....	68
Figura 23. Modelo estructural 3D pre-let-7f-1.....	87
Figura 20 Interacción RBS1:H12 .....	88

## Índice de tablas

Tabla 1 DNA molde para la generación de los precursores de <i>let-7</i> . ....	32
Tabla 2 Cebadores específicos para la amplificación de pre-let-7 mediante PCR.....	33
Tabla 3 Componentes para transcripción <i>in vitro</i> . ....	33



Tabla 4 Cebadores para la amplificación de Pre-let-7 con sitios EcoRI y BamHI. .....	34
Tabla 5 Transcripción <i>in vitro</i> para pre-let-7.....	35
Tabla 6 Mezcla de hibridación molde-cebador para reverso-transcripción.....	42
Tabla 7 Reacción de reverso transcripción .....	43
Tabla 8 Componentes de la reacción de PCR en tiempo real (qPCR) .....	43
Tabla 9 Análisis de similitud estructural de pre-let-7.. .....	50
Tabla 10 Viabilidad de celular evaluada con azul tripano. ....	58

## Índice de graficas

Gráfica 1 Eficiencia de reacción para la qPCR de let-7f-1—luciferasa.	44
Gráfica 3 Actividad de luciferasa en células 293TT transfectadas.	59
Gráfica 4 Actividad de luciferasa post-transfección..	60
Gráfica 5 Amplificación de U6 en células 293TT transfectadas.	61
Gráfica 6 Amplificación del transcrito let-7f-1—luciferasa en células 293TT transfectadas.	62

## Resumen

Durante las últimas décadas el estudio de los microRNAs o miRNAs ha permitido describir fenómenos de expresión y regulación génica en procesos celulares fundamentales. En particular, *let-7* es una familia de miRNAs altamente conservado en diversas especies y directamente involucrado en los procesos de diferenciación y proliferación celular. En el caso de humanos, *let-7* puede generarse a partir de 13 precursores distintos para llegar a un miRNA maduro único y regular los transcritos blanco. Los pre-*let-7* son especialmente regulados por la proteína LIN28A/B, la cual contiene dominios de unión a RNA, que se unen a secuencias consenso presentes en los precursores. La interacción LIN28:pre-*let-7* inhibe la biogénesis debido un efecto estérico o por el marcaje de pre-*let-7* para su degradación. Dicha regulación ha sido ampliamente caracterizada para ciertos precursores de *let-7* y la proteína LIN28A, uno de los dos parálogos en mamíferos de LIN28. Sin embargo, estudios recientes han demostrado que la regulación mediada por LIN28A puede ser diferencial de acuerdo con las características únicas que presenta cada miRNA precursor. Por su parte, la interacción de LIN28B y los precursores se encuentra inconclusa ya que poco se ha investigado sobre las características estructurales que guían esta unión.

En el presente trabajo, se realizó un análisis detallado de las estructuras secundarias de los precursores de *let-7* mediante la metodología SHAPE (del inglés "*Selective 2-Hydroxyl Aciliation analyzed by Primer Extension*"); con la cual, se logró determinar la estructura secundaria de cada precursor. Mediante análisis bioinformáticos, se determinó que los precursores de *let-7* analizados presentaron estructuras estadísticamente distintas, excepto pre-*let-7f-1* y pre-miR-98. Al comparar los resultados estructurales con las constantes de disociación entre LIN28B y cada precursor, se relacionó el patrón estructural presente en pre-*let-7f-1* y pre-miR-98 con una alta unión a LIN28B;

evidenciando un posible efecto de la estructura secundaria en la unión LIN28B:pre-let-7. Esta hipótesis se corroboró mediante la generación de dos mutantes estructurales, las cuales alteraron la estructura secundaria nativa de pre-let-7, sin alterar las secuencias consenso de unión previamente descritas. Ambas mutantes comprometieron negativamente la unión con LIN28B, con lo cual se validó la relevancia estructural de los pre-let-7 para su correcta unión con LIN28B. Dados los resultados obtenidos, se propone una estructura mínima de RNA que favorecería una interacción óptima con LIN28, la cual se fortalece a partir de ensayos estructurales previos de pre-let-7 y el parólogo LIN28A. Finalmente, se realizó una estandarización inicial para mapear la estructura *in cellulo* de pre-let-7, o cualquier otro RNA. Este procedimiento permitirá conocer la estructura secundaria de un RNA en condiciones nativas y correlacionarla con sus efectos celulares en condiciones de homeostasis o patológicas.

## Abstract

The study of microRNAs or miRNAs has revealed several mechanisms that regulate gene expression and cellular processes. In particular, *let-7* is a family of highly conserved miRNAs present in various species and directly involved in cell differentiation and proliferation. In humans, *let-7* can be generated from 13 different precursors which lead to a unique mature miRNA that regulates target transcripts. Pre-*let-7*s are particularly regulated by the LIN28 protein, which harbors RNA-binding domains that recognize consensus sequences present in the precursors. LIN28 binding to pre-*let-7* impairs *let-7* biogenesis by selective degradation or steric effects. This regulation process has been characterized for some *let-7* precursors and the LIN28A protein, one of the two mammalian paralogs of LIN28. However, recent evidence has unveiled that each pre-*let-7* could be differentially regulated by LIN28A. In addition, LIN28B binding to *let-7* remains inconclusive and the relevance of the structural features in this interaction has not been fully described.

In the present work, a detailed analysis of full-length pre-*let-7* secondary structures was performed using a methodology known as SHAPE ("Selective 2-Hydroxyl Acylation analyzed by Primer Extension"). The structural models were analyzed using diverse bioinformatic tools, showing that *let-7* precursors display statistically different structures (excluding pre-*let-7f-1* and pre-miR-98). Moreover, the structural patterns of pre-*let-7f-1* and pre-miR-98 were linked with high LIN28B interaction (as shown by binding analysis) evidencing a possible relationship of RNA secondary structure with LIN28B:pre-*let-7* complex stability. This hypothesis was corroborated by the design of two structural mutants, which changed the secondary structure of pre-*let-7* without altering the described consensus LIN28B binding sequences. Both mutants negatively compromised LIN28B binding to pre-*let-7*, thus validating the relevance of pre-*let-7* structure for a tight LIN28B binding. Given these results, a minimal RNA structure is

proposed for a precise binding with LIN28B, which is strengthened by previous structural analysis of *let-7* precursors and LIN28A.

Finally, a protocol to map the *in cellulo* structure for pre-*let-7* or any other RNA was standardized. This procedure will allow the determination of RNA structures in native conditions and correlate with the cellular effects under homeostasis or pathological conditions.

## Introducción

El ácido ribonucleico o RNA es una biomolécula compuesta por un esqueleto de fosfatos y ribosa, así como 4 bases nitrogenadas: adenina, uracilo, guanina o citosina (abreviadas A, U, G y C respectivamente). A pesar de ser una molécula con sólo tres componentes constitutivos, el RNA presenta una capacidad estructural jerarquizada, diversa y flexible, por lo cual, se encuentra involucrado en prácticamente cualquier proceso celular. Inicialmente, se demostró que el RNA mensajero (mRNA) es la molécula responsable de servir como intermediario entre DNA y la síntesis de proteínas<sup>1</sup>. Investigaciones posteriores señalaron que el RNA no sólo tenía funciones de mensajero, sino también podía generar estructuras tridimensionales y asociarse con proteínas ribosomales para catalizar y dirigir el enlace peptídico<sup>2</sup>. Aunado a estas funciones, algunas moléculas de RNA son capaces de crear interruptores moleculares que promueven la regulación de elementos génicos<sup>3</sup>, conocidos con el nombre de *riboswitches*. Con estos trabajos pioneros, comenzó un amplio y novedoso campo de estudio sobre la estructura y función del RNA *in vivo*<sup>4</sup>. Basados en dicho conocimiento, actualmente se han diseñado, sintetizado<sup>5</sup> o seleccionado *in vitro* nuevos RNAs con funciones regulatorias, de reconocimiento y hasta catalíticas llamados aptámeros, los cuales prometen ser una sustitución viable al uso de anticuerpos<sup>6-9</sup>.

## Caracterización estructural de RNA

Una característica importante para conocer las funciones y mecanismo de acción de un RNA es determinar su estructura. La estructura primaria del RNA se define como la secuencia lineal que se genera por los enlaces covalentes de un nucleótido con otro. La estructura secundaria se cimienta en las interacciones intramoleculares de los nucleótidos, generando bases apareadas o no apareadas que formaran los elementos estructurales más comunes: burbujas, tallos, burbujas asimétricas, pseudo-nudos. El apareamiento, guiado por puentes de hidrógeno, se rige por las interacciones canónicas de hibridación Watson-Crick (A:U, G:C) y por las hibridaciones no canónicas

Wooble (G:U) <sup>10,11</sup>. Además, las interacciones por puente de hidrogeno pueden guiar la formación de estructuras terciarias mediante las hibridaciones no canónicas Hoogsteen (U:A:U, C:G:C, U·G:C), y la formación de *zippers* de ribosa (formación de puentes de hidrogeno por el extremo 2'-OH de la ribosa) <sup>12</sup>. A pesar de que la estructura terciaria del RNA determina su función, actualmente se estima que para RNAs menores a 100 bases el número de motivos terciarios en promedio no es mayor a cinco<sup>13</sup>; por lo cual, la estructura de RNAs pequeños está mayoritariamente determinada por interacciones secundarias.

En años recientes, la caracterización estructural de RNAs ha sido un campo novedoso de estudio, el cual ha brindado nuevo conocimiento sobre sus diversas funciones celulares<sup>14,15</sup>. Una estrategia para evaluar la estructura del RNA ha sido mediante la generación de cristales y su resolución por difracción de rayos X<sup>16</sup>, lo cual se ha denominado como el “*estándar de oro*” para conocer las características tridimensionales de cualquier molécula o complejo molecular. Sin embargo, esta técnica implica una labor larga e intensiva y su resolución resulta en una estructura estática, que no representa necesariamente las características que puede presentar una biomolécula en su contexto nativo<sup>17</sup>. Aunado a esto, la naturaleza poli-aniónica y la gran capacidad dinámica (vibración) presente en los ácidos nucleicos representa una dificultad intrínseca para la generación de arreglos periódicos, equidistantes y paralelamente ordenados requeridos para la generación de cristales<sup>18</sup>; por lo que, la mayoría de las estructuras cristalográficas de RNAs se realizan en presencia de proteínas o con modificaciones artificiales que pueden alterar su naturaleza estructural real.

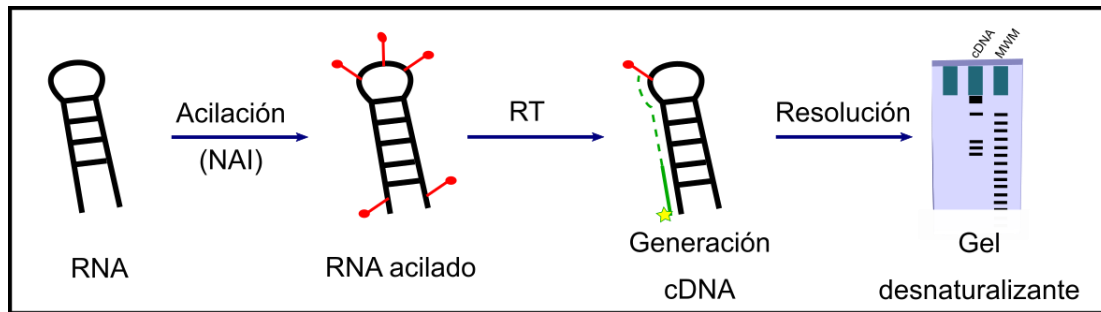
Por otra parte, la determinación estructural mediante el uso de ribonucleasas (RNAsas) ha sido una metodología ampliamente usada para determinar la estructura secundaria de RNAs. Esta técnica consiste en utilizar un conjunto de RNAsas, las cuales generan un corte en el RNA dependiendo de la conformación de cada nucleótido. Este tipo de mapeo estructural presenta

ciertas desventajas, debido a las necesidades catalíticas específicas de cada enzima, así como su incapacidad de realizar determinaciones *in vivo* o *in cellulo*<sup>19</sup>. En conclusión, la información obtenida mediante estas técnicas no ha permitido dilucidar las estructuras y funciones correspondientes en un ambiente nativo; abriendo nuevos retos técnicos que dependen de una gran cantidad de variables en función del contexto en el cual se encuentre un RNA<sup>17,20,21</sup>.

Además de las metodologías de mapeo convencionales, se han utilizado reactivos químicos para determinar el arreglo espacial de diversos RNAs. La primera generación de reactivos sintetizados para este fin, fueron incapaces de atravesar la membrana celular, generaban una alta cantidad de falsos positivos o sólo modificaban una o dos de las cuatro bases nitrogenadas, lo que presentaba el mismo problema que los métodos anteriores: no permitía conocer información completa sobre la estructura del RNA en condiciones nativas<sup>22</sup>. Debido a este problema, surgió una nueva generación de reactivos químicos, por ejemplo, imidazolína de ácido 2-metilnicotínico (NAI), ácido 2-metil-3-furóico (FAI), anhídrido 1-metil-7-nitroisatóico (1M7) y el anhídrido N-metilisatóico (NMIA), que tienen la capacidad de modificar a los 4 ribonucleótidos en el extremo 2'OH de la ribosa, añadiendo un grupo acilo. Esta modificación sólo ocurre si un nucleótido dentro del RNA se encuentra en cadena sencilla, ya que un apareamiento de bases, o bien, el contacto íntimo RNA-proteína inhibe estéricamente la reacción de acilación. Para complementar esta técnica, se acopló una reacción de reverso transcripción usando como molde el RNA acilado. En esta variante, la polimerización se detiene una base antes de encontrarse con una ribosa modificada, debido al impedimento estérico del grupo acilo, con lo cual se pueden determinar las bases que se encontraban en cadena sencilla e inferir aquellas bases que se estructuran en doble cadena (Figura 1). Esta novedosa técnica fue nombrada "acilación selectiva en el extremo 2' hidroxil, analizado por extensión de cebador" o SHAPE (del inglés "*Selective 2-Hydroxyl Acylation analyzed by Primer Extension*")<sup>23</sup>. Es importante mencionar que con un solo reactivo es posible obtener la información estructural de los cuatro nucleótidos, y que



algunos reactivos de SHAPE tienen la capacidad de atravesar la membrana celular. Dado estas características únicas, se ha podido describir estructuras de RNAs en condiciones nativas, lo cual demostró cambios significativos entre los resultados previamente generados *in vitro* e *in silico*<sup>11,24-28</sup>.



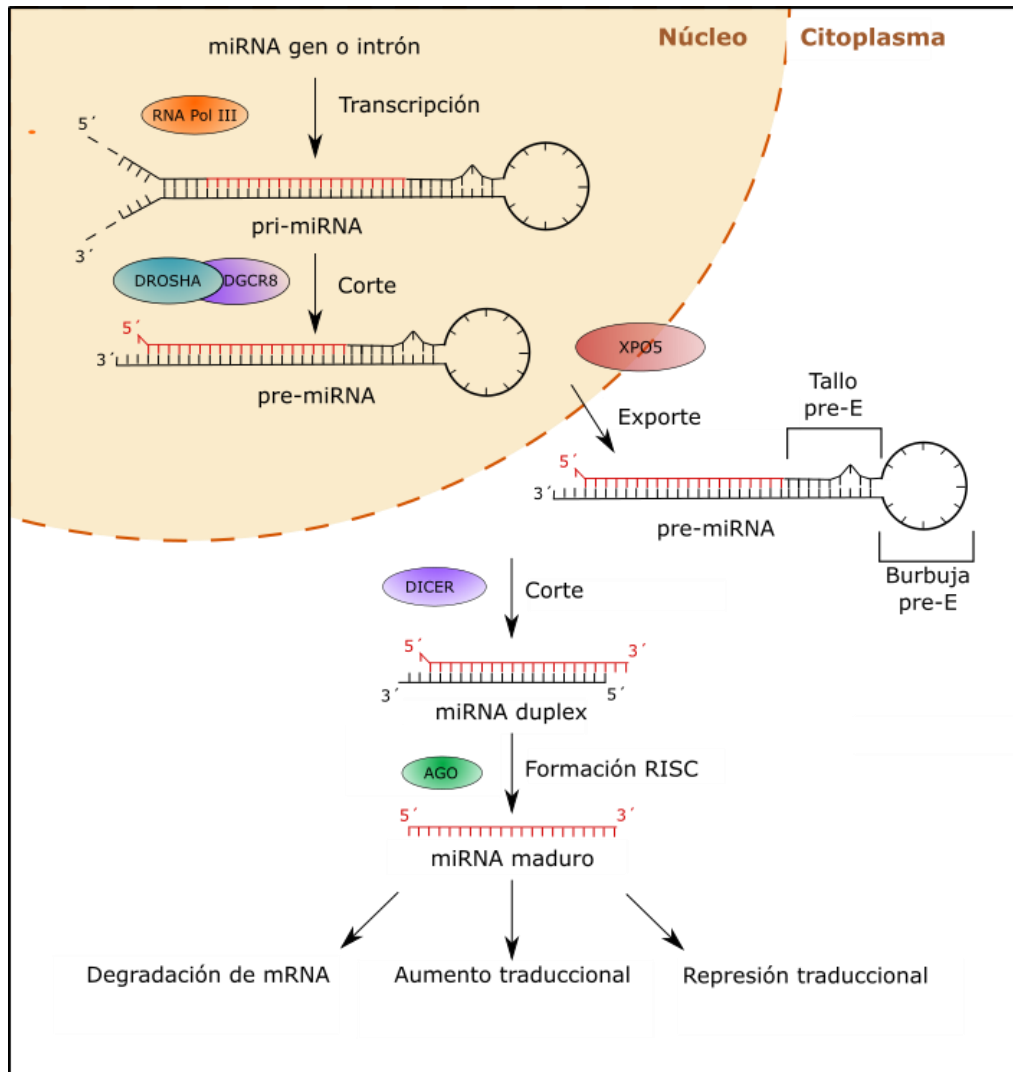
**Figura 1. Determinación estructural de RNAs mediante SHAPE.** Los reactivos compatibles con el método SHAPE (p.ej. NAI) acilan los ribonucleótidos que se encuentran en cadena sencilla, (bastones rojos) tras la reacción se realiza una reverso transcripción (RT) con un oligonucleótido marcado radioactivamente o con algún fluoróforo (marca verde). La RT se detiene una base antes de la base acilada. Los productos de RT (cDNA) se resuelven en un gel de poliacrilamida desnaturalizante, haciendo evidente los nucleótidos en ssRNA y dsRNA.

Durante las últimas décadas, un grupo especial de RNAs denominado micro RNAs (miRNAs) ha sido altamente estudiado debido a su participación en múltiples procesos vitales para la célula; lo cual, ha permitido establecer con más profundidad la relevancia estructural del RNA para los procesos de biogénesis, regulación y función de miRNAs,

### miRNAs: características y biogénesis

Los miRNAs son RNAs no codificantes (ncRNAs) de alrededor de 22 pb que, generalmente, se unen a la región 3' no traducible de mRNAs, provocando principalmente un bloqueo en la traducción y en algunos casos, la degradación del mRNA. De manera interesante, algunos miRNAs pueden tener un efecto contrario, favoreciendo la traducción de mRNAs<sup>29</sup>. Dadas estas características, los miRNAs participan en diversos procesos celulares, incluyendo desarrollo, proliferación, diferenciación, respuesta inmune, apoptosis, tumorigénesis, transducción de señales, diversas enfermedades y desordenes metabólicos.<sup>30</sup>

Aproximadamente, 30% de los miRNAs se generan a partir de intrones, mientras que el resto son expresados a partir de *loci* únicos para miRNAs. Su biogénesis comienza en el núcleo, sintetizados por las RNAs polimerasas II ó III, la cuales producen un transcrito largo conocido como pri-miRNA, dentro del cual se encuentra embebido uno o varios miRNAs; este transcrito pasa por los procesos usuales de corte y empalme (*splicing*), poli-adenilación y *capping*. Tras estas maduraciones, el pri-miRNA será escindido por el *Microprocesador*, un complejo binario de las proteínas DROSHA y DGCR8, el cual genera un RNA de aproximadamente 70 nucleótidos en forma de tallo-burbuja, conocido como miRNA precursor o pre-miRNA, con un extremo 3' saliente de uno o dos nucleótidos. El número de nucleótidos salientes es extremadamente importante, ya que aquellos que sólo contienen un nucleótido saliente necesitan ser adicionados con un nucleótido extra para continuar con los siguientes pasos de biogénesis<sup>31</sup>. Los pre-miRNAs son transportados del núcleo al citoplasma por la exportina-5 (XPO5) donde son procesados por la RNasa tipo III, DICER<sup>32</sup>. La enzima DICER se posiciona en sitios centrales del tallo del pre-miRNA para generar un corte asimétrico. El corte eliminará la región conocida como sitio del pre-elemento o sitio pre-E, y dará como resultado un RNA dúplex asimétrico (miRNA:miRNA\*) de ~22 nucleótidos. Posteriormente, la familia de proteínas Argonauta (AGO1, AGO2, AGO3, o AGO4) se unirá al complejo miRNA-DICER para producir el complejo de silenciamiento inducido por miRNAs (RISC). Tras la formación del complejo, solo una hebra del miRNA (miRNA maduro) será seleccionada y se unirá por hibridación Watson-Crick a su mRNA blanco; lo cual provocará su regulación post-transcripcional (Figura 2)<sup>33,34</sup>. Cabe mencionar que la estructura del extremo 3'UTR es esencial en la regulación por el complejo RISC, ya que si el sitio blanco del mRNA se encuentra altamente estructurado la unión del complejo RISC no se verá favorecida<sup>35</sup>.



**Figura 2. Biogénesis de miRNAs.** La transcripción de miRNAs se lleva a cabo por la RNAPolIII/III generando un transcrito largo (pri-miRNA), el cual es escindido por el Microprocesador (Drosha-DGCR8) para generar un pre-miRNA. El pre-miRNA es exportado al núcleo por la exportina 5 (XPO5) y procesado por DICER, generando un miRNA maduro de doble cadena. El complejo de silenciamiento inducido por RNA (RISC) se forma tras la unión de AGO al complejo miRNA-DICER. Finalmente, se selecciona una de las cadenas la cual se unirá al extremo 3'UTR de mRNAs y promoverá la regulación traduccional.

Debido a que el miRNA maduro es la molécula responsable de los efectos sobre la traducción, la mayoría de los trabajos actuales se han enfocado en validar los blancos potenciales de cada miRNA. Sin embargo, la biogénesis de los

miRNAs está sujeta a múltiples fenómenos de regulación, a nivel de pri-/pre-miRNA, como se ha descrito para la familia de *let-7*.

### **La familia *let-7*: generalidades.**

La familia de *let-7* fue la primera familia de miRNAs descrita, debido a sus efectos en diferenciación y división celular. Inicialmente, se identificó como un regulador embrionario en el nematodo *Caenorhabditis elegans* y posteriormente como una familia de miRNAs altamente conservada en diversos organismos, incluyendo insectos y mamíferos, lo cual permitió postular a *let-7* como un regulador génico ubicuo<sup>36</sup>. Como miRNA maduro, la familia *let-7* presenta la secuencia consenso “UGAGGUAGUAGGUUGUAUAGUU” la cual es constante a través de la escala evolutiva, sugiriendo blancos y funciones similares a través de diversos organismos. En *Drosophila melanogaster* y *C. elegans*, *let-7* sólo presenta una iso-forma, mientras que en humanos la familia *let-7* está compuesta por diez iso-tipos: *miR-98*, *miR-202*, *let-7a*, *let-7b*, *let-7c*, *let-7d*, *let-7e*, *let-7f*, *let-7g* y *let-7i*. Particularmente *let-7a* y *let-7f* pueden provenir de diferentes cromosomas, sub-clasificándose en *let-7a-1*, *let-7a-2*, *let-7a-3* y *let-7f-1*, *let-7f-2*, respectivamente. En condiciones normales la familia de *let-7* está involucrada en la regulación de los procesos de adhesión celular, desarrollo, formación de musculo y regulación génica. En contraste, una baja expresión de *let-7* se ha descrito en varios tipos de cáncer y se ha asociado a una baja supervivencia de pacientes con esta enfermedad. Por este motivo, la regulación de *let-7* ha sido un tema fundamental para entender los mecanismos moleculares de miRNAs<sup>34</sup>.

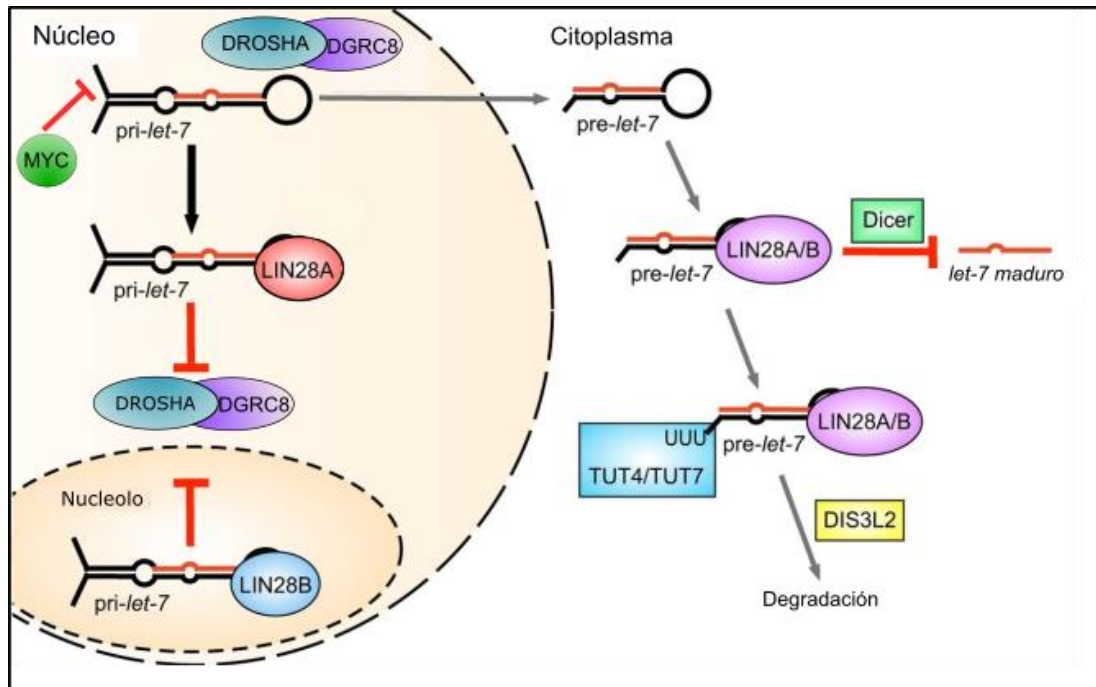
### **Regulación de la familia *let-7*.**

La regulación de la familia de *let-7* puede darse a niveles transcripcionales y post-transcripcionales. Un ejemplo de regulación transcripcional es dada por la proteína MYC, la cual se une a regiones previas al extremo 5´ del promotor o en el promotor del gen que codifica para el *pri-miRNA* de *let-7a-1*, *let-7f-1*, *let-7d*, inhibiendo su transcripción (Figura 3). Asimismo, el mRNA que codifica para

la proteína MYC contiene un sitio de unión para *let-7*, por lo cual se presenta un bucle de regulación negativa entre *let-7* y MYC<sup>37</sup>.

Existen otras regulaciones post-transcripcionales dadas por las enzimas que catalizan la adición de uridinas. La uridil-transferasa TUT-7 añade una uridina extra a aquellos *pre-let-7* que sólo contienen un nucleótido saliente en su extremo 3´ lo cual permite su reconocimiento por DICER y la continuación en su biogénesis<sup>31</sup>. Específicamente, el análisis de diversas estructuras cristalográficas permitió establecer que la estructura de dsRNA con dos nucleótidos salientes en el extremo 3´ presenta la topología ideal para la actividad catalítica de DICER<sup>38-40</sup>. Debido a la adición extra de una uridina, esta desviación de la biogénesis “canónica” se conoce como biogénesis no-canónica.

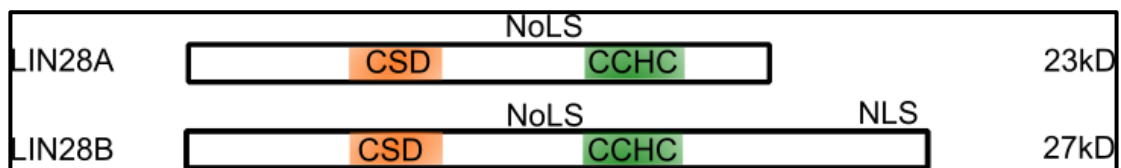
La región pre-E de los *pre-let-7* funciona como un andamio para la unión de diferentes proteínas como, LIN28A/B o KSRP<sup>41</sup>; que a su vez, conforman un nivel decisivo para la regulación post-transcripcional de *let-7*. Las proteínas TUT-7 y TUT-4 pueden oligo-uridilar a los *let-7*, mediante un mecanismo dependiente de la proteína LIN28A/B. La unión de los precursores de *let-7* con LIN28 inhibe el procesamiento por DICER, promueve la oligo-uridilación por TUTs y será una marca para la degradación del precursor por la exoribonucleasa DIS3L2 (Figura 3). La actividad de DIS3L2 se ve estimulada por la presencia de una oligo-uridilación de 10 nucleótidos, mientras que su máxima eficiencia se alcanza cuando los *let-7* contienen una oligo-uridilación de al menos 14 nucleótidos. El análisis cristalográfico de DIS3L2 permitió describir la especificidad de esta interacción, la cual se debe a un acomodo estructural entre los pre-*let-7* y los tres primeros dominios de DIS3L2, los cuales son específicos para la interacción con RNA. Dicha interacción es altamente favorable si los pre-*let-7* contienen una oligo-uridilación suficientemente larga que les permita aproximarse a los dominios específicos de interacción con uridinas en DIS3L2, promoviendo la degradación del precursor<sup>42,43</sup>.



**Figura 3. Regulación de la familia *let-7* por LIN28.** *Let-7* puede ser regulado a nivel transcripcional por la proteína MYC. LIN28A inhibe la acción de DROSHA en el núcleo y puede dirigir la degradación de pre-*let-7* en un mecanismo TUT-dependiente. LIN28B retiene a pre-*let-7* en el núcleo inhibiendo su biogénesis (modificado de <sup>44</sup>).

### Regulación de *let-7* por LIN28

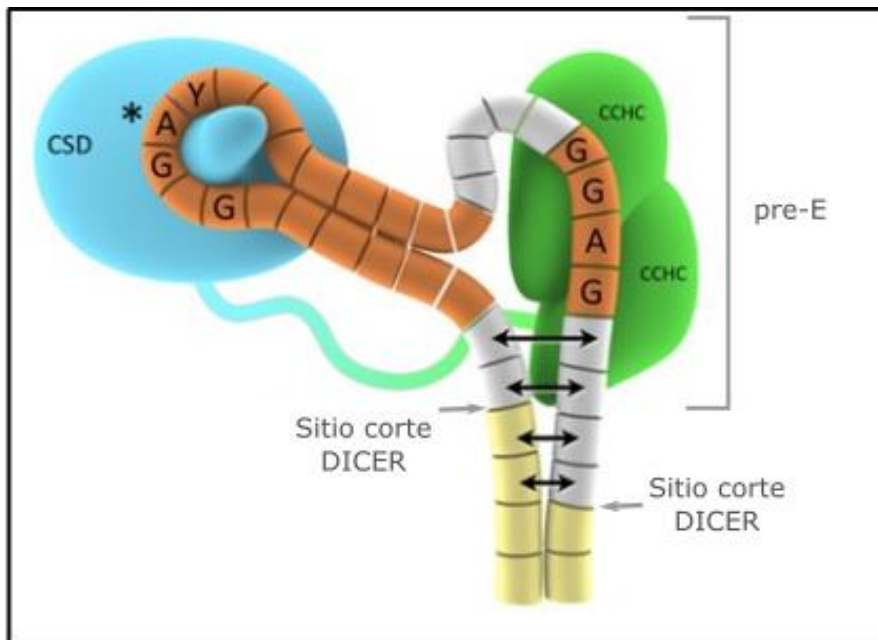
Como se mencionó previamente LIN28, es requerido para la oligo-uridilación de la familia *let-7*. En humanos LIN28 presenta dos parálogos LIN28A (23kDa) y LIN28B (27kDa), los cuales presentan una señal de localización nucleolar (NoLS) y dos dominios de unión a RNA: el dominio “cold shock” (CSD) y el dominio de dedos de zinc (ZnF o CCHC). En el caso particular de LIN28B, éste presenta una señal de localización nuclear (NLS) en el extremo carboxilo (Figura 4)<sup>45</sup>.



**Figura 4. Características de LIN28A y LIN28B.** LIN28A/B contiene dos dominios de unión a *let-7*: el dominio “coldshock” (CSD) y el dominio de dedos de zinc (CCHC). LIN28A/B presentan

una señal de localización nuclear (NLS). LIN28B presenta una señal de localización nucleolar (NoLS) y un extremo carboxilo terminal mayor.

La unión de *let-7* y LIN28A/B se da mediante el CSD con la secuencia consenso “5'-GNGAY-3'”, la cual se encuentra en la burbuja del pre-E. Además, los CCHC se une a la secuencia consenso “GGAG”, la cual se encuentra en la burbuja asimétrica del tallo pre-E (Figura 5)<sup>46</sup>. Ambos dominios permiten la unión de LIN28A/B a los *pri-/pre-let-7* para bloquear su biogénesis. Como se mencionó previamente, la unión de los pre-*let-7* con LIN28A/B inhibe su unión con DICER en el citoplasma, promueve la oligo-uridilación mediada por TUT-4/7 y su degradación por DIS3L2. Además de este mecanismo, LIN28A/B tiene la capacidad de unirse a los *pri-let-7* y pre-*let-7* en el núcleo, inhibiendo la acción del *Microprocesador* y su exporte al citoplasma. En el caso de LIN28B se ha demostrado que se une a los *pri-let-7* en el núcleo y los transporta hacia el nucléolo, gracias a su NoLS. Dicha unión inhibe la interacción de *let-7* con el *Microprocesador* y por ende su maduración, en un mecanismo TUT-independiente<sup>46,47</sup>.

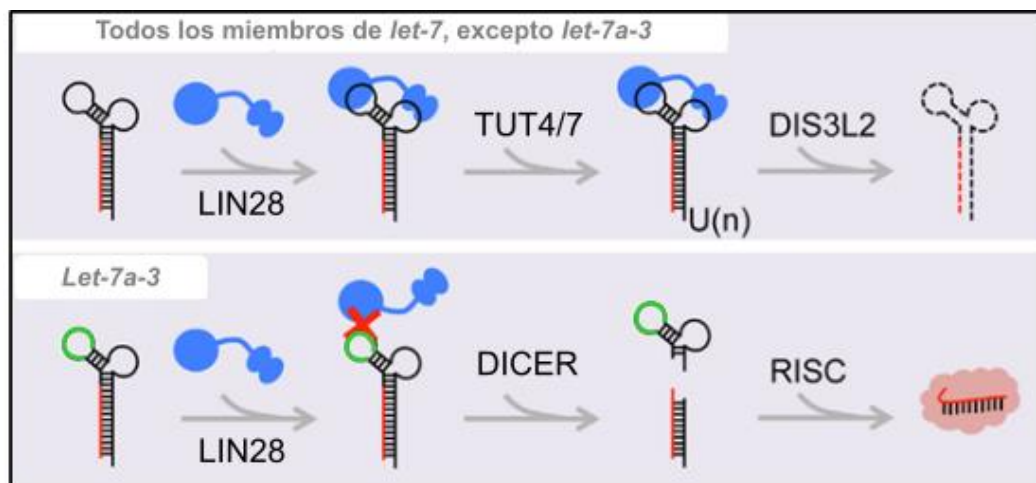


**Figura 5. Sitios de unión de LIN28 con *let-7*.** El dominio “coldshock” (CSD) se une a la secuencia consenso “5'-GNGAY-3'” de la burbuja pre-E. El dominio de dedos de zinc (CCHC)

se unen a la secuencia consenso “5'-GGAG-3'” presente en la burbuja asimétrica del elemento pre-E. La unión de LIN28 con *let-7* bloquea el corte por DICER (Modificado de <sup>46</sup>.)

## Antecedentes directos

Dada la gran cantidad de información correspondiente a la unión, regulación e inhibición de LIN28 hacia *let-7*, se ha fijado a esta interacción como un paradigma regulatorio en miRNAs. Sin embargo, reportes recientes han demostrado que un miembro de la familia de *let-7* puede escapar de la regulación de LIN28. En 2015, se demostró que el precursor murino de *let-7c-2* no puede ser reprimido por LIN28A, proponiendo que este fenómeno sucede por una variación en la estructura primaria que contiene pre-*let-7c-2* en la burbuja pre-E. Los resultados señalan que la sustitución parcial de la secuencia del pre-E de *let-7c-2* (5'-CUCUG-3' → 5'-UCACA-3') restauró parcialmente la unión entre LIN28A y su regulación *in cellulo* (Figura 6)<sup>48</sup>.



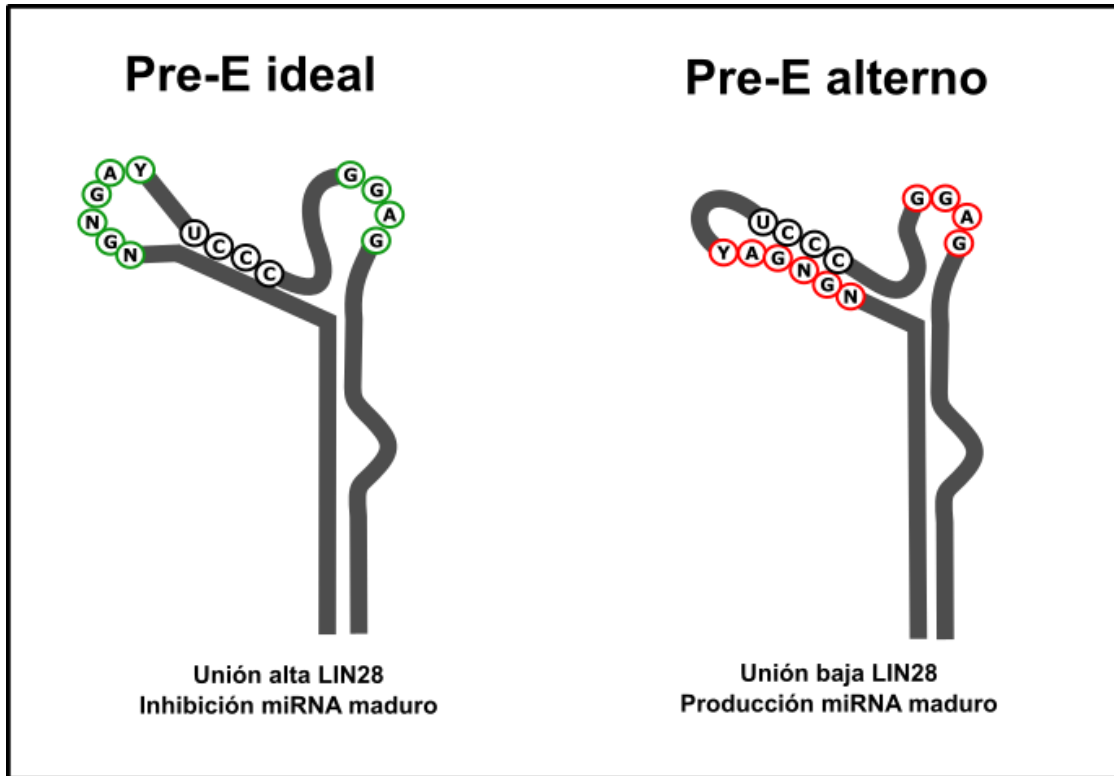
**Figura 6. *Let-7a-3* escapa a la regulación de LIN28A.** LIN28 se une a los precursores de *let-7* favoreciendo su degradación por DIS3L2 en un mecanismo TUT dependiente. *Let-7a-3* contiene una variante nucleotídica en la burbuja del pre-E (círculo verde) que evita su unión con LIN28 y su permite su maduración por DICER y la formación del complejo de silenciamiento inducido por miRNA (RISC) (modificado de <sup>48</sup>).

Una segunda evidencia clave fue la reclasificación de la secuencia consenso 5'-GNGAY-3' a 5'-UGAU-3', lo cual propone que aquellos precursores con



secuencia 5'-UGAU-3' se unen con alta afinidad a LIN28A. Sin embargo, aun con la presencia del motivo 5'-UGAU-3' la unión de los precursores a LIN28A no es homogénea. Además, se observa un caso muy drástico que contradice esta hipótesis, en el cual pre-let-7g (UGAU-positivo) se une de manera idéntica a LIN28A que pre-let-7c (UGAU-negativo)<sup>49</sup>. Estas discrepancias sugieren fuertemente que los factores que permiten una correcta unión pre-let-7:LIN28 no dependen totalmente de las secuencias consenso ni de la estructura primaria.

Aunado a esto, resultados de nuestro grupo de trabajo, así como simulaciones *in silico*<sup>45,50,51</sup>, sugieren que existen diferencias en la estructura secundaria de los pre-let-7 y que dichas diferencias estructurales pueden comprometer la interacción con LIN28 (Figura 7). En este sentido, hipotetizamos que los precursores de *let-7* pueden dividirse en dos grupos: ciertos precursores presentarían una pre-E ideal, el cual favorecería una interacción estable con la proteína LIN28A/B. Por otro lado, ciertos precursores adoptarían una o varias estructuras alternas en la región del pre-E, contribuyendo a una interacción poco favorable con LIN28A/B.



**Figura 7. Estructuras de pre-let-7.** Los precursores de *let-7* pueden presentar un pre-E con características distintas. Se hipotetiza que algunos precursores contienen un pre-E ideal para la interacción con LIN28. Alternativamente, ciertos precursores presentarían una estructura alterna la cual resultaría adversa para la unión con LIN28. Figuras basadas en los resultados de <sup>50,51</sup>.

A pesar de que existen múltiples trabajos sobre la familia *let-7*, la estructura completa de algunos precursores no ha sido caracterizada y para otros precursores como *let-7a*, *let-7e*, *let-7c*, *let-7f-2* no se ha generado ningún modelo estructural basado en evidencia experimental. Además, ninguna de las pasadas determinaciones se ha realizado en condiciones *in cellulo*; por lo cual, poco se sabe sobre las estructuras de cada miembro de la familia *let-7* y si estas se encuentran representadas tanto en condiciones *in vitro* como *in cellulo*. Por otra parte, modelos *in silico* han sugerido que LIN28B, debido su extremo carboxilo terminal de mayor longitud, tendría características de unión distintas, en comparación con LIN28A<sup>52</sup>. Finalmente, de corroborarse que los precursores de *let-7* presentan diferentes estructuras, estas características

podrían explicar su regulación diferencial por LIN28, y por ende, la necesidad evolutiva de mantener distintas estructuras en cada precursor a pesar de que la molécula efectora, es decir los miRNA maduros, son prácticamente idénticos y con la potencialidad de interactuar con los mismos blancos.

## **Justificación**

El estudio de la familia *let-7* ha resultado paradigmático para entender los mecanismos de biogénesis, regulación y funciones de diversos miRNAs; por lo cual, dilucidar las características estructurales de los precursores de esta familia y determinar su relevancia, nos brindaría nuevo conocimiento sobre la participación de dichas características en la regulación de miRNAs. Además, la adición de nuevo conocimiento sobre la estructura del RNA podría ser aplicado para entender las funciones de otros ncRNAs y generar las bases para el diseño racional *de-novo* de ncRNAs con funciones predecibles.

## **Hipótesis**

Si los precursores de la familia *let-7* presentan y conservan diferencias estructurales en distintas especies y condiciones fisiológicas de la célula entonces dichas diferencias podrían resultar en una regulación diferencial por LIN28B.

## **Objetivo general**

Determinar las estructuras de los precursores de *let-7* y relacionarlo con su interacción con LIN28B.

## Objetivos específicos

1. Clonar y sintetizar *in vitro* a los precursores de *miR-98*, *let-7f-1*, *let-7c* y *let-7g*
2. Evaluar la estructura *in vitro* de los precursores de *miR-98*, *let-7f-1*, *let-7c* y *let-7g* mediante la metodología SHAPE.
3. Validar la relevancia estructural de los precursores de *let-7* en su interacción con LIN28B
4. Sintetizar y verificar la capacidad de acilación del reactivo NAI.
5. Estandarizar la transfección de los precursores de *let-7* en la línea celular 293TT.

## **Materiales y métodos**

### **Síntesis del reactivo NAI**

Se decidió realizar la síntesis del reactivo NAI, debido a que ha demostrado las mejores propiedades para la metodología SHAPE *in cellulo*: alta solubilidad, alta permeabilidad a las membranas celulares, amplio rango dinámico y alta relación señal vs. fondo<sup>23</sup>. Para su síntesis se mezcló 1 mmol de carboxidiimidazol (Sigma-Aldrich Inc.) y 1 mmol de ácido 2-metilpiridin-3-carboxílico (Sigma-Aldrich Inc.) con 1 mL de DMSO en un frasco de vidrio estéril. La mezcla se mantuvo en agitación constante por 15 minutos a temperatura ambiente. El producto final se asumió como una mezcla equimolar de NAI e imidazol a una concentración final de 1 M.

### **Verificación del reactivo NAI**

El reactivo NAI sintetizado se verificó de acuerdo a la metodología previamente descrita<sup>53</sup>. Brevemente, se incubó 0.5 fmoles de  $\gamma$ -[<sup>32</sup>P]-ATP (Perkin Elmer, Inc., Waltham, MA) con una concentración final de NAI 100 mM en buffer SHAPE 1X (HEPES 100 mM, MgCl<sub>2</sub> 5 mM y NaCl 100 mM) a temperatura ambiente. La reacción se detuvo a distintos tiempos con buffer STOP (95% formamida, EDTA 5mM) y se resolvió en un gel de acrilamida:bisacrilamida (19:1) al 30%, el cual se expuso en una placa de almacenamiento fluorográfico y se visualizó con el equipo Typhoon TRIO® (GE Healthcare Lifesciences). La adición del grupo acilo se evidenció al comparar el cambio de migración electroforética entre el RNA tratado con NAI o con el vehículo (DMSO).

### **Evaluación *in cellulo* de la acilación del reactivo NAI**

Para ensayos individuales, se utilizó una caja Petri desechable P100 con la línea celular HeLa (ATCC® CCL-2™) en confluencia. Las células se despegaron, lavaron y resuspendieron en 10 mL de PBS; 4.5 mL de la mezcla previa se transfirieron a dos tubos cónicos de 15 mL. La suspensión de células

se centrifugó a 200 RFC durante 3.5 minutos. El sobrenadante se desechó y la pastilla obtenida se resuspendió en 450  $\mu$ L de PBS 1X Gibco™ (KCl 2.6 mM  $\text{KH}_2\text{PO}_4$  1.47 mM NaCl, 137.93 mM y  $\text{Na}_2\text{HPO}_4$  8.09 mM). Posteriormente, se añadió 50 $\mu$ L de NAI 1M o 50  $\mu$ L de DMSO (vehículo). La mezcla se incubó a 37°C durante 20 minutos, se centrifugó la mezcla a 200 RCF durante 3.5 minutos. Finalmente, el RNA modificado y el RNA control se purificaron con el reactivo TRIzol™ (Invitrogen™) siguiendo las instrucciones del proveedor. El RNA obtenido se resolvió en geles de agarosa al 2%.

### **Determinación estructural del aptámero Sc5-c3<sup>54</sup>**

Se utilizaron 17 picomoles del aptámero de RNA Sc5-c3 (amablemente donado por el Dr. Palomino-Vizcaino) disuelto en agua libre de nucleasas que se incubaron a 37°C o 70°C con dos distintas concentraciones finales de NAI (50 mM y 100 mM) en presencia de buffer SHAPE 1X y el inhibidor de ribonucleasas RiboLock™ (4 U); en paralelo el RNA se trató con DMSO (control negativo). El RNA modificado y el control negativo se precipitaron, lavaron y resuspendieron en 20  $\mu$ L de agua libre de nucleasas. Diez  $\mu$ L del RNA del paso anterior se utilizaron como molde para montar reacciones de reverso transcripción en un volumen final de 20  $\mu$ L. Inicialmente se incubó el RNA, 1  $\mu$ L del cebador FAM-TAA TAC GAC TCA CTA TAG-3' 500nM y 0.5 mM de cada dNTP a 65°C por 5 minutos y 4°C por 2 minutos. Tras la incubación, se añadió buffer RT 5X (Tris-HCl 250 mM (pH 8.3), KCl 375 mM,  $\text{MgCl}_2$  15 mM; Invitrogen™) a una concentración final 1X (Invitrogen™), 1 $\mu$ L de DTT 0.1M y 100 U de SuperScript III (Invitrogen™). La mezcla se incubó 50 minutos a 52°C y 15 minutos a 70°C. Los productos de la reacción se resolvieron en un gel de poliacrilamida (19:1) al 18% y urea 7M.

### **Análisis densitométrico de SHAPE y modelado estructural del aptámero Sc5-c3**

Los geles generados en el inciso anterior se escanearon usando el equipo Thyphoon TRIO®, utilizando los parámetros para el fluoróforo fluoresceína

(FAM), con los parámetros “High” y PMT=600V. Las bandas obtenidas se cuantificaron con el software Image Studio™ Lite (LI-COR, Inc) y se normalizaron siguiendo la regla 2%-8%<sup>55</sup>. Los valores de reactividad obtenidos (Ver Anexo) se usaron como valores de pseudo-energía libre en el servidor RNAprobing, usando la metodología de Zarringhalam<sup>56</sup>. El modelo estructural obtenido se visualizó y editó en el software StructureEditor (Mathews Lab).

### **Síntesis *in vitro* de los precursores de *let-7* y sus mutantes**

Para la generación de los precursores se partió de moldes de DNA sintetizados por T4 oligo® (ADN Sintetico S.A.P.I. de C.V.), los cuales contienen el promotor para la RNA polimerasa del fago T7 y la secuencia específica para cada precursor de la familia *let-7* (Tabla 1):

<b>Precursor</b>	<b>Secuencia</b>
<i>let-7c</i>	TAATACGACTCACTATAGGGCATCCGGGTTGAGGTAGT AGGTTGTATGGTTTAGAGTTACACCCTGGGAGTTAACT GTACAACCTTCTAGCTTTCCTTGGAGC
<i>let-7f-1</i>	TAATACGACTCACTATAGGGTCAGAGTGAGGTAGTAGA TTGTATAGTTGTGGGGTAGTGATTTTACCCTGTTCAGGA GATAACTATACAATCTATTGCCTTCCCTGA
<i>let-7g</i>	GCACGTAATACGACTCACTATAGGGAGGCTGAGGTAGT AGTTTGTACAGTTTGAGGGTCTATGATACCACCCGGTA CAGGAGATAACTGTACAGGCCACTGCCTTGCCA
<i>miR-98</i>	CGTAATACGACTCACTATAGGGAAGGATTCTGCTCATGC CAGGGTGAGGTAGTAAGTTGTATTGTTGTGGGGTAGGG ATATTAGGCCCAATTAGAAGATAACTATACAACTTACT ACTTTCCTGGTGTGTGGCATATTCA

<i>let-7a-1</i>	TAATACGACTCACTATAGGGTGGGATGAGGTAGTAGGT TGTATAGTTTTAGGGTCACACCCACCACTGGGAGATAA CTATACAATCTACTGTCTTTCCTA
<i>let-7f-1 mut</i>	TAATACGACTCACTATAGGGTCAGAGTGAGGTAGTAGA TTGTATAGTTGTGGGGTAGTGATTTTAaaaTGTTTCAGGA GATAACTATACAATCTATTGCCTTCCCTGA
<i>let-7g mut</i>	TAATACGACTCACTATAGGGAGGCTGAGGTAGTAGTTT GTACAGTTTGAGGGaggATGATACCACCCGGTACAAtctGA TAACTGTACAGGCCACTGCCTTGCCA

**Tabla 1. Moldes de DNA para la generación de los precursores de *let-7*** En rojo se remarca la secuencia del promotor para T7, en negro la secuencia del precursor y en letras minúsculas, en caso de tenerlas, las mutaciones generadas a la secuencia silvestre del precursor.

Los moldes se purificaron mediante un gel de acrilamida-bisacrilamida (19:1) 10%, urea 7M con el fin de eliminar productos ramificados e incompletos de la síntesis. Posteriormente los moldes se amplificaron mediante PCR con los siguientes cebadores:

Nombre del cebador	Secuencia
let-7f-1 FWD	TAATACGACTCACTATAGGGTCAGAGTG
let-7f-1 REV	CAGGGAAGGCAATAGATTGTATAG
let-7c FWD	GCAGGTAATACGACTCACTATAGGG
let-7c REV	GCT CCAAGGAAAGCTAGAAGG TTG TAC
let-7g FWD	GCACGTAATACGACTCACTATAGGGAGG-
let-7g REV	TGG CAAGGCAGTGGCCTG
miR-98 FWD	CGTAATACGACTCACTATAGGGAGGATTCTGC



miR-98 REV	GAATATGCCACACACCAGGGGAAAG
let-7a-1 FWD	TAATACGACTCACTATAGGGTG
let-7a-1 REV	TAGGAAAGACAGTAGATTGTATAG

**Tabla 2. Cebadores específicos para la amplificación de pre-let-7 mediante PCR.**

El producto de PCR se utilizó como molde para generar RNA *in vitro* de acuerdo con el siguiente protocolo:

Reactivo	Volumen
Agua libre de nucleasas	18 $\mu$ L
Buffer de Transcripción 5X	12 $\mu$ L
rNTPs 2.5 mM	3 $\mu$ L
DTT 0.1M	5 $\mu$ L
RiboLock 4U/ $\mu$ L	1 $\mu$ L
Producto PCR	20 $\mu$ L
T7 RNAPol 20U/ $\mu$ L	1 $\mu$ L

**Tabla 3. Componentes para transcripción *in vitro*.**

El RNA generado se purificó mediante un gel de acrilamida-bisacrilamida (19:1) al 8%, urea 7M con el fin de eliminar productos espurios de la reacción de transcripción

### **Clonación de los precursores de let-7**

Dada la naturaleza de la técnica SHAPE, es necesario un sitio de anclaje para un cebador, el cual permita realizar la reacción de reverso-transcripción. Dicho sitio de anclaje se puede encontrar dentro del RNA problema; sin embargo, esto impide obtener la información estructural de los nucleótidos que comprenden al

sitio de anclaje. Por este motivo, cualquier RNA problema se puede clonar en un vector adecuado, el cual permita generar un transcrito más largo y utilizar un sitio de anclaje exterior a la secuencia a analizar, obteniendo así toda la información estructural del RNA problema.

El molde de DNA de los precursores de *let-7f-1*, *let-7c*, *let-7g* y *miR-98* se amplificó mediante PCR añadiendo los sitios EcoR1 y BamH1 en los extremos 5´ y 3´ de los cebadores, como se muestra en la siguiente tabla:

<b>Nombre cebador</b>	<b>Secuencia</b>
Eco-let7g	ACGAATTCATTTAAGGCTGAGGTAGTAGTTT
Bam-let7g	TTGGATCCTGGCAAGGCAGTGG
Eco-mir98	ACGAATTCATTTAAGGATTCTGCTCATGC
Bam-mir98	TTGGATCCTGAATATGCCACACACC
Eco-let7f1	AGGAATTCTATTATCAGAGTGAGGTAGTAG
Bam-let7f1	CAGGATCCTCAGGGAAGGCAATAG
Eco-let7c	AGGAATTCGCATCCGGGTTGAG
Bam-let7c	CAGGATCCGCTCCAAGGAAAGCTAG

**Tabla 4.** Cebadores para la amplificación de Pre-let-7 con sitios EcoRI y BamHI.

Los productos de PCR se digirieron en presencia de las enzimas de restricción EcoR1 y BamH1 en presencia de Buffer Tango 2X (Tris-acetato 33 mM pH 7.9, Mg(CH<sub>3</sub>COO)<sub>2</sub> 10 mM, K(CH<sub>3</sub>COO) 66mM y BSA 0.1 mg/mL (Thermo Scientific™). El producto de la digestión se purificó mediante precipitación con etanol. En paralelo, el vector pCAP-H12-MS2h se digirió con el mismo protocolo previamente descrito<sup>57</sup>. El producto de la digestión se cargó en un gel de agarosa al 1.0%, se corrió a 100V por 60 minutos, se recortó la banda

correspondiente y se purificó con el kit NZYGelpure (NZYTech©, Lda) siguiendo las instrucciones del proveedor.

La reacción de ligación se realizó con T4 DNA ligasa (Thermo Scientific™) siguiendo las instrucciones del proveedor. Específicamente, se usaron 100 ng del vector y una relación molar 1:7 vector-inserto. Las construcciones generadas (pCAP-let-7c, pCAP-let-7f-1, pCAP-miR-98 y pCAP-let-7g) se verificaron mediante secuenciación automática (Macrogen Inc®), ver Anexo.

### **Transcripción *in vitro* de los precursores de let-7 partiendo de las construcciones pCAP-let-7c, pCAP-let-7f-1, pCAP-miR-98 y pCAP-let-7g**

Para la transcripción *in vitro*, se digirieron 3 µg de los vectores correspondientes con la enzima Sph1. La digestión total del DNA se verifico en un gel de agarosa al 1.0% y el resto de la reacción de digestión se purifico mediante fenol-cloroformo. El DNA purificado se resuspendió en 67 µL y se utilizó como molde para la reacción de transcripción, siguiendo el siguiente protocolo

<b>Reactivo</b>	<b>Volumen</b>
DNA pCAP-let-7 digerido	67 µL
Buffer de transcripción 10X	12 µL
DTT 0.1M	12 µL
rNTPs 2.5 mM	24 µL
RNase OUT™ 40U/µL	2 µL
T7RNAPol	5 µL

**Tabla 5. Transcripción *in vitro* para pre-let-7**

La reacción de transcripción se incubó 4 horas a 37°C y se trató con 2 µL de RQ DNase I (Promega Corporation) durante 30 minutos a 37°C para eliminar el molde de DNA. La síntesis de RNA se verificó en un gel de agarosa al 1.5%, se precipitó con 2 µl de glucógeno 2µg/µL, 0.1 volúmenes de NaCl 1M y etanol para resuspenderse en 40 µL de agua libre de nucleasas. La imagen digital de cada gel se analizó con el software Image J utilizando como referencia el marcador de peso molecular φ29 (patente pendiente CBMSO, Madrid, España) o el marcador GeneRuler 1kb plus DNA Ladder (Thermo Fisher Scientific Inc.) para obtener la concentración de cada RNA sintetizado.

### **Modificación del RNA con el reactivo NMIA**

Para la reacción de acilación, 2 picomoles de cada RNA generado en el inciso anterior se diluyeron en buffer TE 0.5X (EDTA 0.5 mM, Tris-HCl 10 mM, pH final 8.0) en un volumen final de 12 µL. La mezcla se incubó durante 2 minutos a 95°C e inmediatamente se transfirieron a 4°C por 2 minutos. Después, se añadieron 6 µL de buffer SHAPE 3.3X y se incubó a 37°C por 20 minutos (paso de plegamiento). Enseguida, se adicionaron 2 µL de NMIA 65 mM o DMSO (control negativo) a la mezcla anterior. La mezcla resultante se incubó a 37°C por 45 minutos. Finalmente, el RNA modificado y el RNA control se precipitaron, lavaron y se resuspendieron en 10 µL TE 0.5X.

### **Reverso transcripción del RNA modificado**

La totalidad del RNA modificado y RNA control de la sección anterior se utilizaron como molde para generar cDNA. El RNA se incubó con 0.5 µL del cebador NED-TAGCCTTATGCAGTTGCTCTCC<sup>58</sup> (4 µM) a 65°C por 5 minutos, 35°C por 5 minutos y 4°C por 2 minutos. La reacción de reverso transcripción se llevó a cabo a un volumen final de 16 µL usando buffer RT (Invitrogen™), 10 U de RNase OUT Invitrogen™, 1 mM de cada dNTP y 60 U de SuperScript III (Invitrogen™). La mezcla se incubó 30 minutos a 52°C y 15 minutos a 70°C. Los productos de la reacción se resolvieron por electroforesis capilar (Unidad de Genómica IPBLN, Granada, España). Como marcador de peso molecular

se utilizaron 2 pmoles del RNA correspondiente sintetizado *in vitro* y se incubaron con 1  $\mu$ L del cebador FAM-TAGCCTTATGCAGTTGCTCTCC<sup>58</sup> (4  $\mu$ M) a 65°C por 5 minutos, 35°C por 5 minutos y 4°C por 2 minutos. La reacción de reverso transcripción se llevó a cabo a un volumen final de 20  $\mu$ L usando buffer RT 1X (Invitrogen™), 10 U de RNase OUT Invitrogen™, 1 mM de cada dNTP, ddCTP 0.1 mM y 60 U de SuperScript III (Invitrogen™).

## **Análisis de electroferogramas y modelado estructural**

Los electroferogramas generados fueron analizados con el software QuSHAPE<sup>59</sup>. Los valores de reactividad obtenidos se usaron como valores de pseudo-energía libre en el software RNAstructure<sup>60</sup>, con el cual se generaron los modelos de estructura secundaria de los RNAs precursores de *let-7f-1*, *let-7c*, *let-7g* y *miR-98*. Las estructuras obtenidas fueron visualizadas y editadas en el software StructureEditor (Mathews Lab). La estructura secundaria obtenida sirvió como modelo para generar modelos tridimensionales usando el software RNAComposer<sup>61</sup> y editándolos con el software UCSF Chimera<sup>62</sup>

## **Análisis estructural de los precursores de let-7**

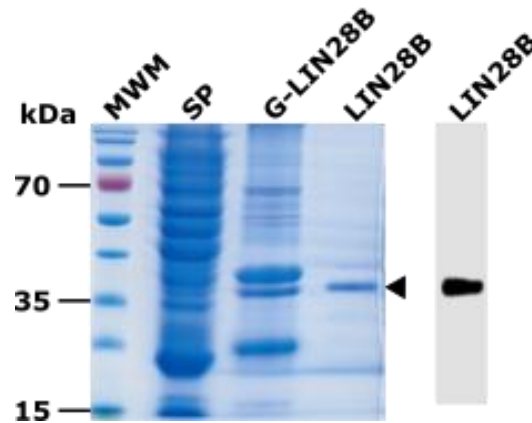
Con el fin de comparar objetivamente los resultados obtenidos, las estructuras generadas mediante SHAPE se analizaron con las herramientas bioinformáticas BEAGLE y LocARNA<sup>63,64</sup>. La herramienta BEAGLE genera alineamientos pareados con el fin de identificar dos estructuras significativamente similares. Se consideró el valor “Z-score” como la medida de referencia para el análisis estadístico, siempre que el valor obtenido fuera igual o mayor a 3.0. De forma complementaria se realizó un análisis extra con LocARNA, una herramienta que permite realizar alineamientos múltiples de acuerdo con las características estructurales de las secuencias analizadas. Sumado al alineamiento, también se genera un árbol guía, el cual establece las similitudes estructurales en grupos jerárquicos; y una estructura secundaria consenso para los RNAs evaluados.

## Inducción de proteína recombinante LIN28B

La inducción de la proteína LIN28B se realizó a partir del plásmido y metodología previamente reportada<sup>65</sup>, incorporando algunos cambios. La construcción pGST-.LIN28B se transformó en *Escherichia coli* Rosetta™ [F<sup>-</sup> *ompT hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup>m<sub>B</sub><sup>-</sup>) gal dcm pRARE<sup>2</sup>(Cam<sup>R</sup>)*] las cuales se incubaron por 30 minutos a 4°C. Tras la incubación, la mezcla se sometió a un choque térmico a 42°C por 2 minutos e inmediatamente fueron trasladadas a 4°C por tres minutos. Posteriormente las bacterias fueron propagadas en cajas Petri con agar LB con ampicilina 200 µg/mL y cloranfenicol 50 µg/mL, las cuales se incubaron a 37°C por 18 horas.

En un matraz de 200mL con deflectores se agregaron 50mL de medio LB adicionado con ampicilina 200 µg/mL y cloranfenicol 50 µg/mL, al cual se añadieron 5 colonias de bacterias transformadas y se incubó a 37°C por 18 horas en agitación constante a 250 rpm. Posteriormente, el cultivo se fraccionó equitativamente en 5 matraces de 200mL con deflectores y se añadió medio con antibióticos hasta un volumen final de 100 mL por matraz. El cultivo se incubó para su crecimiento a 37°C por 3.5 horas con agitación constante a 250 rpm. La inducción se realizó añadiendo IPTG a una concentración final de 1mM en cada frasco e incubando a 37°C por 1.5 horas con agitación constante a 250 rpm. Tras la inducción, las bacterias se recuperaron por centrifugación y se resuspendieron en buffer de lisis (NaCl 100 mM, Tris 20 mM, EDTA 1 mM, IGEPAL 0.5%, AEBSF 0.5 mM y una tableta del inhibidor cOmplet™ (Roche Diagnostics GmbH) por cada 50 mL de buffer de lisis) para lisarse con un equipo Cole-Parmer® Ultrasonic Processor, usando 5 pulsos de 30 segundos a 80% de amplitud. Tras la lisis, el extracto celular se centrifugó a 15,900 RCF durante 3 minutos y la fracción soluble se clarificó a través de un filtro de 0.45 µm. La fracción soluble clarificada se incubó con 50µL de perlas glutatión-agarosa por cada 1.4 mL de fracción soluble a 4°C por una hora en agitación orbital constante. Posteriormente las perlas se lavaron 5 veces con PBS 1X y se incubaron con 250 unidades de Benzonase® (Sigma KGaA) y MgCl<sub>2</sub> 2 mM

durante 1 hora a 37°C. Tras la incubación, se realizaron 10 lavados con PBS 1X y se incubó con 3 unidades de trombina (Invitrogen™) a 4°C durante 48 horas. Finalmente, el sobrenadante obtenido tras la digestión se cuantificó mediante colorimetría usando el reactivo Pierce BCA protein assay kit (Thermo Fisher Scientific Inc.) y se verificó mediante SDS-PAGE e inmuno-detección en fase sólida o Western Blot (Figura 8).



**Figura 8. SDS-PAGE y western blot de LIN28B recombinante.** Gel desnaturalizante con el lisado total (SP), la proteína recombinante LIN28B unida a las perlas de glutación agarosa (G-LIN28B) y la proteína LIN28B purificada (flecha negra). A la derecha, Western Blot de LIN28B purificada. MWM, marcador de peso molecular.

## Ensayos de retardamiento electroforético (EMSA)

Para los ensayos EMSA, se de-fosforilaron los precursores de *let-7a-1*, *let-7f-1mut* y *let-7g* sintetizados *in vitro* con la fosfatasa CIAP (Thermo Fisher Scientific Inc.). En seguida, se marcó el RNA con la cinasa T4 PNK (Thermo Fisher Scientific Inc.) y  $\gamma$ -[<sup>32</sup>P]-ATP (3000 Ci/mmol) (Perkin Elmer, Inc., Waltham, MA), siguiendo las instrucciones del proveedor. Las mezclas de interacción constaron de Tris-HCl 20mM pH 8.6, NaCl 150 mM, MgCl<sub>2</sub> 5 mM, glicerol al 10%, DTT 5 mM, 2 U/ $\mu$ L de RiboLock RNase inhibitor (Thermo Fisher), 2 pmoles de cada RNA marcado radioactivamente y concentraciones crecientes de la proteína LIN28B, todo en un volumen final de 20  $\mu$ L. Las reacciones se incubaron por una hora a temperatura ambiente en agitación rotatoria. Tras la incubación, las mezclas de interacción se resolvieron en geles

nativos de acrilamida-bisacrilamida (37.5:1) al 8% en TBE 0.5X a un voltaje de 100V, durante una hora a temperatura ambiente. En seguida, los geles se incubaron en solución de secado (MeOH 30% y glicerol 5%) durante 15 minutos a temperatura ambiente. Los geles se secaron durante 40 minutos a 80°C en un equipo Gel-Dryer Model 583 (Bio-Rad Laboratories, Inc.) conectado al vacío.

### **Análisis de los ensayos EMSA.**

Los geles secos se expusieron a una placa de almacenamiento fluorográfico, la cual se reveló con un equipo Typhoon TRIO® (GE Healthcare Lifesciences). Las imágenes se analizaron con el programa Image Studio™ Lite (LI-COR, Inc.), cuantificando el complejo retardado en función del complejo libre. Los datos obtenidos se graficaron en función de la concentración de proteína LIN28B agregada y se ajustó al modelo de regresión no lineal de Hill ( $Y=B_{max} \cdot X^h / (K_d^h + X^h$ ; donde :  $B_{max}$ = unión específica máxima,  $K_d$ = constante de disociación y  $h$ =pendiente de Hill), obteniendo los valores de  $K_d$  y  $R^2$  (coeficiente de correlación) para cada precursor evaluado (GraphPad Software v7.0, La Jolla California USA, [www.graphpad.com](http://www.graphpad.com)).

### **Estandarización de la transfección de pre-let-7f-1 en células HEK-293TT**

Las construcciones previamente mencionadas se encuentran dentro de un vector de clonación que contiene el origen de replicación del virus SV40, el cual es responsivo al antígeno T grande presente en la línea celular HEK-293TT. Esta característica permite que cada célula transfectada replique eficientemente el vector y lo segregue a las células hijas. Además, la secuencia que codifica para pre-let-7f-1 se clonó en marco con la secuencia que codifica para luciferasa; esto permite que se genere un transcrito de fusión pre-let-7f-1-luciferasa el cual contendrá en su extremo 5' un CAP y en el extremo 3' una secuencia de poli-adeninas. Finalmente, las células HEK-293TT provenientes de riñón humano embrionario, las cuales pueden transfectarse fácilmente y



permiten una alta expresión de transgenes y la formación de partículas virales<sup>66</sup>, generando una alta cantidad de transcrito estable en cada célula transfectada, lo cual es necesario para realizar un mapeo estructural *in cellulo*; lo que brinda la capacidad de seguir la transfección mediante la actividad de luciferasa.

La estandarización comenzó sembrando 750,000 células de la línea celular HEK-293TT en una caja Petri desechable p60 con 3 mL de medio DMEM suplementado con 5% de suero fetal bovino GIBCO™ (Thermo Fisher Scientific), estreptomina/penicilina 1% (GE Healthcare, Piscataway NJ), gentamicina 10 mg/mL (Thermo Fisher Scientific) y higromicina B 300µg/mL (Sigma-Aldrich Inc). A las 18 horas se eliminó el medio DMEM y se adicionó 2.5 mL de medio Opti-MEM® (Thermo Fischer Scientific). 6 µg de la construcción pCAP-let-7f-1 se mezclaron con 250µL de medio Opti-MEM® y 7µL del agente transfectante Lipofectin™ (Thermo Fischer Scientific) se mezclaron con 250 µL de medio Opti-MEM®. Ambas soluciones (A y B) se incubaron independientemente durante 30 minutos a temperatura ambiente. Tras la incubación, las soluciones se mezclaron en una sola (mezcla de transfección) y se incubaron por 15 minutos a temperatura ambiente. Inmediatamente, la mezcla de transfección se adicionó a las células, las cuales se incubaron a 37°C y 5% CO<sub>2</sub> durante 6 horas. Tras la incubación se retiró el medio Opti-MEM® y se cambió por medio DMEM suplementado con 5% de suero fetal bovino GIBCO™, estreptomina/penicilina 1%, gentamicina 10 mg/mL y higromicina B 300 µg/mL. La viabilidad de celular se evaluó mediante azul tripano en un equipo TC10™ *Automated Cell Counter* (Bio-Rad Laboratory Inc.)

### **Evaluación de la actividad de Luciferasa en células 293TT transfectadas con pCAP-let-7f-1.**

Las células transfectadas se despegaron con una solución de tripsina al 0.05% (Thermo Fisher Scientific) diluida en Verseno 0.5X (Thermo Fisher Scientific) y PBS 0.5X, y se contaron en un equipo TC10™. Inicialmente, se tomó un número

arbitrario de células en un tubo de 1.5 mL, a las cuales se agregó 400 µL del reactivo Bright-Glo™ (Promega Corporation); la mezcla se agitó vigorosamente hasta observarse una solución homogénea. La actividad de luciferasa se midió en un equipo GloMax 20/20 (Promega Corporation). Posteriormente, este mismo procedimiento se utilizó para medir la actividad de luciferasa en todas las células obtenidas a partir de una caja de Peri desechable p60 (aproximadamente 3 millones de células).

### **Evaluación del transcrito de fusión let-7f-1—luciferasa en células transfectadas.**

Cuarenta y dos horas post-transfección, se obtuvo el RNA total de las células transfectadas utilizando el reactivo TRIzol™ (Thermo Fisher Scientific) siguiendo las instrucciones del proveedor. Con el RNA purificado se evaluó la presencia del transcrito de fusión let-7f-1-luciferasa mediante RT-qPCR con el siguiente protocolo:

<b>Componente</b>	<b>Volumen</b>
Agua libre de nucleasas	9 µL
dNTPs 10 mM	1 µL
Oligonucleótido RT Luc (4 µM)	0.5 µL
RNA (698.2ng/µL)	4 µL

**Tabla 6.** Mezcla de hibridación molde-cebador para reverso-transcripción

La mezcla de reacción se incubó a 65°C por 5 minutos, 35°C por 5 minutos y 4°C por 5 minutos; inmediatamente se agregaron los siguientes reactivos:

<b>Componente</b>	<b>Volumen</b>
Buffer RT 5X	4 µL

DTT 0.1 M	1 $\mu$ L
SuperScript®III 200U/ $\mu$ L	0.5 $\mu$ L

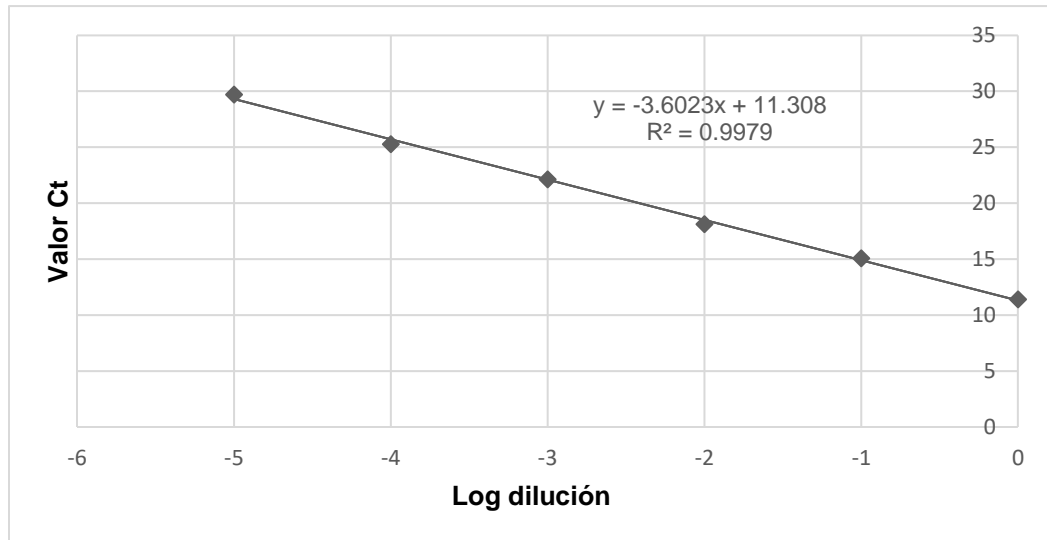
**Tabla 7.** Reacción de reverso transcripción

La mezcla final se incubó 52°C durante 30 minutos seguido de un paso de inactivación a 70°C por 15 minutos. El cDNA obtenido se utilizó para realizar las reacciones de RT-qPCR siguiendo el siguiente protocolo

<b>Componentes</b>	<b>Volumen</b>
Agua libre de nucleasas	5.1 $\mu$ L
Maxima SYBR green <i>Master Mix 2x</i>	6.25 $\mu$ L
Fwd let7f1 10 $\mu$ M	0.3 $\mu$ L
Rev luc10 $\mu$ M	0.3 $\mu$ L
cDNA	0.6 $\mu$ L

**Tabla 8.** Componentes de la PCR en tiempo real (qPCR)

Se determinó la eficiencia de la reacción de la qPCR realizando diluciones seriadas del cDNA y obteniendo el Ct para cada condición (Gráfica 1). Con los datos obtenidos de la regresión lineal, se calculó la eficiencia de reacción con la ecuación “Eficiencia= $-1+10^{(-1/\text{pendiente})}$ ”, obteniéndose un valor del 90%<sup>67</sup>.



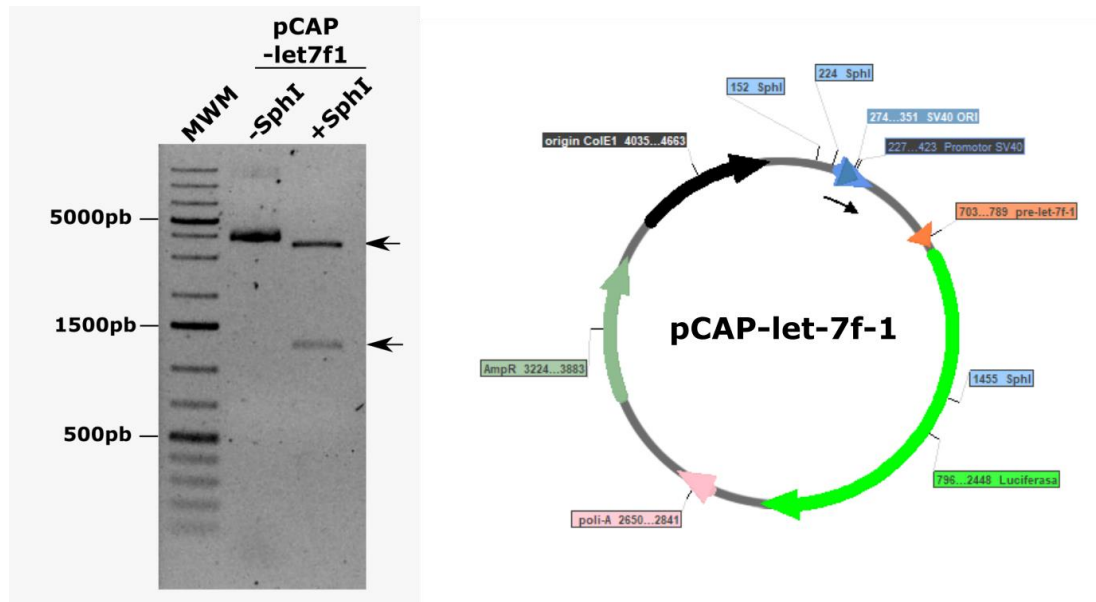
**Gráfica 1. Eficiencia de reacción para la qPCR de let-7f-1—luciferasa.** Valores Ct en función de la cantidad de cDNA usado en cada reacción. La pendiente de la gráfica se utilizó para obtener la eficiencia de reacción con la formula  $E = -1 + 10^{(-1/\text{pendiente})}$

Además, se evaluó al transcrito U6 de acuerdo al protocolo previamente descrito <sup>68</sup>.

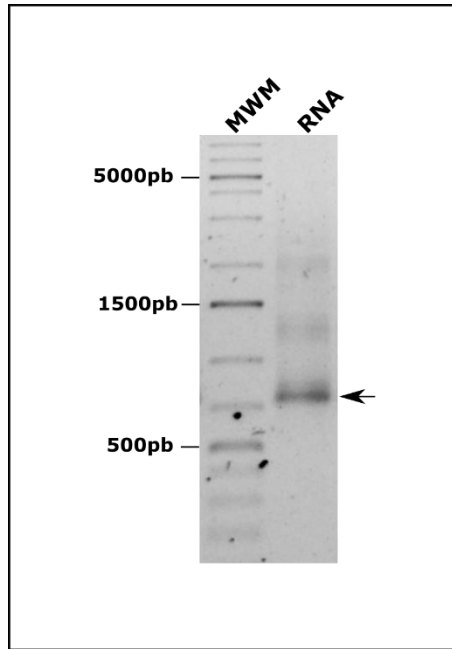
## Resultados

### Síntesis *in vitro* de pre-let-7

Para determinar la estructura secundaria de pre-let-7f-1, pre-let-7c, pre-let-7g y pre-miR-98 se partió de las construcciones pCAP-let-7c, pCAP-let-7f-1, pCAP-miR-98 y pCAP-let-7g (ver Materiales y Métodos) las cuales se digirieron con la enzima Sph1. Obteniendo el patrón de restricción esperado, correspondiente a bandas de aproximadamente 3400 pb y 1200 pb. (Figura 9). Confirmada la digestión, los plásmidos digeridos se purificaron y se utilizaron para la síntesis *in vitro* de RNA (ver Materiales y Metodos), obteniendo una banda mayoritaria de aproximadamente 750 pb (Figura 10).



**Figura 9. Construcción pre-let-7f-1 y digestión con SphI.** Izquierda, gel de agarosa representativo del plasmido pCAP-let-7f-1 digerido con SphI. Los fragmentos observados (flechas negras) corresponden a los pesos esperados, 3409 pb y 1231 pb respectivamente. Derecha, mapa correspondiente a la construcción pCAP-let.7f-1. MWM, marcador de peso molecular; -SphI, plásmido sin digerir; +SphI plásmido digerido.



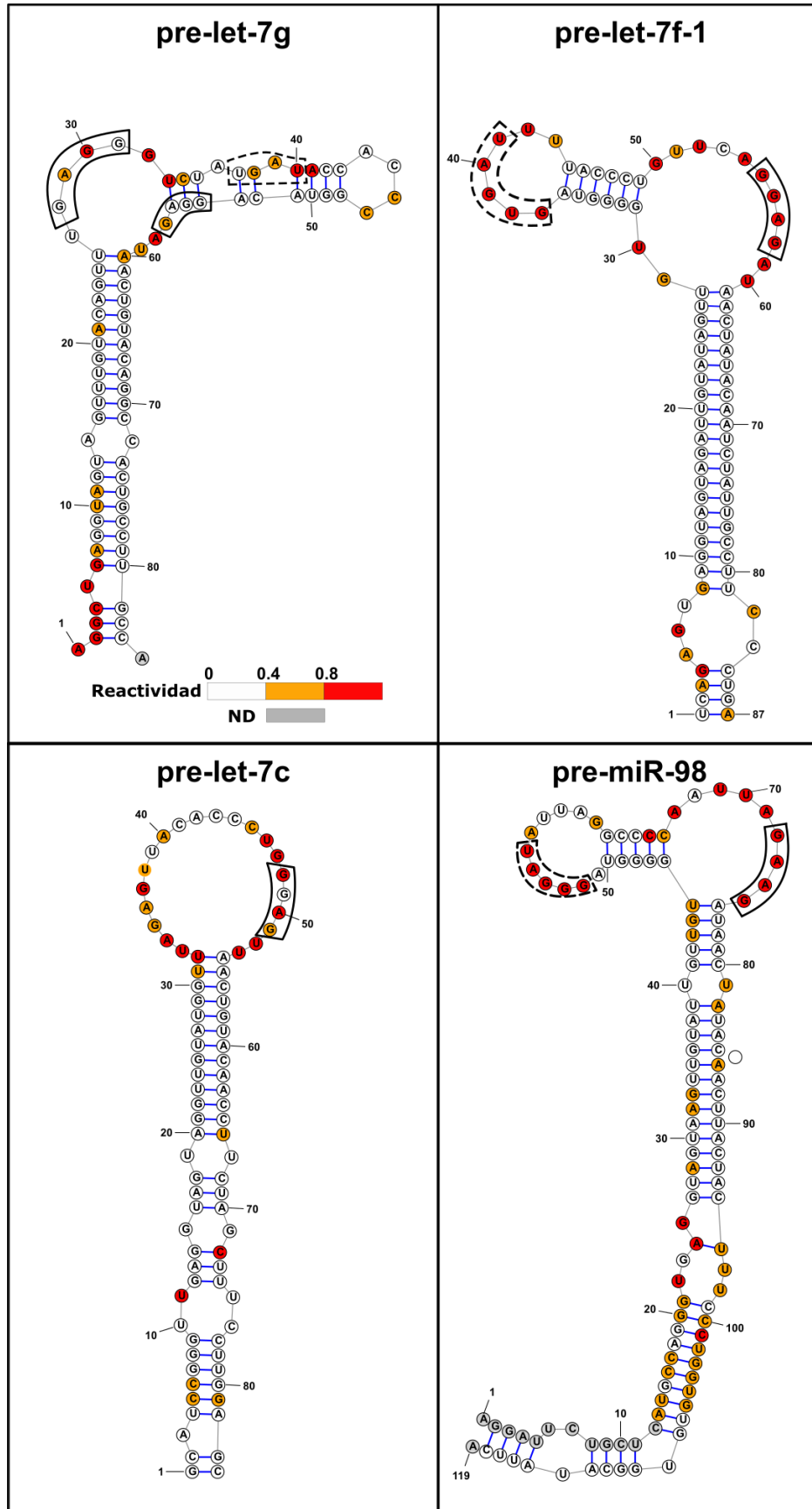
**Figura 10. Transcripción *in vitro* de pre-let-7f-1.** Gel de agarosa representativo de la transcripción *in vitro* de pre-let-7f-1. Se observó una banda mayoritaria de peso esperado (769 nt), correspondiente al RNA de pre-let-7f-1. MWM, marcador de peso molecular.

El RNA obtenido se purificó y cuantificó para determinar la estructura secundaria mediante la técnica SHAPE.

### **Caractrización estructural *in vitro* de pre-let-7**

La estructura de los precursores let-7f-1, let-7g, let-7c y miR-98 se obtuvo mediante la metodología SHAPE, utilizando el reactivo NMIA y resolviendo los productos de las reacciones mediante electroforesis capilar (ver Materiales y Métodos). Los valores de reactividad obtenidos se utilizaron para modelar la estructura secundaria de cada precursor (Figura 11). En todos los casos aquellos nucleótidos que presentaron alta reactividad, y por ende se encuentren preferentemente en ssRNA, se encuentran resaltados en colores rojo y amarillo; mientras que los nucleótidos sin reactividad se encuentran formando estructuras de doble cadena se representan con un fondo blanco. Particularmente, los modelos de pre-let-7f-1 y pre-miR-98 mostraron un patrón estructural común en el pre-E, compuesto de una burbuja cercana al extremo

5,´ en la cual se encuentra la estructura consenso UGAU/GNGAY (Cuadro punteado); seguida de un tallo y una burbuja cercana al extremo 3´, la cual contiene la secuencia consenso GGAG (cuadro continuo). Este mismo patrón burbuja-tallo-burbuja se observó en la estructura de pre-let-7g; sin embargo, la secuencia consenso UGAU/GNGAY se encuentra en el tallo del pre-E, flanqueada por dos secuencias GGAG, ambas presentes en la misma burbuja, pero en posiciones opuestas. Finalmente, la estructura de let-7c presentó una única burbuja con la secuencia consenso GGAG. Podemos observar que cada precursor presenta características estructurales únicas como pre-let-7c y pre-let-7g, o presentar características comunes como pre-miR-98 y pre-let-7f-1.





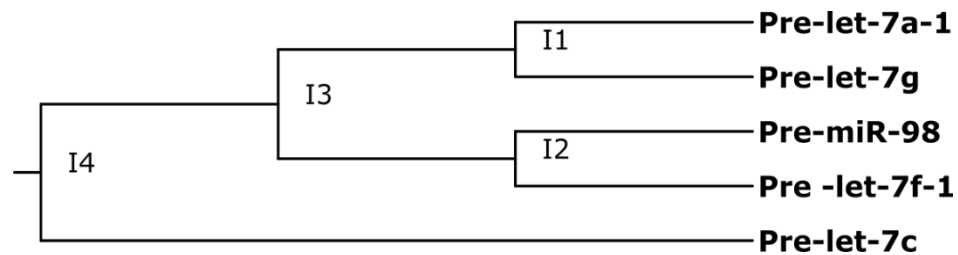
**Figura 11. Estructuras secundarias de pre-let-7.** Estructuras secundarias de pre-let-7g, pre-let-7c, pre-let-7f-1 y pre-miR-98 obtenidas mediante SHAPE. La escala de colores representa la reactividad de cada nucleótido. En líneas sólidas y punteadas se resalta las secuencias consenso GGAG y GNGAY/UGAU respectivamente. ND, nucleótidos sin dato experimental

Con el fin de evaluar de manera objetiva las características estructurales de los precursores de *let-7* obtenidas, se utilizaron las herramientas bioinformáticas BEAGLE y LocARNA (ver Materiales y Métodos), añadiendo la estructura de pre-let-7a-1, mapeada previamente mediante SHAPE<sup>69</sup> (Tabla 9). El resultado del análisis mediante BEAGLE indicó que pre-let-7f-1 y pre-miR-98 son los únicos RNAs que presentan características estructurales similares ( $Z$  score  $\geq 3$ ) mientras que pre-let-7a-1 es el precursor que menos similitud estructural presenta con el resto de los precursores mapeados ( $Z$  score negativo).

RNA-1	RNA-2	Identidad sec (%)	Identidad est (%)	Similitud est (%)	<i>p</i> -value	<i>Z</i> -score
miR-98	let-7f-1	50.42	32.77	63.03	0.0001	3.55
let-7f-1	let-7c	59.77	32.18	56.32	0.0008	2.872
miR-98	let-7c	52.1	16.81	52.1	0.0033	2.635
let-7g	let-7c	61.9	28.57	51.19	0.0068	2.25
let-7f-1	let-7g	72.41	27.59	59.77	0.0201	1.658
miR-98	let-7g	36.97	17.65	46.22	0.2073	0.785
let-7c	let-7a-1	63.1	0	13.1	0.8984	-1.36
let-7f-1	let-7a-1	63.22	0	13.79	0.9578	-1.97
let-7g	let-7a-1	59.52	7.14	15.48	0.9663	-2.36
miR-98	let-7a-1	46.22	0.84	10.08	0.9956	-3.09

**Tabla 9. Análisis de similitud estructural de pre-let-7.** Las estructuras experimentales obtenidas se evaluaron con la herramienta bioinformática BEAGLE. Un Z-score  $\geq 3$  indica estructuras estadísticamente similares.

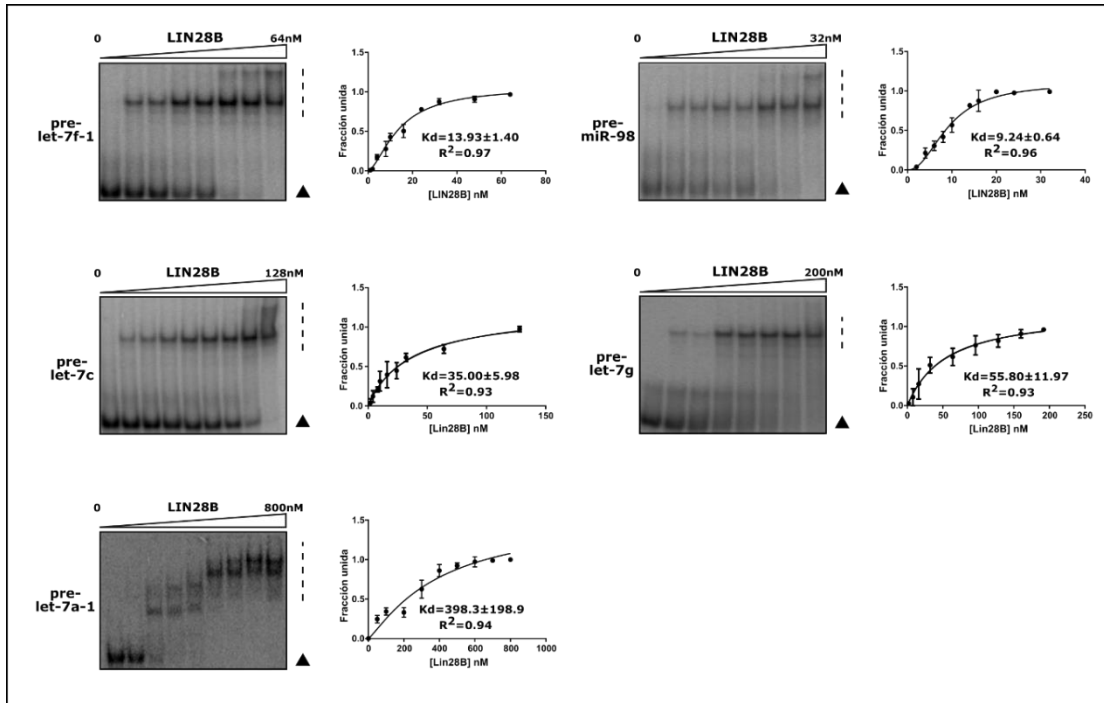
Como un análisis adicional se evaluaron las estructuras obtenidas con la herramienta LocARNA, la cual permite observar gráficamente la cercanía estructural de un conjunto de RNAs. El árbol guía obtenido (Figura 12) mostró que los precursores de miR-98 y let-7f-1 se encuentran estructuralmente relacionados. Pre-let-7g, pre-let-7a-1 y pre-let-7c se encontraron en subgrupos distintos, presentando una mayor similitud estructural pre-let-7g y pre-let-7a-1.



**Figura 12. Árbol guía generado con LocARNA.** Los precursores de let-7 se agruparon en subgrupos de acuerdo con su similitud estructural.

## Ensayos de interacción LIN28B:pre-let-7

Para confirmar que la estructura secundaria del RNA tenía un impacto en la interacción con LIN28B se comparó los resultados estructurales con los ensayos de interacción previamente descritos por nuestro grupo de trabajo<sup>65</sup>, recalculando las Kd con el modelo de Hill (ver Materiales y Metodos). A estos resultados se añadió el ensayo de interacción de pre-let-7a-1 con LIN28B y se calculó su Kd con el método previamente mencionado. (ver Materiales y Métodos).



**Figura 13. Ensayos EMSA entre LIN28B y pre-let-7.** Cada precursor fue incubado con concentraciones crecientes de LIN28B. Las constantes de disociación ( $K_d$ ) se determinaron a partir de tres ensayos independientes con el modelo matemático que presentara la mejor correlación ( $R^2$ ). Triangulo, precursor no unido; líneas punteadas, complejos RNA proteína.

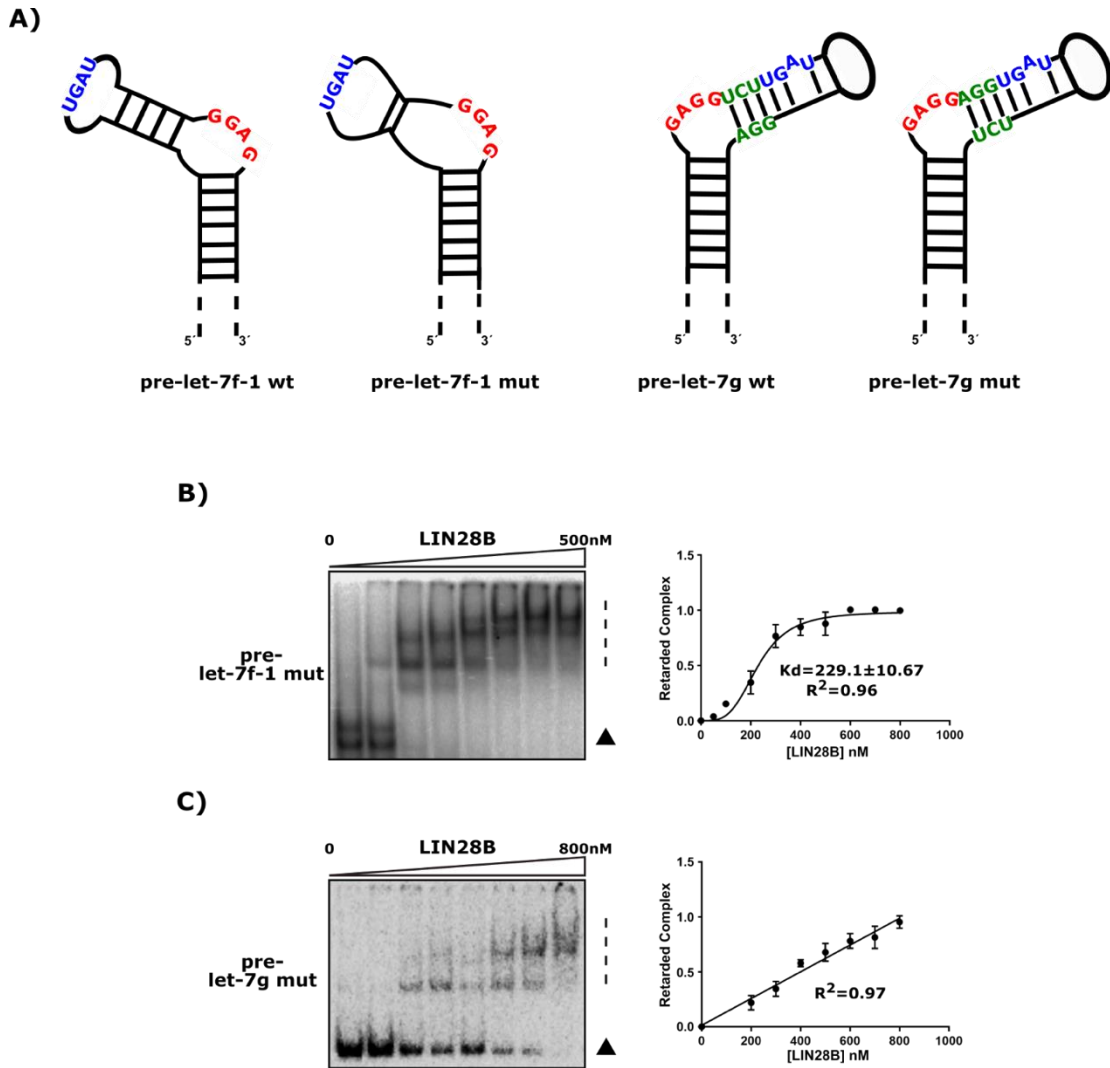
Como se puede observar en la Figura 13, cada precursor presentó diferentes constantes de disociación y un número distinto de complejos al interactuar con LIN28B. Además, aquellos precusores que presentaron la menor  $K_d$ , pre-let-7f-1 y pre-miR-98, presentaron estructuras estadísticamente similares. En contraste, aquellos precusores que presentaron estructuras alternas mostraron una mayor  $K_d$ .

### Evaluación de mutantes estructurales en la interacción pre-let-7:LIN28B

Con el fin de validar que la estructura secundaria tiene efecto en la interacción LIN28B:pre-let-7 se diseñaron mutantes de pre-let-7f-1 y pre-let-7g. La mutante pre-let-7f1mut contiene tres transversiones consecutivas (C/A) en las

posiciones 46-48, este cambio desestabiliza la región de doble cadena del pre-E. Para la mutante let-7gmut se realizó un intercambio de posición de los nucleótidos comprendidos dentro de las posiciones 33-35 (UCU) por aquellos nucleótidos dentro de las posiciones 55-56 (Figura 14, panel A). Como se puede observar, ambas mutantes conservan las secuencias consenso de unión GGAG y GNGAY/UGAU. Los oligonucleótidos molde con estas modificaciones se amplificaron y transcribieron *in vitro*; el RNA se purificó y se utilizó para realizar ensayos EMSA con concentraciones crecientes de la proteína recombinante LIN28B (ver Materiales y Métodos).

Como se muestra en la Figura 14 panel B, ambas mutantes de pre-let-7 tuvieron un impacto negativo en la interacción con LIN28B. Pre-let-7f-1 mut presentó una Kd de 229.1 nM, en comparación con el valor de 13.9 obtenido con la versión silvestre. Al graficar los valores obtenidos en los ensayos de pre-let-7g mut, se observó una gráfica lineal en función de la cantidad de proteína LIN28B agregada. Este comportamiento matemático ha sido ampliamente descrito como una interacción inespecífica, por lo cual el valor de Kd es inestimable e ineficaz<sup>70</sup>. En conclusión, se puede observar que cada cambio estructural afecta de manera distinta la afinidad pre-let-7:LIN28B.

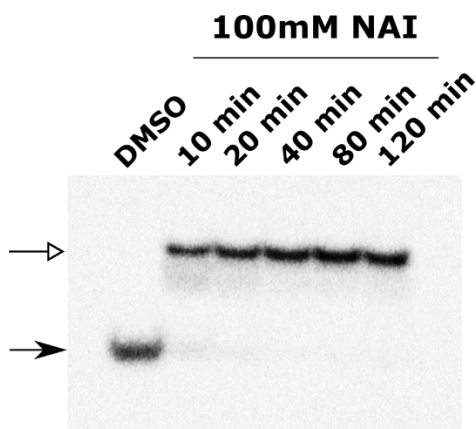


**Figura 14. Cambios estructurales en pre-let-7 comprometen su unión con LIN28B.** A) Caricatura de los precursores de let-7 silvestres (wt) comparadas con su versión mutante (mut). Ensayos de unión (EMSA) de los precursores B) let-7f-1 mut y C) let-7g mut.  $K_d$ , constante de disociación;  $R^2$  constante de disociación. Ensayos por triplicado.

## Síntesis y verificación del reactivo NAI

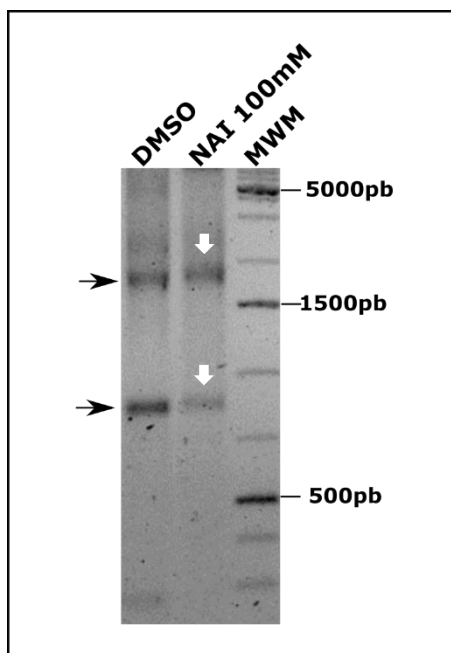
Para la caracterización *in cellulo* es necesario utilizar un reactivo de SHAPE distinto al NMIA ya que éste es incapaz de atravesar la membrana celular. Para ello, se realizó la síntesis del reactivo NAI y su verificación como se describió originalmente<sup>53</sup> (Ver Materiales y Métodos). La verificación del reactivo NAI se realizó con  $\gamma$ -[<sup>32</sup>P]-ATP tratado con 100 mM de NAI, el cual sufrió un cambio en

su migración electroforética en comparación con  $\gamma$ -[ $^{32}\text{P}$ ]-ATP tratado con DMSO (vehículo), esto debido a la adición del grupo acilo donado por reactivo NAI. La reacción de acilación se siguió a través del tiempo, observando que la reacción modifica por completo al  $\gamma$ -[ $^{32}\text{P}$ ]-ATP a los 40 minutos (Figura 15).



**Figura 15. Reacción de acilación con el reactivo NAI.** Cinética de modificación de  $\gamma^{32}\text{P}$ -ATP con 100 mM de NAI. La cinética de reacción es completa a los 40 minutos de incubación a 37°C. Flecha sólida ATP sin modificar, flecha blanca ATP modificado.

Además de la verificación *in vitro*, se realizó una validación rápida de las características *in cellulo* del reactivo NAI. Para ello se trataron células HeLa con 100 mM del reactivo NAI o DMSO, se extrajo el RNA total de cada tratamiento y se evaluó su migración en un gel de agarosa (ver Materiales y Métodos). Se observó que los RNA ribosomales (flechas negras) sufrían un cambio sutil y reproducible en su migración electroforética (Figura 16). Ambas evidencias, *in vitro* e *in cellulo* validaron la síntesis y las características previamente reportadas del reactivo NAI, lo cual permitió su uso para la determinación estructural de RNAs.

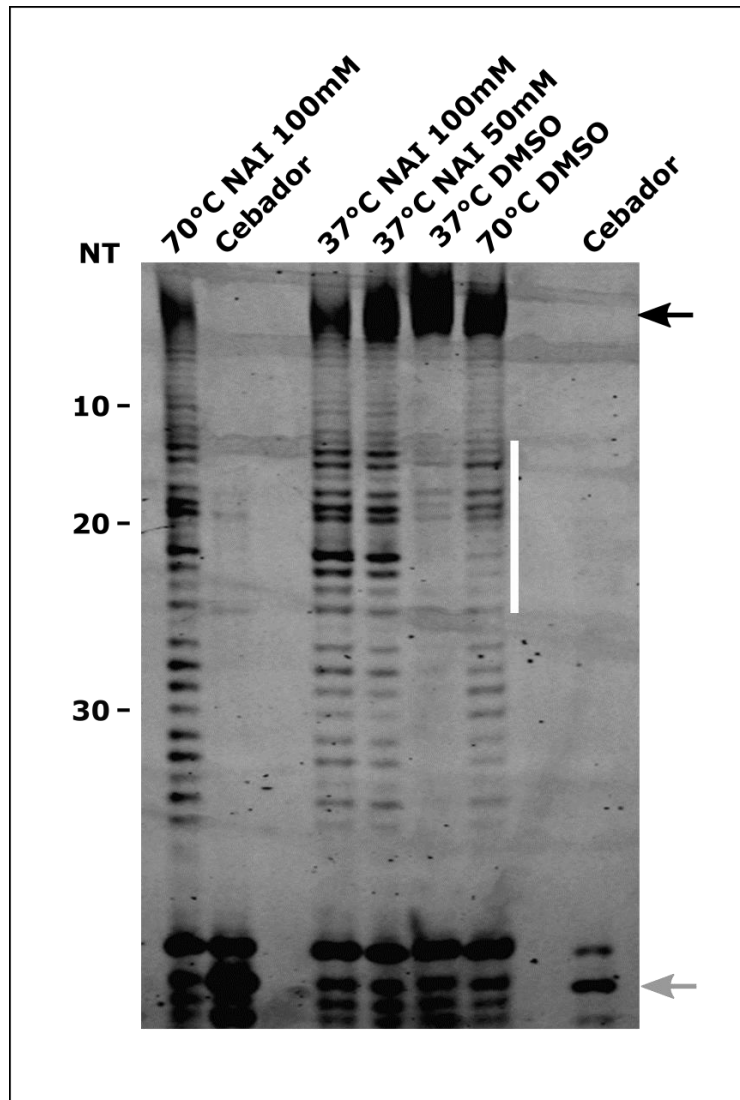


**Figura 16. RNA total modificado *in cellulo* con NAI.** RNA total obtenido de células HeLa tras un tratamiento con NAI 100 mM. El rRNA de las células tratadas con NAI (flechas blancas) presenta un ligero cambio en su migración electroforética en comparación con el rRNA control (DMSO, flechas negras)

## Validación de la metodología SHAPE

Tras corroborar la síntesis y propiedades del reactivo NAI. La estructura secundaria del aptámero Sc5-c3<sup>54</sup> previamente mapeado por ribonucleasas fue obtenida mediante SHAPE (Figura 17). Se observó que el aptámero tratado con DMSO a 37°C presentó una banda mayoritaria de alto peso molecular, correspondiente al transcrito completo (flecha negra). En comparación, el aptámero tratado con DMSO e incubado a 70°C también presentó una banda mayoritaria de alto peso molecular, así como bandas de menor peso; estas últimas, debidas a la degradación normal del RNA molde a altas temperaturas. El aptámero tratado con 50 o 100 mM de NAI presentó bandas intensas a un menor peso molecular, comparado con su contraparte con DMSO, especialmente entre los nucleótidos 13-28 (línea blanca), sugiriendo una zona de ssRNA. Finalmente, el aptámero tratado con 100 mM de NAI e incubado a

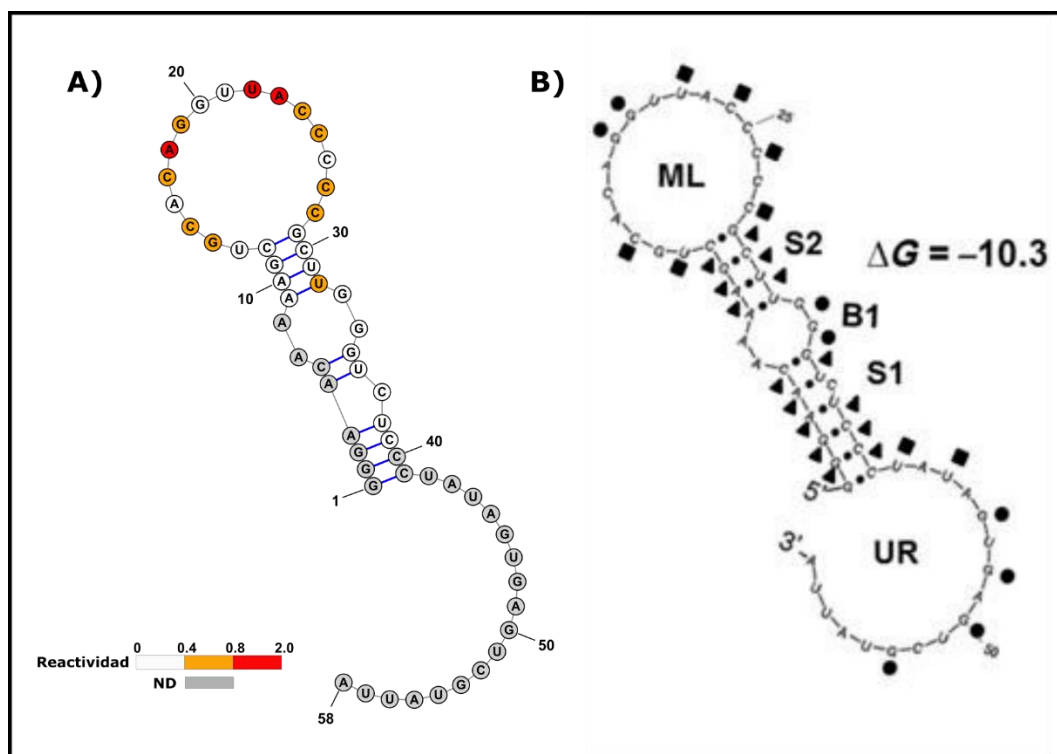
70°C presentó señales intensas en prácticamente cada nucleótido, lo cual es esperado debido a que a altas temperaturas la estructura secundaria del RNA se encontrará preferentemente como cadena sencilla. Dada esta característica, el RNA incubado a 70°C con 100 mM de NAI se usó como marcador de peso molecular.



**Figura 17. Patrón de reactividad del aptámero Sc5-c3 determinada por SHAPE.** El aptámero Sc5-c3 se trató con 50 mM y 100 mM de NAI a 37°C, obteniendo una zona de alta reactividad (línea blanca), correspondiente a regiones de ssRNA (nt 13-28). El aptámero tratado con DMSO (control negativo) se incubó a 37°C y 70°C. Como marcador de peso se incubó el aptámero con 100 mM de NAI a 70°C. Flecha gris: cebador fluorescente; flecha negra: producto completo de la reverso-transcripción.



Las bandas obtenidas se cuantificaron, normalizaron y se utilizaron como datos de pseudo-energía libre para generar un modelo estructural del aptámero Sc5-c3 (Ver Materiales y Métodos). Los nucleótidos más reactivos se muestran en colores rojo y amarillo en comparación con los nucleótidos en color blanco, que no presentaron reactividad significativa. Los nucleótidos de los cuales no se obtuvo información experimental fueron marcados en color gris (Figura 18). El modelo obtenido mediante SHAPE es congruente con los datos experimentales previamente publicados<sup>54</sup>; con lo cual se validó la técnica de SHAPE con el reactivo NAI previamente sintetizado *in-house*.



**Figura 18. Estructura Secundaria del aptámero SC5-c3.** A) Estructura secundaria del aptámero SC5-c3 obtenida mediante SHAPE. B) Estructura del aptámero SC5-c3 previamente caracterizada<sup>54</sup>. Escala de colores indican la reactividad de cada nucleótido. ND: nucleótidos sin dato experimental.

## Transfección de la construcción pCAP-let-7f-1 en células HEK-293TT

Para el uso de la metodología SHAPE propuesta, es necesario aumentar la abundancia relativa del RNA que se desee estudiar *in cellulo* con el fin de promover una mayor síntesis del transcrito, el cual servirá como molde para el paso de reverso-transcripción. En este paso crítico, el número de copias de RNA estará en correlación directa con la cantidad de cDNA generado y tendrá que ser suficiente para obtener una señal medible durante el paso de lectura basado en electroforesis capilar. Para ello, se utilizó la construcción pCAP-let-7f-1 previamente descrita (ver Materiales y Métodos) la cual contiene características importantes para favorecer la síntesis de RNA en células HEK-293TT.

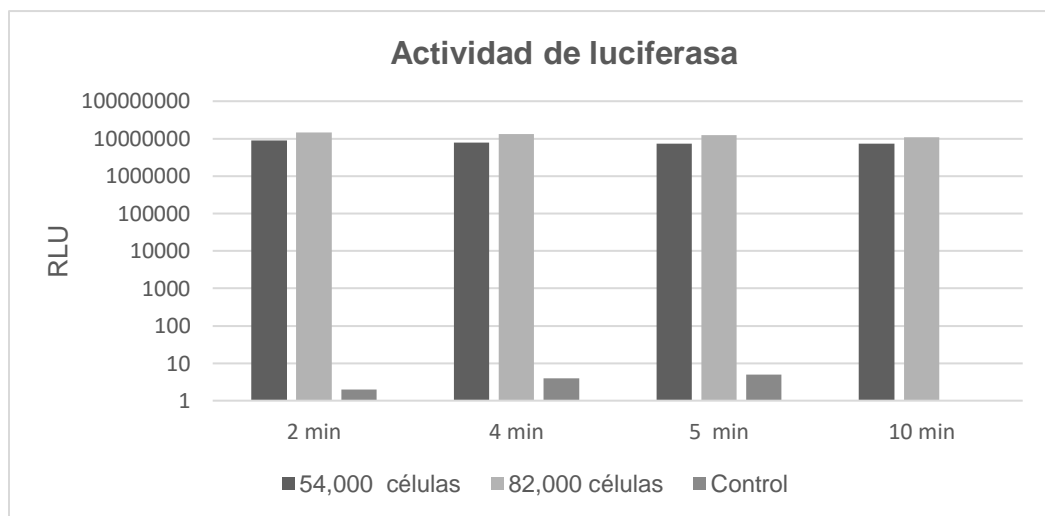
La construcción pCAP-let-7f-1 se transfectó en células HEK-293TT (ver Materiales y Métodos) y se evaluó la viabilidad celular obtenida con el fin de verificar que el proceso de transfección no fuera tóxico (Tabla 10). Encontrando que la viabilidad obtenida es comparable a transfectar construcciones no relacionadas (pSV1-YFP y pCR3.1-YFP).

Construcción	# de células	# células vivas	% de viabilidad	% promedio viabilidad
pSV1-YFP	3.21x10 <sup>6</sup>	3.07x10 <sup>6</sup>	94	95.5
pCR3.1-YFP	2.01x10 <sup>6</sup>	1.89x10 <sup>6</sup>	96	96.5
pCAP-let-7f-1	2.7x10 <sup>6</sup>	2.63x10 <sup>6</sup>	97	97

**Tabla 10. Viabilidad de celular evaluada con azul tripano.** Células HEK-293TT se transfectaron con la construcción pCAP-let-7f-1 y dos construcciones no relacionadas (pSV1-YFP y pCR3.1 YFP). Todas las transfecciones presentaron un porcentaje de viabilidad similar (N=2).

Establecido que la transfección de pCAP-let-7f-1 no resultaba en una disminución de la viabilidad celular, se prosiguió a determinar la actividad de luciferasa con el fin de verificar de manera indirecta la presencia del transcrito de fusión pre-let-7f-1-luciferasa (ver Materiales y Métodos). Las

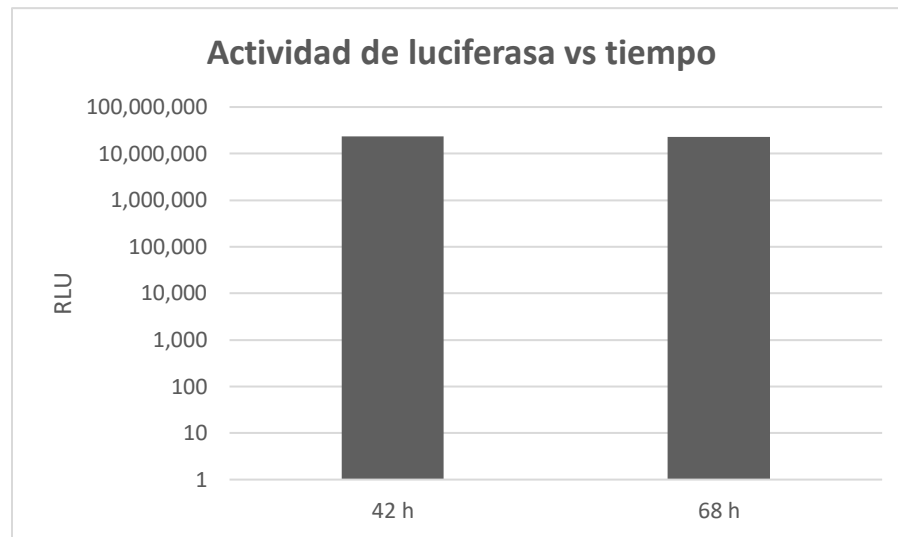
determinaciones se hicieron 66 horas post-transfección con  $5.4 \times 10^4$  y  $8.2 \times 10^4$  células. La actividad de luciferasa se midió a los 2, 4, 5 y 10 minutos post lisis con el fin de evaluar la estabilidad de la señal durante el tiempo. En paralelo se midió la actividad de luciferasa a partir de  $10^5$  células HEK-293TT transfectadas con la construcción pSV1-YFP (control negativo). Los resultados demostraron que la actividad de luciferasa fue estable durante los tiempos evaluados y cuantitativamente similar aun usando un distinto número de células, sumado a la nula actividad de luciferasa al transfectar la construcción pSV1-YFP (Gráfica 3). Dichos resultados permiten establecer que la medición de la actividad de luciferasa al transfectar el plásmido pCAP-let7f1 resulta una metodología fácil y robusta para corroborar una transfección eficiente en la línea celular HEK-293TT.



**Gráfica 2. Actividad de luciferasa en células HEK-293TT transfectadas.** Actividad de luciferasa en 54,000 u 82,000 células transfectadas con el plásmido pCAP-let-7f-1 medido a los 2, 4, 5 y 10 post-lisis. En paralelo se midieron 100,000 células transfectadas con pSV1-YFP (control negativo). RLU, unidades relativas de luz; min, minutos; N=1.

Posteriormente se evaluó la actividad de luciferasa a las 42 y 66 horas post-transfección, lisando todas las células contenidas en una caja p60 (aproximadamente entre  $2-3 \times 10^6$  de células por caja). Como se observa en la Gráfica 3 la actividad de luciferasa fue prácticamente idéntica entre las 42 y 68

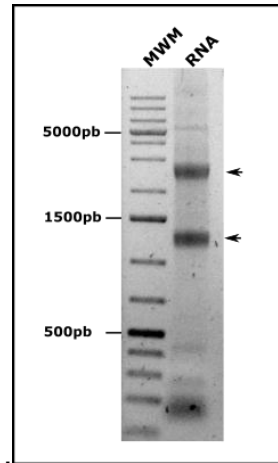
horas post transfección. Considerando estos resultados se decidió establecer 42 horas post-transfección como la condición final para cosechar las células.



**Gráfica 3. Actividad de luciferasa post-transfección.** Actividad de luciferasa en células HEK-293TT transfectadas con pCAP-let-7f-1 a las 42 horas y 68 horas post-transfección. RLU, unidades relativas de luz; h, horas; N=1.

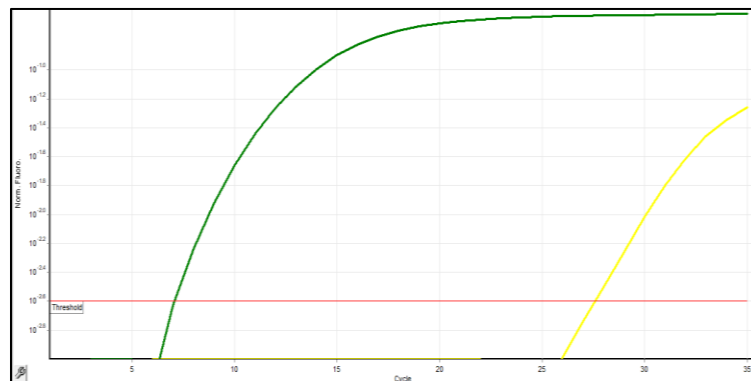
Como ultima determinación, se evaluó el transcrito de fusión pre-let-7f-1—luciferasa mediante RT-qPCR (ver Materiales y Métodos) con el fin de validar de manera directa la presencia del transcrito y normalizarlo en función del RNA constitutivo U6.

A partir de células HEK-293TT transfectadas con el plásmido pCAP-let-7f-1, se obtuvo el RNA total, se cuantificó y se evaluó su integridad mediante electroforesis en gel de agarosa (ver Materiales y Métodos). El RNA obtenido mostró el patrón característico esperado, observándose claramente los RNAs ribosomales 28S y 18S (Figura 19, flechas negras).



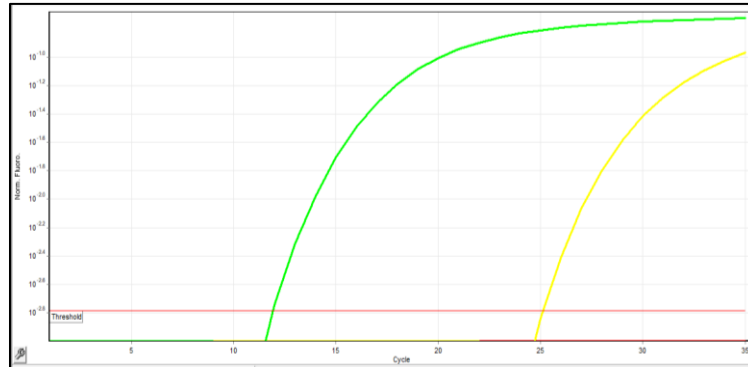
**Figura 19. RNA total de células HEK-293TT transfectadas.** RNA total de células 293TT transfectadas con la construcción pCAP-let-7f-1. Flechas negras señalan los RNAs ribosomales 28S y 18S. MWM, marcador de peso molecular.

Verificada la calidad del RNA purificado, se utilizó para realizar las reacciones de amplificación por RT-qPCR (ver Materiales y Métodos) para el gen U6 y el transcrito de fusión let-7f-1—luciferasa. La amplificación de U6 fue correcta en el RNA evaluado (Gráfica 5, línea verde) obteniendo un Ct de 7.07; además, se observó la amplificación de DNA genómico en el control RT menos (Gráfica 5, línea amarilla), aproximadamente  $10^6$  veces más diluido en comparación al RNA correspondiente a U6.



**Gráfica 4 Amplificación de U6 en células 293TT transfectadas.** El RNA U6 se amplificó mediante RT-qPCR (línea verde) a partir de células 293TT transfectadas con la construcción pCAP-let-7f-1. Línea amarilla, control RT-menos; línea negra, control negativo.

El fragmento correspondiente al transcrito let-7f-1—luciferasa también presentó una amplificación correcta (Gráfica 6, línea verde) con un valor de Ct de 12.32; además, se observó la amplificación de DNA correspondiente al vector de expresión en el control RT menos (Gráfica 6, línea amarilla), aproximadamente  $10^4$  veces más diluido en comparación al RNA correspondiente al transcrito de fusión.



**Gráfica 5. Amplificación del transcrito let-7f-1—luciferasa en células HEK-293T transfectadas.** El RNA let-7f-1—luciferasa se amplificó mediante RT-qPCR (línea verde) a partir de células HEK-293TT transfectadas con la construcción pCAP-let-7f-1. Línea amarilla, control RT-menos; línea negra, control negativo.

Dados los resultados anteriores, se concluye que se estandarizaron las condiciones para la transfección de precursores de *let-7*, lo cual permite generar una cantidad considerable del transcrito *in cellulo*. El aumento de la abundancia relativa permitirá la obtención de su estructura secundaria mediante la técnica SHAPE.

## Discusión

El estudio detallado de biomoléculas a un nivel estructural ha permitido comprender las características conformacionales que les permiten interactuar con otras moléculas orgánicas e inorgánicas y modelar estructuras tridimensionales que permiten predecir las condiciones para su interacción. De manera específica, este tipo de análisis ha sido esencial para entender las características que gobiernan la interacción de RNAs y sus proteínas de unión. Históricamente, las características y dominios estructurales que conforman a una proteína de unión a RNA han sido protagónicos para entender cómo se une a cierto RNA con alta especificidad y afinidad; describiéndose así los dominios *Cold Shock*, ZnF, RRM, KH, WD40 o los dominios DEAD-box. La información de estos dominios logró clasificar de manera sistemática la interacción de proteínas con RNAs; sin embargo, dejaba en un papel secundario a la estructura del RNA y eludía la creciente información de proteínas con regiones intrínsecamente desordenadas (IDRs) las cuales se unen a RNAs sin presentar algún dominio estructurado. Estas nuevas observaciones, han evidenciado que las características fisicoquímicas y estructurales del RNA tienen un papel fundamental para su unión con otras biomoléculas<sup>71,72</sup>.

En el caso de los *let-7*, esta familia se ha utilizado como modelo para revelar las características claves que debe tener un pri-miRNA para ser escindido por el Microprocesador<sup>73</sup>; así como las estructuras diferenciales que promueven el corte por la enzima Dicer<sup>31</sup> y aquellas que permiten una eficiente degradación por la exoribonucleasa DIS3L2<sup>74</sup>. Dichas características resultan decisivas para degradar al pri-/pre-miRNA para dar lugar a un miRNA maduro; inclusive han sido extendidas para conocer los procesos de biogénesis de otros miRNAs. Este trabajo se fundamentó en los precursores de *let-7* para elucidar el papel de su estructura secundaria en la interacción con la proteína reguladora LIN28B.

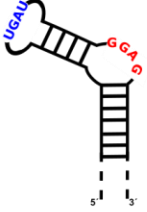
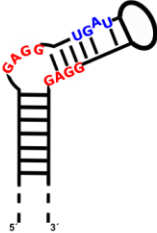
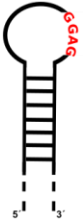
Desde una visión general, pre-let-7f-1 y pre-miR-98 son similares en la región pre-E: ambas moléculas coinciden en tener la secuencia consenso GNGAY/UGAU en el extremo 5´ y un sitio GGAG en el extremo 3´, ambos en una conformación de ssRNA. Además, dichos extremos se encuentran separados por un tallo (dsRNA) enriquecido en guaninas y citosinas, lo cual brinda alta estabilidad a este elemento estructural. En contraste pre-let-7g contiene dos sitios GGAG ubicados en los extremos 5´ y 3´, separados por una región tallo-burbuja, la cual contiene la secuencia GNGAY/UGAU en dsRNA. Finalmente, el pre-E de pre-let-7c contiene sólo una secuencia GGAG en ssRNA (Figura 10).

En el caso de la estructura de pre-let-7f-1, la región pre-E coincide con la estructura cristalográfica previamente establecida en conjunto con LIN28B<sup>46</sup>, incluidos los nucleótidos 29, 30, 50 y 51. Estos nucleótidos han sido controversiales ya que las estructuras *in silico* los modelaban en dsRNA, mientras que los resultados cristalográficos en presencia de LIN28B los mostraban en ssRNA. Dichas discrepancias se atribuían a que LIN28B propiciaba la separación de ambas cadenas tras su unión en el pre-E<sup>46</sup>; sin embargo, los resultados obtenidos en este trabajo muestran que estos nucleótidos se encuentran en ssRNA previo a la unión con LIN28B. Cabe mencionar que las estructuras obtenidas corresponden a precursores de let-7 de longitud completa, sin ninguna alteración en el pre-E y la determinación estructural se realizó en solución. En contraste, los ensayos cristalográficos, y otras determinaciones estructurales basadas en ribonucleasas, realizaban modificaciones o deleciones<sup>46,75</sup> a los precursores debido a la dificultad metodológica para realizar caracterizaciones experimentales con RNAs completos, lo cual puede explicar las diferencias obtenidas. Sumado a lo anterior, las secuencias consenso GNGAY/UGAU y GGAG mostraron una alta reactividad SHAPE, evidenciando que estos nucleótidos son altamente accesibles para una interacción estable con LIN28B. Estas características de reactividad también se encontraron en pre-miR-98 el segundo precursor más largo de la familia *let-7*.



Pre-let-7g presentó un patrón similar (burbuja-tallo-burbuja) en el pre-E; sin embargo, los nucleótidos 28-31 y 54-57, los cuales contienen la secuencia consenso GGAG, no presentaron un patrón de reactividad tan claro como pre-let-7f-1. Además, la secuencia consenso GNGAY/UGAU, la cual presenta alta reactividad SHAPE, se encuentra en una burbuja asimétrica dentro del tallo pre-E. Ambas características sugieren que la interacción con LIN28B puede verse comprometida, lo cual coincide con los resultados de afinidad entre LIN28B y pre-let-7g (Figura 12). Este patrón, en el cual pre-let-7f-1 tiene una afinidad mayor a LIN28B que pre-let-7g, coincide con resultados previamente publicados con LIN28A<sup>46</sup>

En el caso de pre-let-7c observamos que el pre-E no cumple con el mismo patrón estructural, presentando una burbuja apical de ~20 nucleótidos. Sin embargo, la única secuencia consenso GGAG en el pre-E presenta alta reactividad SHAPE, al igual que los nucleótidos aledaños, evidenciando una zona altamente accesible para los ZnF de LIN28B. Dadas estas características, se sugiere que LIN28B sólo interacciona con pre-let-7c mediante los ZnF para su regulación; esta hipótesis se fortalece con resultados recientes, los cuales evidenciaron que los ZnF son necesarios y suficientes para la regulación de pre-let-7. Este dominio de unión a RNA pueden unirse por sí solo a pre-let-7 de una manera altamente estable, reclutar a TUT/7 y promover la degradación del precursor por acción de DIS3L2<sup>76</sup>.

<b>Estructura</b>	<b>Precursor</b>	<b>Kd</b>
	<b>Pre-miR-98</b>	<b>9.24 nM</b>
	<b>Pre-let-7f-1</b>	<b>13.93 nM</b>
	<b>Pre-let-7g</b>	<b>55.80 nM</b>
	<b>Pre-let-7c</b>	<b>35.00 nM</b>

**Figura 20. Estructuras ideales y alternas para la interacción con LIN28B.** Los precursores con la mayor afinidad (pre-let-7f-1 y pre-miR-98) presentan estructuras “ideales” para la interacción con LIN28. Pre-let-7c y pre-let-7g presentan estructuras alternas, las cuales interactúan con menor afinidad con LIN28B.

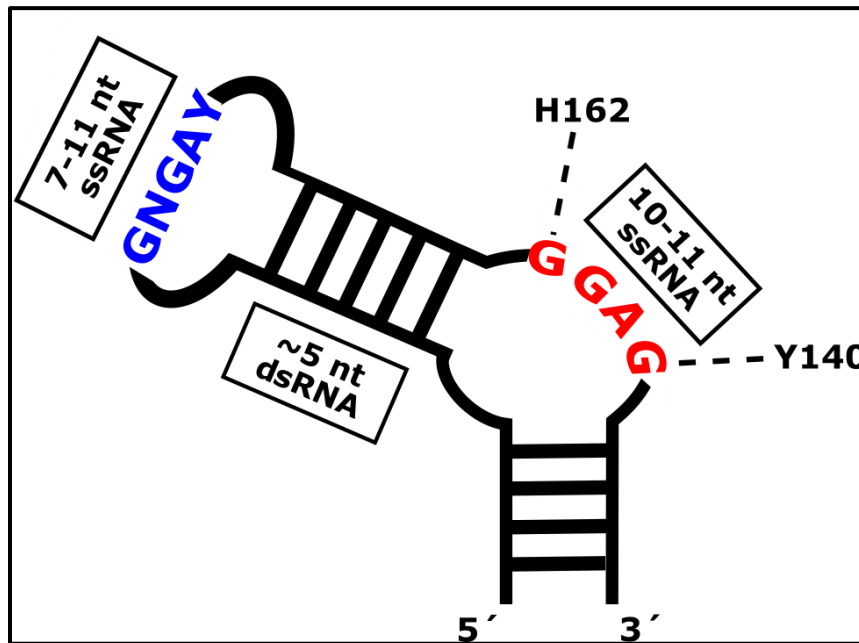
En la Figura 20 se resumen las estructuras y constantes de disociación obtenidas, evidenciando que los precursores de *let-7* evaluados experimentalmente mostraron estructuras secundarias diversas las cuales se relacionan con sus constantes de afinidad. Sin embargo, pre-let-7f-1 y pre-miR-98 presentaron características estructurales estadísticamente similares (Tabla 9). Estos resultados coinciden con los ensayos de afinidad LIN28B:pre-let-7 (Figura 12), observando que ambos precursores presentan la mejor afinidad con LIN28B. Debido a esta relación estructura-afinidad, proponemos a pre-let-7f-1 y pre-miR-98 como estructuras ideales para la interacción con LIN28B,

mientras que las estructuras presentes en los precursores de *let-7g*, *let-7c* y *let-7a-1* resultan adversas para esta interacción.

Dentro de los trabajos previos enfocados en los precursores *let-7*, estudios de dinámica molecular propusieron que la burbuja del extremo 5', la cual contiene la secuencia consenso UGAU/GNGAY, sería un factor altamente importante para una interacción estable entre LIN28A y pre-*let-7*<sup>50</sup>. En dicho trabajo, se compara la burbuja de pre-*let-7d* la cual contiene 7 nucleótidos en ssRNA, contra la estructura de pre-*let-7c-2* con sólo 3 nucleótidos en ssRNA; sugiriendo que las características de la burbuja de pre-*let-7d* son más favorables para la interacción con LIN28A. Sumado a esto, el análisis de la estructura cristalográfica permitió establecer las interacciones que guían la unión con los ZnF. En este caso, los residuos presentes en las posiciones H162 y Y140 generan un *pocket* hidrofóbico para la interacción estable con la primera y última guanina de la secuencia consenso GGAG; mientras que los residuos aledaños V170, K160 y R138 estabilizan la interacción formando puentes de hidrógeno con los átomos que conforman a las bases nitrogenadas<sup>46</sup>.

Además de estas características estructurales, los experimentos realizados con pre-*let-7f-1mut* confirman que la región de doble cadena puede estabilizar la interacción de manera importante (Figura 13). Esto concuerda con hipótesis previas, las cuales sugerían que la región de doble cadena funciona como un "nudo de corbata", el cual permite una conformación estructural más estable entre la secuencia consenso GAGAY/UGAU y el CSD<sup>46</sup>. Finalmente, el intercambio de posición de la secuencia GGAG del extremo 5' al extremo 3', demostró que dicho cambio compromete la interacción con LIN28B, en este caso, la unión específica se abatió a su totalidad. Dicho resultado podría explicar la conservación de los nucleótidos correspondientes a la secuencia consenso GGAG en el extremo 3' y la secuencia GNAGAY/UGAU en el extremo 5'. Esta conservación se ha reportado a través de distintos precursores de *let-7* en distintos organismos<sup>51,77,78</sup>. Considerando los antecedentes previamente mencionados, en conjunto con la evidencia experimental obtenida, se ha

propuesto una estructura mínima ideal para la unión eficiente de un RNA con la proteína LIN28 (Figura 21).



**Figura 21. Estructura mínima de unión con LIN28B.** Para la correcta unión entre un RNA y LIN28B se propone una estructura mínima con las secuencias consenso GNGAY/UGAU y GGAG en conformación de ssRNA, separadas por un tallo altamente estable el cual estabiliza la interacción<sup>46</sup>. La primera y última G del GGAG interaccionan fuertemente con los residuos H162 y Y140 de LIN28<sup>46</sup>, respectivamente. Se propone que la burbuja con el GNGAY/UGAU tenga al menos 7 nucleótidos para la correcta unión del CSD<sup>50</sup>

Este modelo experimental resulta relevante ya que un oligonucleótido pequeño con dichas características estructurales podría competir eficientemente con los precursores endógenos de pre-let-7; evitando así su degradación por LIN28A/B. Dicha molécula inhibitoria se sumaría a las propuestas de pequeños inhibidores que se han probado experimentalmente como estrategias terapéuticas para reprimir los efectos de alta proliferación y des-diferenciación provocados por LIN28A/B en distintos tipos de cáncer<sup>79</sup>.

Finalmente, la estructura secundaria de RNAs puede variar en distintos contextos, en los cuales existen condiciones variables en la presencia de iones,

proteínas, temperatura, moléculas orgánicas, etc<sup>17</sup>. Por estos motivos, los estudios más relevantes sobre RNA se han llevado *in vivo* o *in cellulo*<sup>25,27,53</sup>. Debido a las características intrínsecas presentes en cada célula se pueden obtener estructuras que nos permitan modelar con mayor exactitud la interacción entre RNAs y proteínas; permitiendo correlacionar estas estructuras con su función biológica. Por ese motivo fue que en el presente trabajo se establecieron las condiciones metodológicas para mapear estructuralmente a los precursores de *let-7 in cellulo*. Los resultados obtenidos nos permiten concluir que la determinación estructural de estos precursores es viable, ya que pre-let-7f-1 se puede sobre-expresar de manera eficiente en la línea celular HEK-293TT, sin presentar un efecto tóxico para las células (Tabla 10). Además, se logró la síntesis correcta del reactivo NAI con el cual se obtuvo la caracterización estructural del aptámero Sc5-c3<sup>54</sup> (Ver Materiales y Métodos así como Resultados). Dado los resultados obtenidos *in vitro* y el modelo mínimo de unión, resulta especialmente relevante comparar las observaciones en este contexto contra las estructuras *in cellulo*. Dicha comparación permitirá revelar si la estructura mínima, establecida con los resultados *in vitro* se conserva en pre-let-7f-1 o si existe otra conformación estructural más estable, guiada por las condiciones presentes en el contexto celular. Los resultados serían un elemento clave para entender las estructuras posibles de precursores de miRNAs *in cellulo* y cómo estos patrones estructurales pueden determinar la regulación post-transcripcional de un RNA.

## Conclusiones

- Los precursores del *let-7* presentan estructuras secundarias estadísticamente distintas.
- La estructura secundaria de cada precursor tiene relación con la afinidad que presentan hacia LIN28B. Los cambios en afinidad pueden explicar la regulación diferencial ejercida por LIN28B a cada pre-*let-7*.
- El cambio de las características estructurales compromete la interacción con LIN28B de manera distinta.
- Se establecieron las condiciones para obtener la secundaria de pre-*let-7* *in cellulo*.

## Perspectivas

- Determinar la afinidad *in vitro* del RNA mínimo con la proteína LIN28B.
- Determinar la estructura *in cellulo* de los precursores de *let-7*.
- Comparar las estructuras *in vitro* con las estructuras *in cellulo*.
- Transfectar el RNA mínimo, así como variantes, en un modelo celular para evaluar un posible efecto inhibitorio contra LIN28B.

## Referencias

1. Brenner, S., Jacob, F. & Meselson, M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* **190**, 576–581 (1961).
2. Scherrer, K., H, L. & Darnell, J. E. Demonstration of an unstable RNA and of a precursor to ribosomal RNA in HeLa cells. *Proc. Natl. Acad. Sci. U. S. A.* **49**, 240–8 (1963).
3. Lemay, J.-F. & Lafontaine, D. A. Core requirements of the adenine riboswitch aptamer for ligand binding. *RNA* **13**, 339–50 (2007).
4. Cech, T. R. & Steitz, J. A. The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. *Cell* **157**, 77–94 (2014).
5. McKeague, M., Wong, R. S. & Smolke, C. D. Opportunities in the design and application of RNA for gene expression control. *Nucleic Acids Res.* **44**, 2987–99 (2016).
6. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
7. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–10 (1990).
8. Leija-Montoya, A. G., Benítez-Hess, M. L. & Alvarez-Salas, L. M. Application of Nucleic Acid Aptamers to Viral Detection and Inhibition. in *Nucleic Acids - From Basic Aspects to Laboratory Tools* (InTech, 2016). doi:10.5772/61773.
9. Ashrafuzzaman, M., Tseng, C.-Y., Kaptj, J., Mercer, J. R. & Tuszynski, J. A. A computationally designed DNA aptamer template with specific binding to phosphatidylserine. *Nucleic Acid Ther.* **23**, 418–26 (2013).

10. Fallmann, J. *et al.* Recent advances in RNA folding. *J. Biotechnol.* (2017) doi:10.1016/j.jbiotec.2017.07.007.
11. Rouskin, S., Zubradt, M., Washietl, S., Kellis, M. & Weissman, J. S. Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* **505**, 701–705 (2013).
12. Devi, G., Zhou, Y., Zhong, Z., Toh, D.-F. K. & Chen, G. RNA triplexes: from structural principles to biological and biotech applications. *Wiley Interdiscip. Rev. RNA* **6**, 111–128 (2015).
13. Xin, Y., Laing, C., Leontis, N. B. & Schlick, T. Annotation of tertiary interactions in RNA structures reveals variations and correlations. *RNA* **14**, 2465–77 (2008).
14. Soemedi, R. *et al.* The effects of structure on pre-mRNA processing and stability. *Methods* (2017) doi:10.1016/j.ymeth.2017.06.001.
15. Foley, S. W., Kramer, M. C. & Gregory, B. D. RNA structure, binding, and coordination in *Arabidopsis*. *Wiley Interdiscip. Rev. RNA* e1426 (2017) doi:10.1002/wrna.1426.
16. Spitale, R. C. & Wedekind, J. E. Exploring ribozyme conformational changes with X-ray crystallography. *Methods* **49**, 87–100 (2009).
17. Ganser, L. R., Kelly, M. L., Herschlag, D. & Al-Hashimi, H. M. The roles of structural dynamics in the cellular functions of RNAs. *Nature Reviews Molecular Cell Biology* vol. 20 474–489 (2019).
18. Ke, A. Crystallization of RNA and RNA-protein complexes. *Methods* **34**, 408–414 (2004).
19. Gerashchenko, M. V. & Gladyshev, V. N. Ribonuclease selection for ribosome profiling. *Nucleic Acids Res.* **45**, e6–e6 (2017).
20. Bevilacqua, P. C., Ritchey, L. E., Su, Z. & Assmann, S. M. Genome-



- Wide Analysis of RNA Secondary Structure. *Annu. Rev. Genet.* **50**, 235–266 (2016).
21. Cruz, J. A. *et al.* The dynamic landscapes of RNA architecture. *Cell* **136**, 604–9 (2009).
  22. Cordero, P., Kladwang, W., VanLang, C. C. & Das, R. Quantitative Dimethyl Sulfate Mapping for Automated RNA Secondary Structure Inference. *Biochemistry* **51**, 7037–7039 (2012).
  23. Lee, B. *et al.* Comparison of SHAPE reagents for mapping RNA structures inside living cells. *RNA* **23**, 169–174 (2017).
  24. Smola, M. J. *et al.* SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the *Xist* lncRNA in living cells. *Proc. Natl. Acad. Sci.* **113**, 10322–10327 (2016).
  25. Diaz-Toledano, R., Lozano, G. & Martinez-Salas, E. In-cell SHAPE uncovers dynamic interactions between the untranslated regions of the foot-and-mouth disease virus RNA. *Nucleic Acids Res.* **45**, 1416–1432 (2017).
  26. Ding, Y. *et al.* In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **505**, 696–700.
  27. Guo, J. U. & Bartel, D. P. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* (80-. ). **353**, (2016).
  28. Strobel, E. J., Yu, A. M. & Lucks, J. B. High-throughput determination of RNA structures. *Nat. Rev. Genet.* **19**, 615–634 (2018).
  29. Li, L. C. *et al.* Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 17337–17342 (2006).
  30. Paul, P. *et al.* Interplay between miRNAs and human diseases. *J. Cell. Physiol.* (2017) doi:10.1002/jcp.25854.

31. Heo, I. *et al.* Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. *Cell* **151**, 521–532 (2012).
32. Yi, R., Qin, Y., Macara, I. G. & Cullen, B. R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **17**, 3011–6 (2003).
33. Winter, J., Jung, S., Keller, S., Gregory, R. I. & Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* **11**, 228–234 (2009).
34. Ha, M. & Kim, V. N. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **15**, 509–524 (2014).
35. Low, J. T. *et al.* SHAPE-directed Discovery of Potent shRNA Inhibitors of HIV-1. *Mol. Ther.* **20**, 820–828 (2012).
36. Wang, X. *et al.* Regulation of let-7 and its target oncogenes (Review). *Oncol. Lett.* **3**, 955–960 (2012).
37. Roush, S. & Slack, F. J. The let-7 family of microRNAs. *Trends Cell Biol.* **18**, 505–516 (2008).
38. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. *Nature* **426**, 465–469 (2003).
39. MacRae, I. J. *et al.* Structural Basis for Double-Stranded RNA Processing by Dicer. *Science (80-. ).* **311**, (2006).
40. Ma, J.-B., Ye, K. & Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318–322 (2004).
41. Trabucchi, M. *et al.* The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* **459**, 1010–1014 (2009).

42. Faehnle, C. R., Walleshauser, J. & Joshua-Tor, L. Mechanism of Dis3L2 substrate recognition in the Lin28-let-7 pathway. *Nature* **514**, 252–6 (2014).
43. Ustianenko, D. *et al.* Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* **19**, 1632–8 (2013).
44. Tzialikas, J. & Romer-Seibert, J. LIN28: roles and regulation in development and beyond. *Development* **142**, 2397–2404 (2015).
45. Mayr, F. & Heinemann, U. Mechanisms of Lin28-Mediated miRNA and mRNA Regulation—A Structural and Functional Perspective. *Int. J. Mol. Sci.* **14**, 16532–16553 (2013).
46. Nam, Y., Chen, C., Gregory, R., Chou, J. & Sliz, P. Molecular Basis for Interaction of let-7 MicroRNAs with Lin28. *Cell* **147**, 1080–1091 (2011).
47. Piskounova, E. *et al.* Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. *Cell* **147**, 1066–1079 (2011).
48. Triboulet, R., Pirouz, M. & Gregory, R. I. A Single Let-7 MicroRNA Bypasses LIN28-Mediated Repression. *Cell Rep.* **13**, 260–6 (2015).
49. Ustianenko, D. *et al.* LIN28 Selectively Modulates a Subclass of Let-7 MicroRNAs. *Mol. Cell* **71**, 271-283.e5 (2018).
50. Sharma, C. & Mohanty, D. Molecular Dynamics Simulations for Deciphering the Structural Basis of Recognition of Pre-let-7 miRNAs by LIN28. *Biochemistry* **56**, 723–735 (2017).
51. Zamora-Contreras, A. & Alvarez-Salas, L. M. *Differential let-7 family expression does not correlate with LIN28B expression in cervical cells.* (2017).
52. Borrego-Diaz, E. *et al.* A potential regulatory loop between Lin28B:miR-212 in androgen-independent prostate cancer. *Int. J. Oncol.* **45**, 2421–

2429 (2014).

53. Spitale, R. C. *et al.* RNA SHAPE analysis in living cells. *Nat. Chem. Biol.* **9**, 18–20 (2013).
54. Leija-Montoya, A. G., Benítez-Hess, M. L., Toscano-Garibay, J. D. & Alvarez-Salas, L. M. Characterization of an RNA aptamer against HPV-16 L1 virus-like particles. *Nucleic Acid Ther.* **24**, 344–355 (2014).
55. Duncan, C. D. S. & Weeks, K. M. SHAPE analysis of long-range interactions reveals extensive and thermodynamically preferred misfolding in a fragile group I intron RNA. *Biochemistry* **47**, 8504–13 (2008).
56. Zarringhalam, K., Meyer, M. M., Dotu, I., Chuang, J. H. & Clote, P. Integrating Chemical Footprinting Data into RNA Secondary Structure Prediction. *PLoS One* **7**, e45160 (2012).
57. Francisco-Velilla, R., Fernandez-Chamorro, J., Dotu, I. & Martinez-Salas, E. The landscape of the non-canonical RNA-binding site of Gemin5 unveils a feedback loop counteracting the negative effect on translation. *Nucleic Acids Res.* **46**, 7339–7353 (2018).
58. Lozano, G., Francisco-Velilla, R. & Martinez-Salas, E. Ribosome-dependent conformational flexibility changes and RNA dynamics of IRES domains revealed by differential SHAPE. *Sci. Rep.* **8**, 5545 (2018).
59. Karabiber, F., McGinnis, J. L., Favorov, O. V. & Weeks, K. M. QuShape: Rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA* **19**, 63–73 (2013).
60. Reuter, J. S. & Mathews, D. H. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **11**, 129 (2010).

61. Purzycka, K. J. *et al.* Automated 3D RNA Structure Prediction Using the RNAComposer Method for Riboswitches1. *Methods Enzymol.* **553**, 3–34 (2015).
62. Pettersen, E. F. *et al.* UCSF Chimera?A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
63. Will, S., Joshi, T., Hofacker, I. L., Stadler, P. F. & Backofen, R. LocARNA-P: Accurate boundary prediction and improved detection of structural RNAs. *RNA* **18**, 900–914 (2012).
64. Mattei, E., Pietrosanto, M., Ferrè, F. & Helmer-Citterich, M. Web-Beagle: a web server for the alignment of RNA secondary structures. *Nucleic Acids Res.* **43**, 493–497 (2015).
65. Franco-Urquijo, P. A. Evaluación de la interacción de la proteína LIN28B con los miRNAs de la familia de let-7. (Cinvestav, 2018).
66. Laboratory of Cellular Oncology. 293TT cells. <https://ccrod.cancer.gov/confluence/display/LCOTF/293TT> (2018).
67. Rasmussen, R. Quantification on the LightCycler. in *Rapid Cycle Real-Time PCR* 21–34 (Springer Berlin Heidelberg, 2001). doi:10.1007/978-3-642-59524-0\_3.
68. González-Torres, A. *et al.* MYPT1 is targeted by miR-145 inhibiting viability, migration and invasion in 2D and 3D HeLa cultures. *Biochem. Biophys. Res. Commun.* **507**, 348–354 (2018).
69. Liu, Z. *et al.* Cryo-EM Structure of Human Dicer and Its Complexes with a Pre-miRNA Substrate. *Cell* **173**, 1191-1203.e12 (2018).
70. Motulsky, H. J. & Neubig, R. R. Analyzing Binding Data. *Curr. Protoc. Neurosci.* **52**, 7.5.1-7.5.65 (2010).

71. Järvelin, A. I., Noerenberg, M., Davis, I. & Castello, A. The new (dis)order in RNA regulation. *Cell Communication and Signaling* vol. 14 1–22 (2016).
72. Zagrovic, B., Bartonek, L. & Polyansky, A. A. RNA-protein interactions in an unstructured context. *FEBS Letters* vol. 592 2901–2916 (2018).
73. Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231–235 (2004).
74. Faehnle, C. R., Walleshauser, J. & Joshua-Tor, L. Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. *Nature* **514**, 252–256 (2014).
75. Lightfoot, H. L. *et al.* A LIN28-dependent structural change in pre-let-7g directly inhibits dicer processing. *Biochemistry* **50**, 7514–7521 (2011).
76. Wang, L. *et al.* LIN28 Zinc Knuckle Domain Is Required and Sufficient to Induce let-7 Oligouridylation. *Cell Rep.* **18**, 2664–2675 (2017).
77. Liang, T., Yang, C., Li, P., Liu, C. & Guo, L. Genetic Analysis of Loop Sequences in the Let-7 Gene Family Reveal a Relationship between Loop Evolution and Multiple Isoforms. *PLoS One* **9**, e113042 (2014).
78. Newman, M. A., Thomson, J. M. & Hammond, S. M. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* **14**, 1539–1549 (2008).
79. Zhou, J., Ng, S. B. & Chng, W. J. LIN28/LIN28B: An emerging oncogenic driver in cancer stem cells. *Int. J. Biochem. Cell Biol.* **45**, 973–978 (2013).
80. Martinez-Salas, E., Embarc-Buh, A. & Francisco-Velilla, R. Emerging Roles of Gemin5: From snRNPs Assembly to Translation Control. *Int. J. Mol. Sci.* **21**, 3868 (2020).

81. Fernandez-Chamorro, J. *et al.* Identification of novel non-canonical RNA-binding sites in Gemin5 involved in internal initiation of translation. *Nucleic Acids Res.* **42**, 5742–5754 (2014).
82. Francisco-Velilla, R. *et al.* RNA-protein coevolution study of Gemin5 uncovers the role of the PXSS motif of RBS1 domain for RNA binding. *RNA Biol.* 1–11 (2020) doi:10.1080/15476286.2020.1762054.

## Anexo

### Datos reactividad SHAPE:

pre-let-7f-1			pre-let-7g		
seqNum	seqRNA	normDiff	seqNum	seqRNA	normDiff
1	U	0	1	A	1.41
2	C	0	2	G	1.39
3	A	0.41	3	G	1.04
4	G	1.19	4	C	1.22
5	A	0.7	5	U	2.38
6	G	1.29	6	G	1.67
7	U	0.39	7	A	0.57
8	G	0.66	8	G	0.26
9	A	0.25	9	G	0.16
10	G	0	10	U	0.51
11	G	0	11	A	0.47
12	U	0	12	G	0.07
13	A	0.13	13	U	0.17
14	G	0	14	A	0.35
15	U	0	15	G	0.23
16	A	0	16	U	0
17	G	0	17	U	0.1
18	A	0	18	U	0.18
19	U	0	19	G	0.4
20	U	0.06	20	U	0.33
21	G	0.02	21	A	0.46
22	U	0	22	C	0.26
23	A	0	23	A	0.21
24	U	0	24	G	0.09
25	A	0.01	25	U	0.21
26	G	0.01	26	U	0
27	U	0.15	27	U	0
28	U	0.37	28	G	0
29	G	0.55	29	A	0.76
30	U	0.96	30	G	1.88
31	G	0.2	31	G	0
32	G	0.02	32	G	1.39
33	G	0.03	33	U	1.29
34	G	0	34	C	0.69
35	U	0	35	U	0.19



36	A	0.27	36	A	0.07
37	G	2.92	37	U	0.23
38	U	2.04	38	G	0.68
39	G	5.09	39	A	0.69
40	A	2.77	40	U	1.05
41	U	2.29	41	A	1.38
42	U	1.2	42	C	0.37
43	U	0.75	43	C	0.27
44	U	0.05	44	A	0.26
45	A	0.01	45	C	0.36
46	C	0	46	C	0.51
47	C	0	47	C	0.49
48	C	0	48	G	0.21
49	U	0.05	49	G	0.12
50	G	1.37	50	U	0.17
51	U	0.79	51	A	0.14
52	U	1.02	52	C	0.24
53	C	0.36	53	A	0.33
54	A	1.12	54	G	0.36
55	G	1.85	55	G	0.12
56	G	4.26	56	A	0.4
57	A	3.75	57	G	0.49
58	G	3.34	58	A	0.93
59	A	2.69	59	U	0.77
60	U	3.01	60	A	0.46
61	A	0.12	61	A	0.13
62	A	0	62	C	0.25
63	C	0.01	63	U	0
64	U	0	64	G	0.01
65	A	0	65	U	0.04
66	U	0.09	66	A	0.13
67	A	0.28	67	C	0.15
68	C	0.18	68	A	0.13
69	A	0.09	69	G	0.05
70	A	0.05	70	G	0.04
71	U	0.25	71	C	0.1
72	C	0.28	72	C	0.07
73	U	0.02	73	A	0.04
74	A	0.01	74	C	0
75	U	0	75	U	0.37
76	U	0.01	76	G	0
77	G	0.07	77	C	0.41
78	C	0	78	C	0

79	C	0.02	79	U	0
80	U	0.19	80	U	0.22
81	U	0.3	81	G	0
82	C	0.45	82	C	0.23
83	C	0.34	83	C	0.42
84	C	0.31	84	A	0.46
85	U	0.11			
86	G	0.07			
87	A	0.6			

pre-let-7c			miR-98		
seqNum	seqRNA	reactivity	seqNum	seqRNA	reactivity
1	G	0.34	13	A	0.43
2	C	0.2	14	U	0.59
3	A	0.14	15	G	0.1
4	U	0.16	16	C	0.56
5	C	0.4	17	C	0.47
6	C	0.42	18	A	0.23
7	G	0.32	19	G	0
8	G	0.07	20	G	0.42
9	G	0	21	G	0.6
10	U	0.04	22	U	1.13
11	U	1.85	23	G	0.31
12	G	0	24	A	1.24
13	A	0	25	G	1.08
14	G	0	26	G	0.37
15	G	0.04	27	U	0
16	U	0.07	28	A	0.4
17	A	0.01	29	G	0
18	G	0	30	U	0
19	U	0.01	31	A	0
20	A	0.02	32	A	0.83
21	G	0.01	33	G	0.42
22	G	0	34	U	0
23	U	0	35	U	0.1
24	U	0.02	36	G	0.11
25	G	0.12	37	U	0.18
26	U	0.18	38	A	0.02
27	A	0.03	39	U	0
28	U	0.06	40	U	0.01
29	G	0.05	41	G	0
30	G	0.17	42	U	0.42

31	U	0.78	43	U	0.69
32	U	0.97	44	G	0.6
33	U	1.09	45	U	0.61
34	A	1.18	46	G	0
35	G	0.68	47	G	0.02
36	A	0.69	48	G	0.17
37	G	1.19	49	G	0
38	U	0.7	50	U	0.12
39	U	0.17	51	A	0.25
40	A	0.47	52	G	3.5
41	C	0.13	53	G	1.83
42	A	0	54	G	5.78
43	C	0.07	55	A	3.03
44	C	0.26	56	U	1.99
45	C	0.75	57	A	0.75
46	U	0.96	58	U	0.51
47	G	0.97	59	U	0.18
48	G	0.98	60	A	0.34
49	G	0	61	G	0.62
50	A	0.94	62	G	0.29
51	G	0.83	63	C	0.06
52	U	0.89	64	C	0.16
53	U	0.94	65	C	1.5
54	A	0.18	66	C	0.81
55	A	0.11	67	A	1.07
56	C	0.03	68	A	0.36
57	U	0.03	69	U	1.87
58	G	0	70	U	2.76
59	U	0.07	71	A	5.38
60	A	0.06	72	G	4.88
61	C	0.01	73	A	4.97
62	A	0.04	74	A	4.17
63	A	0.07	75	G	3.42
64	C	0	76	A	0.06
65	C	0.04	77	U	0
66	U	0.56	78	A	0
67	U	0.12	79	A	0.02
68	C	0.11	80	C	0.17
69	U	0.07	81	U	0.59
70	A	0.17	82	A	0.4
71	G	0.17	83	U	0.04
72	C	1.03	84	A	0
73	U	0	85	C	0.3

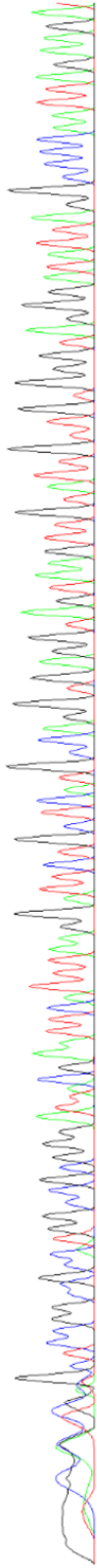
74	U	0.04	86	A	0.73
75	U	0.21	87	A	0.13
76	C	0.27	88	C	0.15
77	C	0.07	89	U	0
78	U	0	90	U	0.02
79	U	0.03	91	A	0
80	G	0.18	92	C	0
81	G	0.7	93	U	0.36
82	A	0.09	94	A	0.24
83	G	0.06	95	C	0.33
84	C	0.04	96	U	0.75
			97	U	0.75
			98	U	0.45
			99	C	0.01
			100	C	0.57
			101	C	1.27
			102	U	0.52
			103	G	0.59
			104	G	0.72
			105	U	0.59
			106	G	0.79
			107	U	0.19
			108	G	0.32
			109	U	0.13
			110	G	0.03
			111	G	0.09
			112	C	0.08
			113	A	0.18
			114	U	0.05
			115	A	0.22
			116	U	0.24
			117	U	0
			118	C	0
			119	A	0

## Secuenciaciones construcciones pCAP-let7

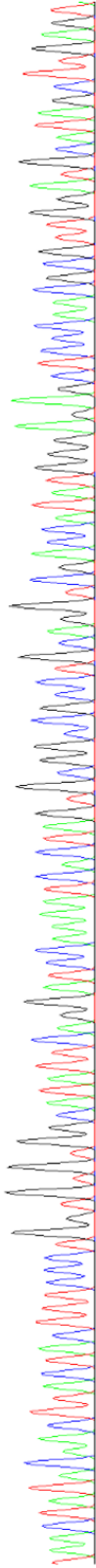
Las secuencias obtenidas se compararon con las secuencias de cada precursor presentes en la base de datos miRBase (<http://www.mirbase.org/>)

## Pre-miR-98

T CCGTCCCGGC TCCGAGCTCCGGGCGGTGGCCCGGGG AAT ACG AATTCAATTAAGG AATCTCGTCAATG CCAAGGTG AGGTAGTAAAGTTGTATTGTGTGGGTAGGGATATTAGGCCCAATTAGAAAG A 130

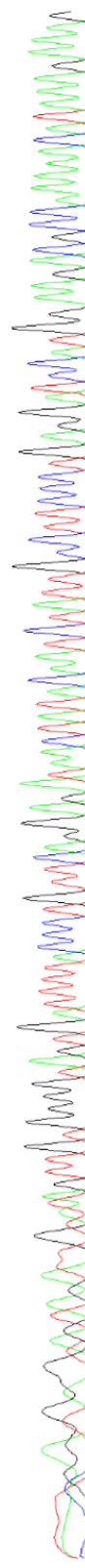


TAACTATACAACTTACTACTTCCCTGGTGTGGCATATTTCAGGATCCAAAAATCACTAGTCCGGCCGCCCTGCAGGTCGACCAATATGGGAGAGCTCCCAACGCCGTTGGATGCAATAGCITTGAGT 250



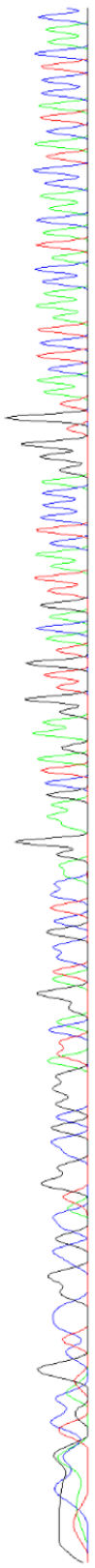
## Pre-let-7f-1

CATTT CAGGCGTG AGGTAGTAGATTGTATGTTTGGGGTATGTAATTTTACCCTGTTCCAGGAGATAACTATACAAATCTATTGCCCTCCCTGAGGATCCATGG AAGACGCCAAAAACATAAAAGAAAAG 130

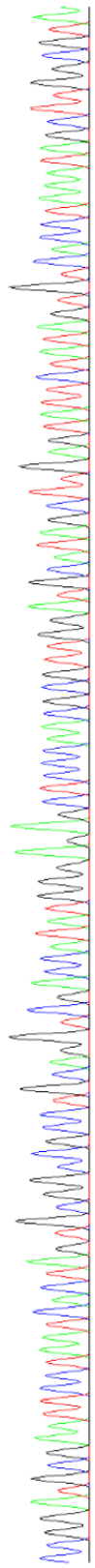


### Pre-let-7c

10 G A T A C A A T C G C T G C T C C G C C C C C G G G T C A G G A T C C C G C T C C A G G A A G C T A G A A G G T T G T A C A G T T A A C T C C C A G G T G T A A C T C T A A A C C A T A C A A C C T A C T A C C T C A A C  
 20  
 30  
 40  
 50  
 60  
 70  
 80  
 90  
 100  
 110  
 120

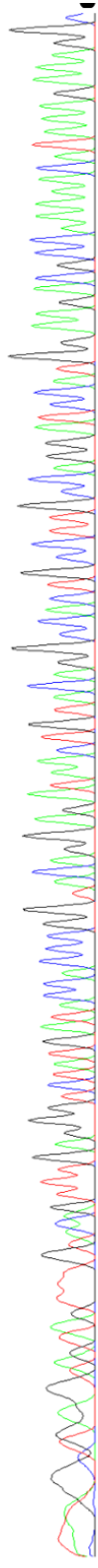


130 C C G G A T G C G A A T T C C C T A A T C A C T A G T G C G G C C C C C T G C A G G T C G A C C A T A T G G A G A G C T C C C A A C G C C G T T G G A T G C A T A G C T T G A G T A F T T C T A T A G T G T C A C C T A A A T A G C T T G G C G T A  
 140  
 150  
 160  
 170  
 180  
 190  
 200  
 210  
 220  
 230  
 240

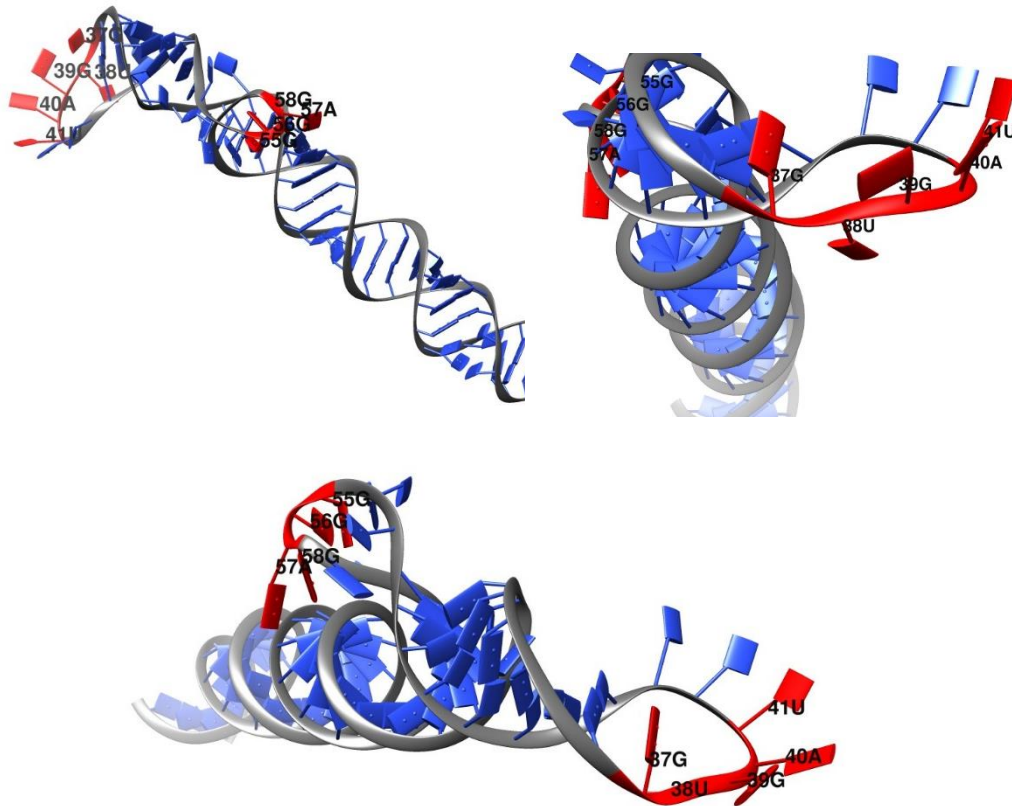


### Pre-let-7g

10 C T A A A G C C T G C G G T A G T A G T T T G T A G T A G T T T G T A G G T T G A G G G T C T A T G A T A C C A C C C C G G T A C A G G A G A T A A C T G T A C A G C C C A C T G C C T T G C C A G G A T C C A T G G A A G A C C C A A A A A C A T A A A G A A A G G C  
 20  
 30  
 40  
 50  
 60  
 70  
 80  
 90  
 100  
 110  
 120



## Modelo tridimensional pre-let-7f-1

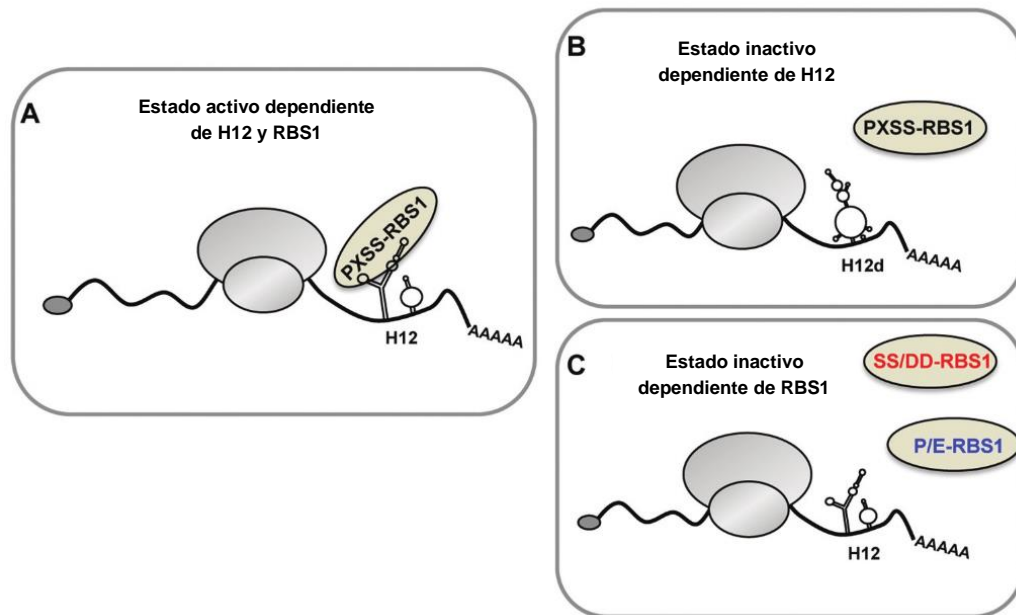


**Figura 22. Modelo estructural 3D pre-let-7f-1.** Se muestran diferentes ángulos del modelo 3D diseñado con RNAComposer de pre-let-7f-1 (ver Materiales y Métodos). En rojo y letras negras se resaltan las secuencias consenso GGAG y GNGAY.

## GEMIN5, dominio RBS1 y el RNA H12

Un ejemplo interesante (y fundamental para el desarrollo técnico de este proyecto) es la proteína GEMIN5 y su interacción con el fragmento de mRNA denominado H12. La proteína GEMIN5 es una proteína multifuncional perteneciente al complejo SMN; su principal función como complejo es de chaperona molecular en la biogénesis y función de ribonucleoproteínas nucleares pequeñas (snRNPs). GEMIN5 contiene múltiples dominios, siendo los más relevantes el dominio WD40 y los dominios no canónicos de unión a RNA, RBS1 y RBS2<sup>80</sup>. En particular, el estudio detallado de GEMIN5 evidenció

que el dominio RBS1 se une al propio mRNA de GEMIN5 en una porción nombrada como H12<sup>57</sup>; además, diversos análisis demostraron que la región directamente involucrada en la interacción con H12 es una IDR<sup>81</sup>. En particular, la participación en este proyecto, encabezado por la Dra. Encarnación Martínez-Salas del Centro de Biología Molecular Severo Ochoa, permitió conocer las condiciones adecuadas para llevar a cabo la metodología SHAPE *in vitro*. Siguiendo dicha metodología, se determinó la estructura de H12, la mutante H12d y se realizaron ensayos preliminares de “*fingerprinting*” entre H12 y RBS1; lo cual permitió trasladar dicha metodología a la caracterización estructural de los precursores de *let-7*. La interacción GEMIN5:H12 ejemplifica la relevancia de un RNA para su correcta unión con una proteína y enfatiza la relevancia estructural del RNA como elemento clave para la correcta función de una proteína (Figura 23)<sup>82</sup>.



**Figura 23 Interacción RBS1:H12.** La proteína GEMIN5 interacciona con su propio mRNA a través del dominio RBS1 en un sitio conocido como H12; dicha unión promueve la traducción del mRNA (panel A). La desestabilización del sitio H12 inhibe la unión y el efecto traduccional de RBS1 (Panel B). En contraparte, cambios puntuales en los residuos P<sub>1297</sub>S<sub>1299</sub>S<sub>1300</sub> presentes en el dominio RBS1 abaten la unión a H12 (Panel C).





## A Pilot Study on the HPV Type Frequency in a Federal High-Specialty Hospital of Mexico City: Is HPV16 Our Main Problem?

Rangel-Guerrero Sergio Israel<sup>1</sup> · Hernández-Sánchez José Basilio<sup>2</sup> · Thompson-Bonilla María del Rocío<sup>2</sup> · Rocha-Gonzalez Hector<sup>3</sup> · Alvarez-Salas Luis Marat<sup>1</sup>

Accepted: 17 February 2020 / Published online: 14 March 2020  
© Springer Nature Switzerland AG 2020

### Abstract

Human papillomavirus (HPV) is the most common sexually transmitted infection worldwide and a necessary condition for cervical cancer (CeCa) development. HPV types 16 and 18 are the most common in cervical lesions (CLs); therefore, specific vaccines have reduced cervical neoplasia. In 2015, HPV16 presented the highest frequency in CLs (~28%); however, types 31, 45, 52, and 58 were also detected, suggesting that other HPV types may have impact on healthcare. Here, we typified 30 patients with CLs from the “Hospital Regional 1° de Octubre” (HR10), in Mexico City, for the most reported high-risk papillomavirus (HPV16, 18, 31, 33, 45, 58, and 59). The results showed that HPV16 frequency is dissimilar to international and previous national data, suggesting that a deeper knowledge of HPV type distribution is required before choosing a vaccination scheme for HPV in Mexico City.

**Keywords** HPV · Cervical dysplasia · Cervical lesions · Typification · Vaccines

### Introduction

HPV infection has been proved as a necessary condition for cervical cancer (CeCa) development. Although most HPV occurrences do not cause symptoms and resolve spontaneously, persistent infection may result in disease. From over 100 different strains, HPV types 16 and 18 have been correlated with about 70% of CLs worldwide which may progress to CeCa [1, 2].

In Mexico, 4000 deaths by CeCa are registered and 8000 new cases are diagnosed annually, ranking as the second leading cause of female mortality by cancer [3]. In 2015, HPV16 presented the highest frequency in cervical tumors (57%), while HPV18 showed a minor frequency (< 8%) in agreement with HPV incidence worldwide [4]. Nevertheless, HPV types 31, 45, 52, and 58 were also detected in Mexican females, suggesting an important and unattended factor for CeCa and HPV infection control [4, 5].

HPV-related pathologies have been reduced in developed countries by prophylactic vaccination against the most prevalent HPV types (16 and 18), resulting in a dramatic decrease (~50%) of CLs and CeCa cases [6]. In Mexico City, free application of prophylactic vaccines against HPV16/18 has been approved for 10–12-year-old girls, thus becoming the leading strategy to reduce CeCa incidence, a major public health issue by decades [7]. However, no current data is available on the vaccine impact, HPV frequency, or HPV genotypes in Mexico City.

In this case report, HPV presence was typified in low-grade (LSIL) and high-grade (HSIL) squamous intraepithelial lesions obtained from 30 patients from the HR10 in Mexico City. Our data suggests that HPV16 frequency could be dissimilar to international and previous reports in Mexico. Moreover, other HPV types 31, 45, 58, and 59 were detected,

This article is part of the Topical Collection on *Medicine*

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s42399-020-00245-9>) contains supplementary material, which is available to authorized users.

✉ Alvarez-Salas Luis Marat  
lalvarez@cinvestav.mx

<sup>1</sup> Laboratorio de Terapia Génica, Departamento de Genética y Biología Molecular, Centro de Investigación y de Estudios Avanzados del I.P.N., Mexico City, Mexico




<sup>2</sup> “Hospital Regional 1° de Octubre”, Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado, Mexico City, Mexico

<sup>3</sup> Sección de Estudios de Posgrado e Investigación, Escuela Superior de Medicina del I.P.N., México City, Mexico

RESEARCH PAPER



## RNA-protein coevolution study of Gemin5 uncovers the role of the PXSS motif of RBS1 domain for RNA binding

Rosario Francisco-Velilla <sup>a</sup>, Azman Embarc-Buh <sup>a</sup>, Sergio Rangel-Guerrero<sup>b</sup>, Sudipto Basu<sup>c,d</sup>, Sudip Kundu<sup>c,d</sup>, and Encarnacion Martinez-Salas <sup>a</sup>

<sup>a</sup>Centro de Biología Molecular Severo Ochoa, CSIC-UAM, Madrid, Spain; <sup>b</sup>Laboratorio de Terapia Génica, Centro de Investigación y de Estudios Avanzados del I.P.N., Ciudad de México, Mexico; <sup>c</sup>Department of Biophysics, Molecular Biology and Bioinformatics, University of Calcutta, Kolkata, India; <sup>d</sup>Center of Excellence in Systems Biology and Biomedical Engineering, TEQIP Phase-III, University of Calcutta, Kolkata, India

### ABSTRACT

Regulation of protein synthesis is an essential step of gene expression. This process is under the control of cis-acting RNA elements and trans-acting factors. Gemin5 is a multifunctional RNA-binding protein organized in distinct domains. The protein bears a non-canonical RNA-binding site, designated RBS1, at the C-terminal end. Among other cellular RNAs, the RBS1 region recognizes a sequence located within the coding region of Gemin5 mRNA, termed H12. Expression of RBS1 stimulates translation of RNA reporters carrying the H12 sequence, counteracting the negative effect of Gemin5 on global protein synthesis. A computational analysis of RBS1 protein and H12 RNA variability across the evolutionary scale predicts coevolving pairs of amino acids and nucleotides. RBS1 footprint and gel-shift assays indicated a positive correlation between the identified coevolving pairs and RNA-protein interaction. The coevolving residues of RBS1 contribute to the recognition of stem-loop SL1, an RNA structural element of H12 that contains the coevolving nucleotides. Indeed, RBS1 proteins carrying substitutions on the coevolving residues P<sub>1297</sub> or S<sub>1299</sub>S<sub>1300</sub>, drastically reduced SL1-binding. Unlike the wild type RBS1 protein, expression of these mutant proteins in cells failed to enhance translation stimulation of mRNA reporters carrying the H12 sequence. Therefore, the PXSS motif within the RBS1 domain of Gemin5 and the RNA structural motif SL1 of its mRNA appears to play a key role in fine-tuning the expression level of this essential protein.

### ARTICLE HISTORY

Received 24 February 2020  
Revised 9 April 2020  
Accepted 24 April 2020

### KEYWORDS

RNA-protein interaction;  
coevolving pairs; Gemin5;  
RNA-binding; RNA SHAPE  
reactivity; translation control

### Introduction

RNA-binding proteins (RBPs) play a pivotal role in gene expression control [1]. Studies carried out over the years have established a number of conventional RNA-binding domains (RBDs) according to their structural composition and RNA recognition features [2]. However, recent RNA-capture procedures have increased the list of RBPs harbouring previously unknown RBDs [3,4], often including protein-protein interaction sites as well as intrinsically disordered regions (IDRs) [5,6]. Despite the knowledge gained of novel RBPs, non-conventional RBDs are rather diverse precluding the identification of amino acids directly involved in RNA-binding from the primary sequence of a putative RBD. This heterogeneity hampers the identification of novel RBPs by conventional methodologies.


Gemin5 is an abundant predominantly cytoplasmic protein involved in small nuclear ribonucleoproteins (snRNPs) assembly and translation control [7]. The protein was originally described as the RBP of the survival of motor neuron (SMN) complex [8]; subsequent studies reported a role in translation control [9–11]. Distinct domains of the protein are responsible for the recognition of specific targets, either RNAs or proteins [12]. Beyond the WD40 repeat domain located at the N-terminal region, which is involved in the recognition of snRNAs [13] and in ribosome

interaction [14], our group identified a non-conventional RNA-binding site (designated RBS1) on the C-terminal region of the protein (Fig. 1A). The RBS1 domain is responsible for the recognition of viral IRES elements [15], as well as for specific cellular mRNAs [16]. Previous attempts to study the RNA-binding site of RBS1 by NMR indicated that this region of the protein does not adopt a uniform three-dimensional structure [15], consistent with the notion of an intrinsically disordered region.

We have recently shown that the RBS1 domain interacts with cellular RNAs in human cells [16]. Of note, one of the most abundant targets of RBS1 was an internal region of Gemin5 mRNA, which included two hits (designated H12 for hits 1–2) (Fig. 1A). The results obtained from functional analysis provided the basis for a positive feedback loop involving H12 RNA and the RBS1 domain. First, RBS1 stimulates the translation of RNA reporters bearing the H12 sequence. Second, RBS1-mRNA interaction was dependent on RNA structure as a destabilized RNA (H12d) displayed a decreased binding *in vitro* and in living cells. Third, expression of RBS1 in human cells leads to an increase of Gemin5 mRNA translation [16], counteracting the negative effect of Gemin5 on global protein synthesis. However, the specific RBS1 residues responsible for H12 RNA recognition remain elusive.

**CONTACT** Encarnacion Martinez-Salas  emartinez@cbm.csic.es  Centro de Biología Molecular Severo Ochoa, CSIC-UAM, Nicolás Cabrera 1, Madrid 28049, Spain

\*These authors are Co-first authors.

 Supplemental data for this article can be accessed here.

© 2020 Informa UK Limited, trading as Taylor & Francis Group



## Structural insights of the pre-let-7 interaction with LIN28B

Sergio Israel Rangel-Guerrero<sup>a</sup>, Pablo Alberto Franco-Urquijo<sup>a</sup>, Encarnación Martínez-Salas<sup>b</sup>, and Luis Marat Alvarez-Salas<sup>a</sup>

<sup>a</sup>Laboratorio de Terapia Génica, Departamento de Genética y Biología Molecular, Centro de Investigación y de Estudios Avanzados del IPN, México City, Mexico; <sup>b</sup>Centro de Biología Molecular Severo Ochoa, Universidad Autónoma de Madrid, CSIC-UAM, Madrid, Spain

### ABSTRACT

The Let-7:LIN28 regulatory loop is a paradigm in miRNA regulation. LIN28 harbors two RNA binding domains, which interact with well-conserved sequences in pre-let-7 RNAs, the GNGAY and the GGAG motifs. Here, the differential binding between LIN28B and pre-let-7 members was associated with the structural characteristics of the pre-let-7 family mapped by SHAPE, uncovering diverse structural patterns within pre-let-7 members. Pre-let-7 mutants supported a relevant role of the GGAG motif location and the preE-stem stability for the interaction with LIN28B. Based on these results, we propose a core RNA structure for LIN28B interaction.

### ARTICLE HISTORY

Received 1 May 2020  
Accepted 30 November 2020



### KEYWORDS


MicroRNA; let-7; LIN28B;  
RNA structure; SHAPE

## 1. Introduction

MicroRNAs (miRNAs) are small non-coding RNAs of ~21 nt that target and regulate expression of several mRNAs.<sup>[1]</sup> The miRNA biogenesis starts in the nucleus by RNA Pol II/III transcription, generating a long primary transcript known as pri-miRNA.<sup>[2,3]</sup> Pri-miRNAs are cleaved in the nucleus by the *Microprocessor* complex (DROSHA/DGCR8) generating a ~70 nt RNA hairpin known as pre-miRNA.<sup>[4]</sup> After cleavage, pre-miRNAs are transported from the nucleus to the cytoplasm by the Exportin-5 complex,<sup>[5]</sup> where they are further processed by DICER<sup>[6]</sup> producing a functional double-stranded miRNA. DICER/miRNA complexes associate with Argonaute proteins<sup>[7]</sup> to generate the miRNA-induced silencing complex (RISC);<sup>[8]</sup> which specifically modulates translation of its target mRNA.

The let-7 family is a wide-spread evolutionary conserved miRNA family, present in several metazoan species.<sup>[9,10]</sup> In humans, the let-7 family includes 13 members located in nine different chromosomes.<sup>[11]</sup> Let-7

CONTACT Luis Marat Alvarez-Salas  [lalvarez@cinvestav.mx](mailto:lalvarez@cinvestav.mx)  Laboratorio de Terapia Génica, Departamento de Genética y Biología Molecular, Centro de Investigación y de Estudios Avanzados del IPN, México City 07360, Mexico.

 Supplemental data for this article is available online at <https://doi.org/10.1080/15257770.2020.1859116>.

© 2020 Taylor & Francis Group, LLC