

# Estudio computacional de las interacciones entre el DMSO y ARN

Pedro Andrés Guzmán Cagua, Cód: 2200218

Trabajo de Grado para Optar por el Título de Químico

Directora

Doctora en Ciencias - Química. Martha Cecilia Daza Espinosa

Codirector

Dr. rer. nat. Markus Hans Oliver Doerr



Grupo de Bioquímica Teórica - GBQT

Universidad Industrial de Santander

Facultad de Ciencias

Escuela de Química

Pregrado

2025

## Agradecimientos

Les agradezco profundamente a mi padre y madre por la sangre, sudor y lágrimas derramadas en estos 23 años. A mis hermanos por su apoyo. A los profesores Martha Daza y Markus Doerr por guiarme durante este proceso, por el apoyo económico de transporte y para participar en el congreso SICTeI 2024. Al Grupo de Bioquímica Teórica por el apoyo económico brindado mediante auxilias (proyectos 3744 y 4331). A las personas que han regado mis raíces y en las que he dejado una semilla. Y al Dr. Charly García por su música. Say No More.

## Tabla de contenido

	Pag.
Agradecimientos .....	2
Introducción .....	10
1. Objetivos .....	11
1.1 Objetivo general.....	11
1.2 Objetivos específicos .....	11
2. Estado del arte y planteamiento del problema .....	12
3. Marco teórico .....	14
3.1 Estructura del ARN .....	16
3.2 Interacciones no covalentes .....	21
3.2.1 Enlaces de hidrógeno .....	22
3.3 Método de predicción de la estructura del ARN: FARFAR2.....	23
3.4 Dinámica molecular .....	25

3.5	Campo de fuerza .....	25
3.6	Dimetil sulfóxido (DMSO).....	27
4.	Metodología .....	29
4.1	Selección del fragmento de ARN.....	29
4.2	Determinación de la estructura secundaria del fragmento de ARN .....	29
4.3	Dinámicas moleculares .....	29
4.3.1	Dinámicas de ARN en agua .....	31
4.3.2	Dinámicas de ARN en solución de DMSO al 5% v/v .....	32
5.	Resultados .....	34
5.1	Determinación y caracterización estructural del fragmento de ARN .....	34
5.1.1	Dinámica conformacional del fragmento de ARN en fase acuosa y en presencia de DMSO	38
5.1.2	Estabilidad de la estructura secundaria del ARN: Análisis de los enlaces de hidrógeno intramoleculares .....	40
5.1.3	Caracterización de las interacciones intermoleculares no covalentes: ARN-agua y ARN-DMSO .....	42
6.	Conclusiones .....	64
	Recomendaciones .....	65
	Referencias bibliográficas.....	66
	Apendice .....	71

### Lista de Tablas

<b>Tabla 1.</b> Caracterización de los pares de bases, tipos de apareamientos, motivos de estructura secundaria y apilamiento entre bases del fragmento de ARN.....	36
---	----

### Lista de Esquemas

<b>Esquema 1.</b> Diagrama de flujo para la producción de las dinámicas moleculares .....	30
<b>Esquema 2.</b> Visualización de los enlaces de hidrógeno formados entre el dmsó y el fragmento de ARN.....	57

### Lista de Figuras

<b>Figura 1.</b> Representación esquemática de las escalas de tiempo relevantes para la dinámica estructural del ARN.....	13
<b>Figura 2.</b> Enlaces fosfodiéster del esqueleto covalente del ADN Y del ARN .....	14
<b>Figura 3.</b> Bases púricas y pirimidínicas de los ácidos nucleicos .....	15
<b>Figura 4.</b> Patrones de enlace de hidrógeno en los pares de bases definidos por Watson Y Crick .....	16
<b>Figura 5.</b> Niveles estructurales del ARN .....	17
<b>Figura 6.</b> Motivos estructurales del ARN .....	17
<b>Figura 7.</b> Estructura tridimensional del ARN .....	19
<b>Figura 8.</b> Bordos de interacción de las bases nitrogenadas y familias sus geométricas .....	20
<b>Figura 9.</b> Potencia de energía de enlace.....	21
<b>Figura 10.</b> Algoritmo de FARFAR2 de predicción estructural de ARN. ....	24

<b>Figura 11.</b> Esquema ilustrativo de las principales contribuciones de un campo de fuerza de dinámica molecular .....	26
<b>Figura 12.</b> Estructura del DMSO .....	28
<b>Figura 13.</b> Estructura secundaria más estable del fragmento de ARN predicha con FARFAR2	34
<b>Figura 14.</b> Motivos de estructura secundaria del fragmento de ARN de 30 nucleótidos perteneciente al gen humano DRD4 .....	35
<b>Figura 15.</b> Representación 3D del fragmento de ARN para el análisis del apareamiento y apilamiento entre bases nitrogenadas.....	37
<b>Figura 16.</b> Visualización 3D del fragmento de ARN .....	38
<b>Figura 17.</b> Flexibilidad local del fragmento de ARN.....	39
<b>Figura 18.</b> Flexibilidad global del fragmento de ARN .....	40
<b>Figura 19.</b> Estabilidad de los enlaces intramoleculares del fragmento de ARN.....	41
<b>Figura 20.</b> Distribución radial de las moléculas de agua (OW) alrededor de los grupos donadores de enlaces de hidrógeno de los residuos del fragmento de ARN .....	43
<b>Figura 21.</b> Distribución radial de las moléculas de agua (HW) alrededor los grupos aceptores de enlaces de hidrógeno de los residuos del fragmento de ARN.....	46
<b>Figura 22.</b> Distribución radial de las moléculas de DMSO (O-DMSO) alrededor los grupos donadores de enlaces de hidrógeno de los residuos del fragmento de ARN .....	50
<b>Figura 23.</b> Conteo de enlaces de hidrógeno ARN-DMSO.....	54
<b>Figura 24.</b> Distribución de las distancias de enlace de hidrógeno ARN-DMSO .....	55
<b>Figura 25.</b> Nucleótidos con mayores tiempos de residencia de enlace de hidrógeno con el DMSO .....	55
<b>Figura 26.</b> Conteo de enlaces de hidrógeno del ARN y el backbone con el agua .....	60

**Figura 27.** Conteo de enlaces de hidrógeno de cada nucleótido con el agua..... 61

**Figura 28.** Distribución de las distancias de enlace de hidrógeno ARN-agua..... 64

### Lista de Apéndices

**Figura A. 1.** Paisaje de energía libre 3D y las conformaciones de menor energía libre en cada réplica del fragmento de ARN ..... 78

**Figura A. 2** Mapa de calor de la estabilidad de los enlaces de hidrógeno intramoleculares entre bases nitrogenadas ..... 79

**Figura A. 3.** Grupos donadores y aceptores de enlaces de hidrógeno de los nucleótidos del ARN ..... 81

**Figura A. 4.** Comparación del SASA de cada nucleótido en las simulaciones..... 82

**Figura A. 5.** Mapa de calor de los tiempos de residencia de los enlaces de hidrógeno ARN-DMSO..... 83

**Figura A. 6.** Mapa de calor de los tiempos de residencia de los enlaces de hidrógeno ARN-agua del sistema de ARN en fase acuosa..... 84

**Figura A. 7.** Mapa de calor de los tiempos de residencia de los enlaces de hidrógeno ARN-agua del sistema de ARN CON DMSO al 5% V/V..... 86

### Glosario

**ARN:** Ácido ribonucleico

**DMSO:** dimetil sulfóxido

**DRD4:** Receptor dopamínico D4

**FARFAR2:** Fragment Assembly of RNA with Full-Atom Refinement 2

**CHARMM36:** Chemistry at Harvard Macromolecular Mechanics (Campo de fuerza)

**RDF:** Función de distribución radial

**RMSD:** Desviación cuadrática media

**RMSF:** Fluctuación cuadrática media

**SASA:** Área superficial accesible por el solvente

**MD:** Dinámica Molecular.

**PME:** Particle Mesh Ewald (Algoritmo para interacciones electrostáticas).

**LINCS:** Linear Constraint Solver (Algoritmo para restricciones).

**TIP3P:** Transferable Intermolecular Potential 3 Point (El modelo de agua).

**NVT:** Ensemble Canónico (Algoritmo de equilibración).

**NPT:** Ensemble Isotérmico-Isobárico (Algoritmo de equilibración).

**GROMACS:** Groningen Machine for Chemical Simulations (software usado).

**Hairpin:** Horquilla (Motivo estructural del ARN).

**Wobble:** Apareamiento de bases tipo Wobble (Tipo de apareamiento no canónico).

**Backbone:** Columna vertebral del ARN conformada por la cadena de enlaces fosfodiéster.

## Resumen

**Título:** Estudio computacional de las interacciones entre el DMSO Y ARN

**Autor:** Pedro Andrés Guzmán Cagua

**Palabras clave:** Dinámica molecular, ARN, DMSO, interacciones electrostáticas.

**Descripción:** El dimetilsulfóxido (DMSO) es un solvente orgánico ampliamente utilizado como agente desnaturalizante para mejorar la eficiencia de pruebas de diagnóstico como la PCR, al desestabilizar la estructura secundaria del ARN. Sin embargo, el mecanismo de interacción mediante el cual el DMSO induce la desestabilización de la estructura secundaria del ARN aún no es completamente elucidado. Por lo tanto, en el presente trabajo se aborda esta brecha de conocimiento evaluando el efecto del DMSO al 5% v/v sobre la estructura secundaria de un fragmento de 30 nucleótidos del gen humano DRD4 mediante simulaciones de dinámica molecular en la escala de microsegundos.

Los resultados muestran que en la escala de tiempo simulada (2 $\mu$ s), el DMSO al 5% v/v no induce la desnaturalización global de la estructura secundaria del ARN. Sin embargo, la molécula de ARN presenta una dinámica conformacional ligeramente más flexible en comparación con las simulaciones en fase acuosa. Se identificó que el DMSO forma enlaces de hidrógeno transitorios y de bajo tiempo de residencia con el ARN. Estas interacciones muestran una ligera preferencia por las guaninas y uracilos de motivos estructurales flexibles, pero no se observó que el DMSO compitiera directamente por los enlaces de hidrógenos intramoleculares entre las bases nitrogenadas que estabilizan la estructura secundaria.

**Abstract**

**Title:** Computational study of interactions between DMSO and RNA

**Author:** Pedro Andrés Guzmán Cagua

**Keywords:** Molecular Dynamics, RNA, DMSO, electrostatic interactions.

**Description:** Dimethyl sulfoxide (DMSO) is an organic solvent widely used as a denaturing agent to improve the efficiency of diagnostic tests, such as PCR, by destabilizing the secondary structure of RNA. However, the interaction mechanism by which DMSO induces this destabilization is not yet fully elucidated. Therefore, this work addresses this knowledge gap by evaluating the effect of 5% v/v DMSO on the secondary structure of a 30-nucleotide fragment of the human DRD4 gene through microsecond-scale molecular dynamics simulations.

The results show that on the simulated timescale (2  $\mu$ s), 5% v/v DMSO does not induce global denaturation of the RNA secondary structure. However, the RNA exhibits slightly more flexible conformational dynamics compared to simulations in the aqueous phase. It was identified that DMSO forms transient hydrogen bonds with low residence times with the RNA. These interactions show a slight preference for guanine and uracil residues in flexible structural motifs, but DMSO was not observed to compete directly for the intramolecular hydrogen bonds that stabilize the secondary structure.

## Introducción

Las metodologías computacionales, como la dinámica molecular, se han convertido en herramientas indispensables para analizar las propiedades estructurales y las interacciones de biomoléculas a nivel atómico. El avance de esta área investigativa ha impulsado el desarrollo de aplicaciones terapéuticas y farmacológicas al facilitar una comprensión profunda de los mecanismos moleculares que gobiernan la función de proteínas, ácidos nucleicos y membranas biológicas y su interacción con diversos sustratos.

Entre estas biomoléculas, el ARN ha emergido como un objetivo central de estudio en la última década, debido a su rol fundamental en procesos biológicos clave, desde la transcripción y traducción de la información genética en los seres vivos (ARNm y ARNt, ARNr) (*Lehninger Principles of Biochemistry, 8th Edition* | Macmillan Learning US, 2021.) hasta la regulación de la expresión genética (micro-ARN). La función del ARN está ligada a su capacidad de adoptar complejas estructuras tridimensionales. Sin embargo, esta complejidad estructural del ARN plantea desafíos para su estudio y resulta un obstáculo en aplicaciones biotecnológicas, ya que el complejo estructural del ARN se compone de cuatro niveles estructurales: estructura primaria (secuencia lineal de nucleótidos), estructura secundaria (enlaces de hidrógeno entre pares de bases nitrogenadas, estructura terciaria (plegamiento tridimensional por interacciones intramoleculares) y estructura cuaternaria (complejos con otras biomoléculas).

Un área de estudio del ARN es la amplificación de secuencias virales por transcripción inversa mediante la técnica PCR con fin de diagnóstico. Donde se usan solventes como el DMSO para desestabilizar y promover la desnaturalización del ARN (Wilcox et al., 2019) para mejorar la eficiencia de la transcripción, la capacidad de diagnóstico y reducir el número de resultados falsos negativos causados por las estructuras secundarias estables en secuencias de ARN viral que inhiben

la eficiencia de la transcripción en las pruebas PCR (Cadena-Caballero et al., 2022; Wilcox et al., 2019).

A pesar de su uso extendido y eficacia demostrada (Cadena-Caballero et al., 2022; Hardjasa et al., 2010; Wilcox et al., 2019), existe una laguna en el conocimiento del mecanismo de interacción a nivel atómico entre el DMSO y el ARN. Estudios afines se han enfocado en otros aspectos y sistemas (Stachura et al., 2018; Tariq et al., 2022; Tunçer et al., 2018), y no responden cómo el DMSO promueve la desestabilización de la estructura secundaria del ARN y si existe preferencia de interacción con motivos estructurales y residuos específicos. Para abordar esta brecha, este trabajo de grado estudia las interacciones del DMSO con la estructura secundaria de un fragmento de ARN de 30 ribonucleótidos perteneciente al gen humano DRD4 (Human Dopamine Receptor D4, NM\_000797.4 (Nöthen et al., 1994)) mediante dinámica molecular a escala de microsegundos para caracterizar las interacciones con resolución atómica.

## 1. Objetivos

### 1.1 Objetivo general

- Evaluar el efecto del DMSO sobre la estructura secundaria de un fragmento de ARN de 30 ribonucleótidos.

### 1.2 Objetivos específicos

- Determinar el tipo de interacciones no covalentes que definen la estructura secundaria de una secuencia de 30 ribonucleótidos.
- Evaluar las interacciones no covalentes entre el DMSO y el ARN.

## 2. Estado del arte y planteamiento del problema

El estudio computacional del ARN ha experimentado un crecimiento notable en el último lustro, impulsado en gran medida por la pandemia del virus SARS-CoV-2. En este contexto, una investigación colaborativa entre el Centro de Supercomputación y Cálculo Científico de la UIS (SC3UIS) y el Grupo de Investigación en Cómputo Avanzado y a Gran Escala (CAGE) de la Universidad Industrial de Santander (UIS) realizaron un estudio que combinó parte experimental y análisis computacional, en el que se demostró la mejora en la precisión y velocidad de diagnóstico de las pruebas PCR mediante el uso de agentes desnaturizantes de secuencias de ARN con Dimetilsulfóxido (DMSO) y Tetraetilamonio (TEA) (Cadena-Caballero et al., 2022). Esta investigación constituye el punto de partida de la presente investigación.

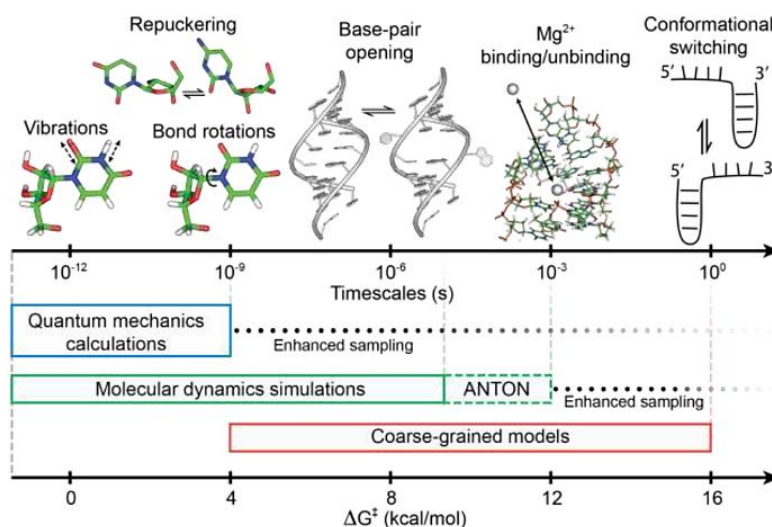
Sin embargo, a pesar de la evidencia experimental sobre la eficacia del DMSO para desestabilizar la estructura de los ácidos nucleicos y su uso para optimizar técnicas de diagnóstico (Chakrabarti & Schutt, 2001; Hardjasa et al., 2010; Lestarini et al., 2019; Wilcox et al., 2019), la comprensión del mecanismo de interacción a nivel atómico entre el DMSO y el ARN es limitada (Lee et al., 2013). Los estudios computacionales análogos existentes se han centrado en las interacciones DMSO-agua (Stachura et al., 2018), el efecto de diferentes concentraciones de DMSO en la alteración de procesos celulares mediante la modulación conformacional de biomoléculas, como la transición conformacional de ADN-B a ADN-Z (Tunçer et al., 2018) y el efecto estabilizante indirecto del DMSO sobre los G-quadruplexes mediante la modulación de las interacciones solvente-catión-ADN (Tariq et al., 2022).

Por lo tanto, mediante el presente trabajo de grado se propone estudiar las interacciones entre el DMSO y el ARN que promueven la desestabilización de la estructura secundaria mediante simulaciones de dinámica molecular con el paquete Gromacs, metodología idónea para investigar

las interacciones inter e intramoleculares con resolución atómica. Se seleccionó el campo de fuerza CHARMM36 usado ampliamente para modelar ácidos nucleicos (Huang & MacKerell Jr, 2013; Šponer et al., 2018). Además, se usa el software FARFAR2 (Watkins et al., 2020) para la predicción y determinación de la estructura de partida.

A pesar de las ventajas del método, existen limitaciones asociadas con el coste computacional ya que se requieren simular escalas de tiempo grandes ( $\mu\text{s}$  a  $\text{s}$ ) para observar transiciones conformacionales drásticas y acoplamiento ligando-ARN (Šponer et al., 2018), *Figura 1*. los resultados de este estudio no sólo ampliarán el conocimiento fundamental sobre este sistema, sino que también fortalecerán las capacidades metodológicas del Grupo de Bioquímica Teórica (GBQT).

**Figura 1.** Representación esquemática de las escalas de tiempo relevantes para la dinámica estructural del ARN



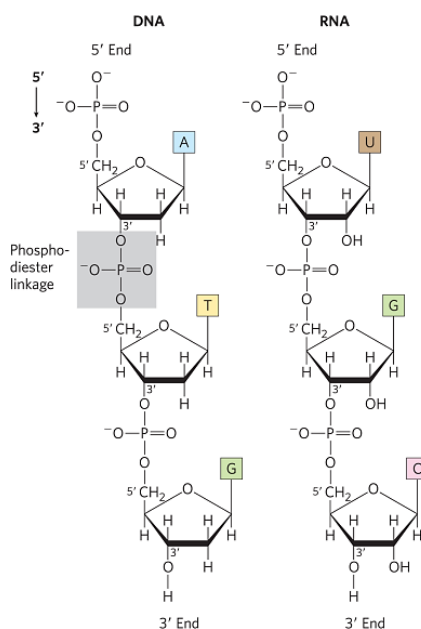
*Nota.* Figura tomada de Šponer, J., et al. (2018). RNA Structural Dynamics as Captured by Molecular Simulations: A Comprehensive Overview. *Chemical Reviews*, 118(8), 4177–4338.

### 3. Marco teórico

Los ácidos nucleicos son biomoléculas de alto interés biológico debido a su rol en los procesos de conservación y evolución de la vida. El ARN es la principal biomolécula encargada de los procesos de transcripción y traducción de la información genética en las células, siendo el encargado de almacenar, regular, y participar en la expresión de la información genética. En estos procesos biológicos (i.e. la transcripción, traducción y regulación de la expresión genética) participan diferentes tipos de ARN como el ARNm, ARNt, ARNr, microARN y ncARN, que poseen una compleja conformación tridimensional que guarda una fuerte relación con su función biológica (Lehninger Principles of Biochemistry, 8th Edition | Macmillan Learning US, 2017).

Los ácidos nucleicos (ARN y ADN) son polinucleótidos cuya unidad estructural son los ribonucleótidos y desoxirribonucleótidos, unidos mediante enlaces fosfodiéster entre el grupo fosfato ( $5'PO_4^{3-}$ ) y el  $3'OH$  de la ribosa o de la desoxirribosa, en sentido  $5'$  a  $3'$ , y se diferencian principalmente por la presencia del grupo  $2'OH$  de la ribosa del backbone, *Figura 2*.

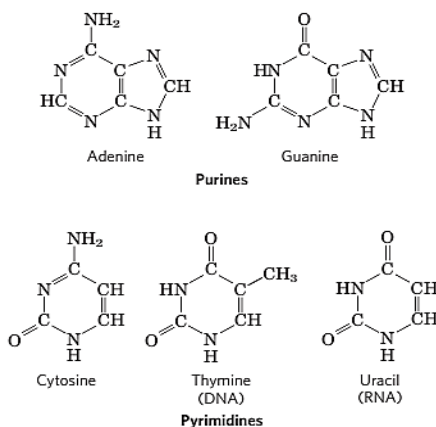
**Figura 2.** Enlaces fosfodiéster del esqueleto covalente del ADN y del ARN



*Nota.* Tomado de Nelson, D. L., & Cox, M. M. (2017). *Lehninger principles of Biochemistry* (8th ed., p. 1038). Macmillan learning.

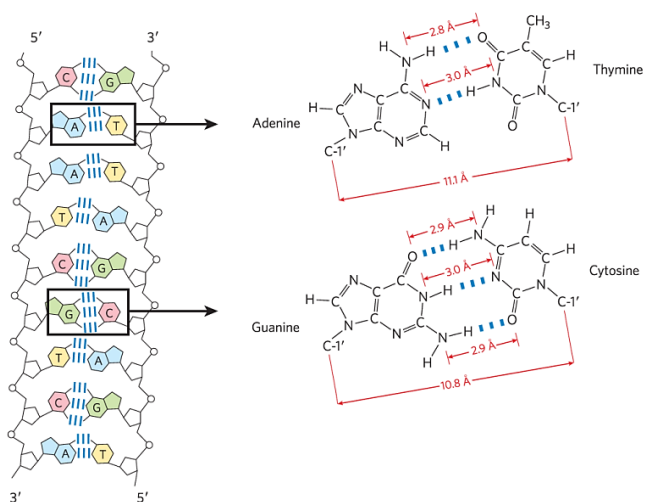
Los nucleótidos, unidades monoméricas, esencialmente están conformados por tres grupos, una ribosa o una desoxirribosa, un grupo fosfato y una base nitrogenada (purina o pirimidina). Las pirimidinas (i.e. citosina (CYT), uracilo (URA) y timina (THY)) presentan una estructura de anillo de seis átomos y las purinas adenina (i.e. (ADE) y guanina (GUA)) tienen una estructura de doble anillo de nueve átomos, *Figura 3*.

**Figura 3.** Principales bases púricas y pirimidínicas de los ácidos nucleicos



*Nota.* Tomado de Nelson, D. L., & Cox, M. M. (2017). *Lehninger Principles of biochemistry* (8th ed., p. 1030). Macmillan learning.

Las bases nitrogenadas interactúan complementariamente mediante enlaces de hidrógeno, formando apareamientos de diferentes tipos (e.g. canónicos o de Watson-Crick entre Adenina-Uracilo (AU) y Guanina-Citosina (GC)). Durante el apareamiento (A=U) se forman dos enlaces de hidrógeno y durante el apareamiento (G≡C) se forman tres enlaces de hidrógeno *Figura 4*.

**Figura 4.** Patrones de enlace de hidrógeno en los pares de bases definidos por Watson y Crick

*Nota.* Tomado de Nelson, D. L., & Cox, M. M. (2017). *Lehninger principles of biochemistry* (8th ed, p. 1044). Macmillan learning.

Adicionalmente a los apareamientos canónicos, existe otro tipo de apareamiento entre bases presente en estructuras específicas de ARN con funciones relevantes en procesos biológicos (e.g. El apareamiento tipo Wobble, presente en el ARNt (entre G=U o I=U; donde I = Imidazol, una base nitrogenada derivada de la guanina (Varani & McClain, 2000) y otros apareamientos no canónicos como los apareamientos de tipo Hoogsteen y de Sugar Edge, más información en la sección 3.1.

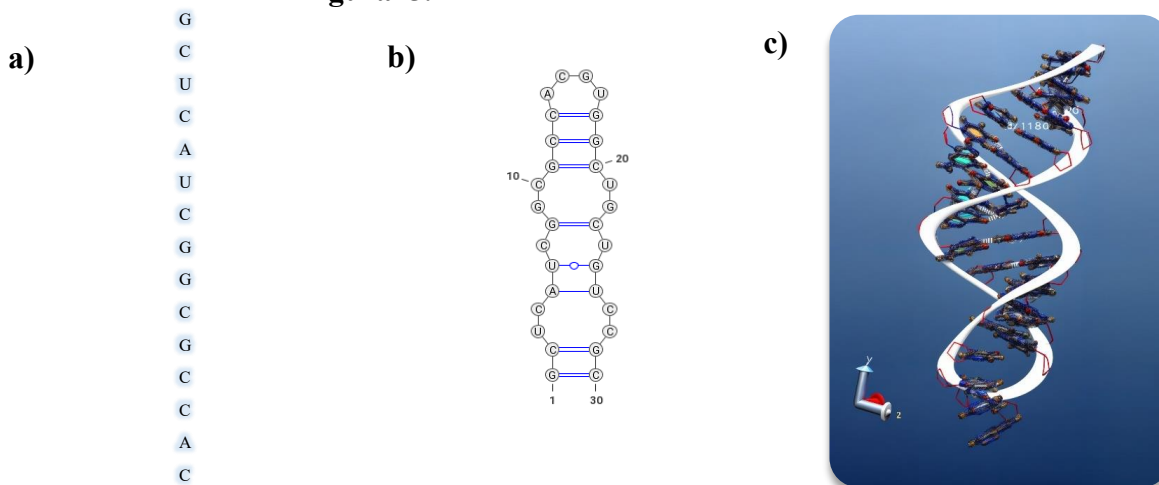
### 3.1 Estructura del ARN

La estructura del ARN comprende de arquitecturas jerárquicas compuestas por motivos estructurales recurrentes que son estabilizados por interacciones no covalentes intramoleculares (i.e. enlaces de hidrógeno, interacciones  $\pi$ - $\pi$ , interacciones de Van der Waals y fuerzas de London) entre los residuos del ARN (e.g. bases nitrogenadas, ribosas y fosfatos). Los motivos estructurales del ARN cumplen roles específicos asociados a estructuras funcionales nativas. La estructura del

ARN se compone de cuatro niveles de estructura: estructura primaria, estructura secundaria, estructura terciaria y estructura cuaternaria.

La estructura primaria es la base de la formación de los demás niveles estructurales, y está conformada por la secuencia linear de nucleótidos de la molécula de ARN, *Figura 5.a*.

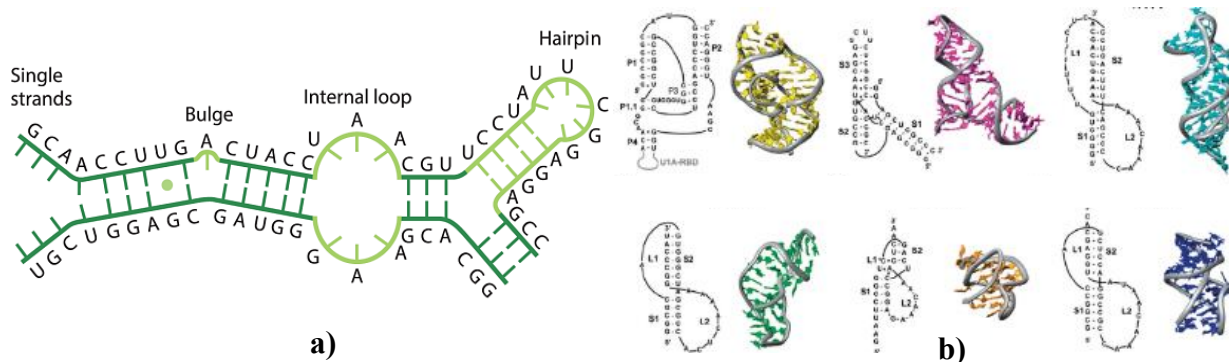
**Figura 5.** Niveles estructurales del ARN



*Nota.* **a)** Estructura primaria, secuencia linear de ribonucleótidos. **b)** Estructura secundaria estabilizada por enlaces de hidrógeno entre pares de bases nitrogenadas, representación 2D representación hecha con VARNA (Darty et al., 2009), **c)** Estructura tridimensional, representación hecha con VMD (Humphrey et al., 1996).

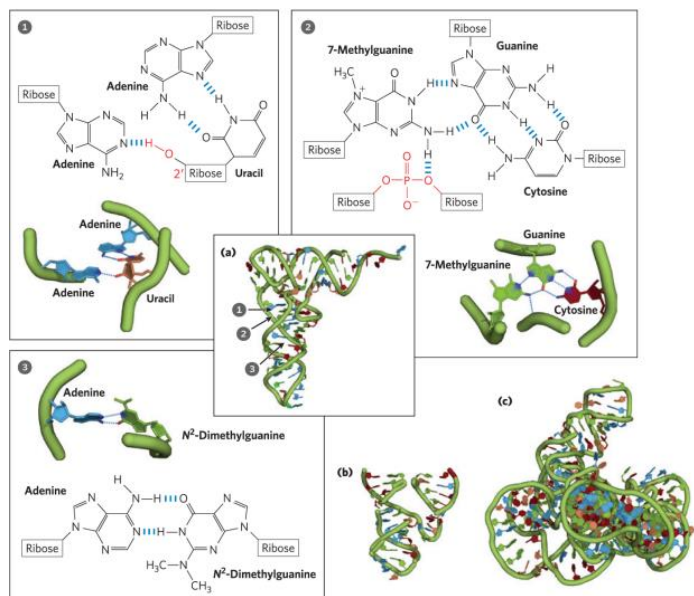
La estructura secundaria está determinada por los enlaces de hidrógeno entre los grupos donadores (-NH<sub>2</sub>, -NH y -CH) y los grupos aceptores (-CO y -N) de las bases nitrogenadas. Estas interacciones a lo largo de la cadena forman patrones recurrentes o también llamados motivos estructurales, los cuales dirigen el plegamiento de la molécula, *Figura 5.b* y *Figura 6.a*. Los motivos más comunes de estructura secundaria del ARN son: Stem (tallo), secuencia consecutiva de 4 pares de bases en adelante, tridimensionalmente se observa como una doble hélice; bulge (bulto), presencia de una base desapareada en medio de un tallo; internal loop (loop interno), zonas abiertas por bases desapareadas; hairpin (horquilla), secuencia de pares de bases en forma de tallo

con un extremo abierto tipo loop; y pseudoknots (pseudonudos), apareamiento entre bases de diferentes segmentos de la misma molécula, relevantes en el plegamiento tridimensional de la molécula, estos motivos se observan en la *Figura 6*.



*Nota.* **a)** Estructura secundaria del ARN, **b)** Secuencias y estructuras de pseudonudos de ARN. Tomado de Nelson, D. L., & Cox, M. M. (2017). *Lehninger Principles of Biochemistry* (8th ed). 1044). Macmillan learning. Staple, D. W., Butcher, S. E. (2005). Tomado de In D. W. Staple, S. E.; Butcher, S. E. Pseudoknots: RNA structures with diverse functions. *PLoS Biology*, 3(6) («(PDF) Pseudoknots», 2024).

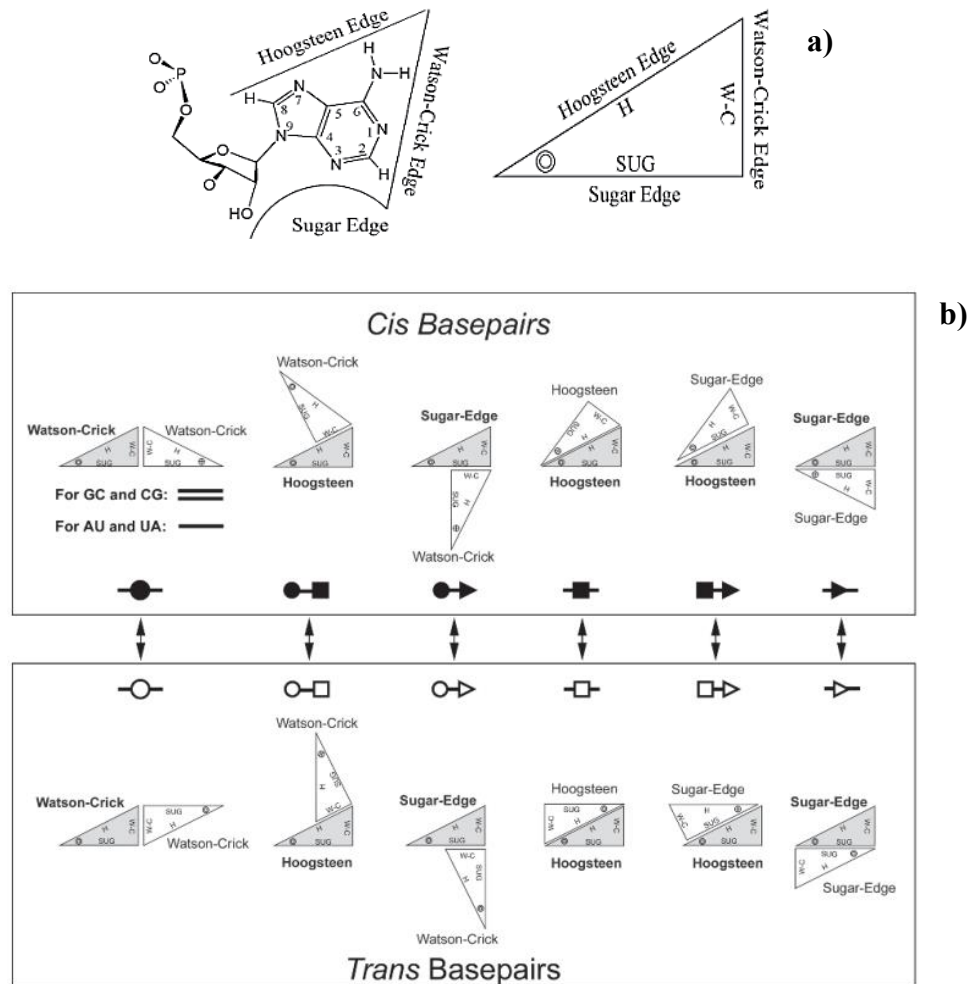
La estructura terciaria del ARN está asociada a las interacciones que dirigen el plegamiento tridimensional de la molécula para adoptar la conformación más estable en términos de energía libre de Gibbs ( $\Delta G$ ). La conformación 3D del ARN está asociada a funciones en procesos estructurales o enzimáticos. Las interacciones que hacen parte de los motivos de estructura terciaria del ARN son: los apareamientos canónicos y no canónicos entre bases, el apilamiento de bases, interacciones base-backbone, e interacciones entre el backbone de diferentes segmentos de la estructura, *Figura 7*. (e.g. fosfato-ribosa, fosfato-fosfato y ribosa-ribosa). Algunos motivos de estructura terciaria comunes son: apilamiento coaxial, kissing loops, kink turns, U•A-U triplets and G-quadruplexes (Tariq et al., 2022; Walter et al., 2009).

**Figura 7.** Estructura tridimensional del ARN

Nota. Tomado de Nelson, D. L., & Cox, M. M. (2017). *Lehninger Principles of Biochemistry* (8th ed). 1074). Macmillan learning.

Los apareamientos no canónicos se forman debido a la alta flexibilidad del backbone del ARN, ya que ocurren torsiones y variaciones de la orientación en la que interactúan las bases nitrogenadas, principalmente por torsiones del enlace glucosídico base-ribosa. Las bases presentan tres bordes de interacción para formar enlaces tipo puente de hidrógeno, estos bordes son: Borde de Watson-Crick (W-C), borde de Azúcar (SUG) y borde de Hoogsteen (H) (Hoehndorf et al., 2011), *Figura 8 a*).

**Figura 8.** Bordes de interacción de las bases nitrogenadas y familias sus geométricas



Nota. Tomado de Nasalean, L., Stombaugh, J. and Zirbel, C. L., & Leontis, N. B. (2009). *Non-Protein Coding RNAs (Vol. 13, pp. 6 y 9)*. Springer, Berlin, Heidelberg.

Los apareamientos no canónicos entre bases nitrogenadas se organizan en 12 familias geométricas, formados mediante la combinación de los seis tipos acercamientos entre los tres bordes de interacción (i.e. Watson-Crick, Hoogsteen, Sugar) en sentido cis- o trans-, *Figura 8*.

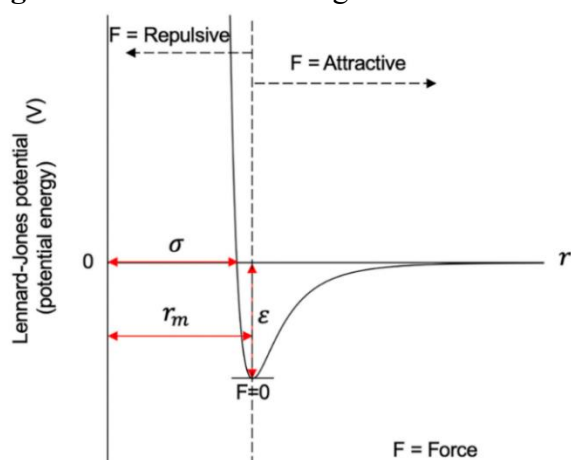
Por último, las interacciones del ARN con otras biomoléculas (e.g ARN-ARN, ARN-ADN y ARN-proteína) se denominan la estructura cuaternaria del ARN.

### 3.2 Interacciones no covalentes

Más allá del concepto de enlace químico; definidos como iónicos, metálicos y covalentes, se conoce la capacidad de interacción de la materia mediante interacciones no covalentes, fuerzas atractivas y repulsivas encargadas de que las moléculas o átomos permanezcan unidos e impiden su colapso. Las interacciones no covalentes son fuerzas débiles y reversibles y no implican el intercambio o la transferencia de electrones, a diferencia de los enlaces covalentes.

Las fuerzas electromagnéticas de átomos y moléculas producen las siguientes interacciones no covalentes: interacciones electrostáticas, los enlaces de hidrógeno, las interacciones de Van der Waals, fuerzas de dispersión de London y los efectos hidrofóbicos. Son fundamentales para la estructura y la función de las macromoléculas biológicas, ya que influyen en el plegamiento de las proteínas, la estabilidad del ADN y el reconocimiento molecular. La intensidad de la fuerza de interacción entre moléculas o átomos depende de la distancia y la orientación relativa de estas partículas. El comportamiento general de la energía de interacción se representa en la *Figura 9*.

**Figura 9.** Potencia de energía de enlace



Nota. El punto mínimo de energía de la gráfica es la distancia en que las fuerzas atracción y repulsión se compensan y se define como la distancia de enlace en equilibrio.

Las interacciones comúnmente estudiadas son las interacciones de Van del Waals, formadas por dipolos permanentes con dipolos permanentes o inducidos, y los enlaces de hidrógeno, formados entre un átomo electronegativo con un par de electrones libres que se une a un grupo de un hidrógeno y un átomo electronegativo. Estas interacciones son las principales encargadas de estabilizar estructura y definir el plegamiento de las biomoléculas, y están presentes en las interacciones de las biomoléculas con las moléculas del medio en que se encuentran (Ingold, 1940).

### **3.2.1 Enlaces de hidrógeno**

Los enlaces de hidrógeno juegan un rol fundamental en la estructura de gran variedad de sistemas, presentes en la estructura del agua y en la estructura de las biomoléculas de los seres vivos. Los enlaces de hidrógeno se caracterizan por ser interacciones estabilizantes formados entre un grupo donador AH; un hidrógeno unido a un átomo electronegativo y un grupo aceptor B; un átomo electronegativo con un par de electrones libres. Los átomos aceptores más electronegativos son F, O y N.

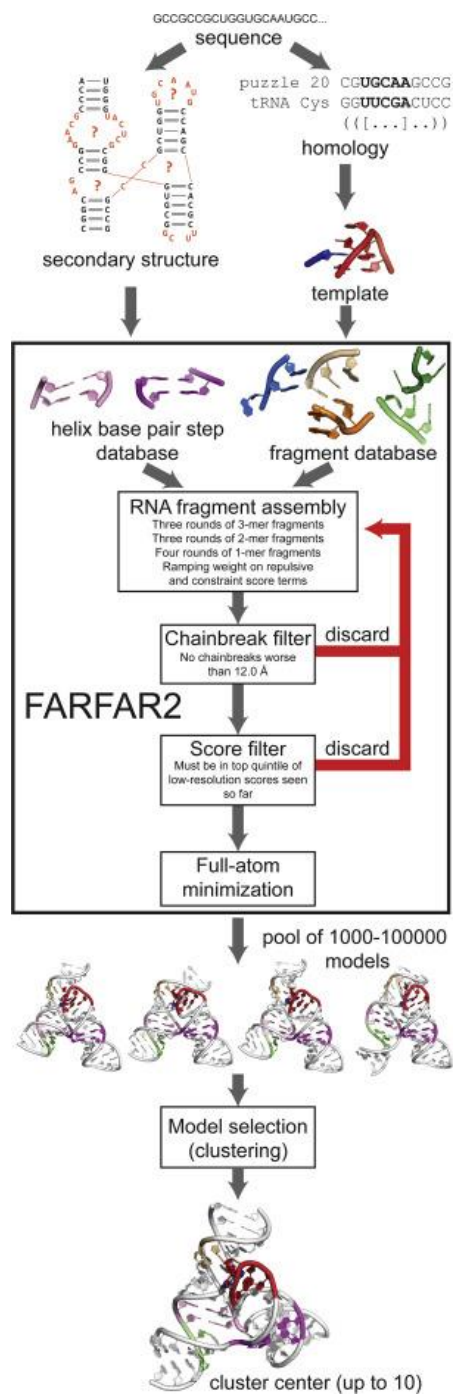
La fortaleza de los enlaces de hidrógeno depende de la electronegatividad de los átomos A y B involucrados, de la distancia entre ellos, el ángulo y la densidad electrónica de los grupos que interactúan. Los enlaces de hidrógeno se catalogan como fuertes, cuando la distancia entre el hidrógeno y el aceptor ( $H\cdots B$ ) está entre 1.2 y 1.5 Å. Estos enlaces tienen un considerable carácter covalente y energías de enlace altas ( $> 15$  kcal/mol). Presentan una mayor flexibilidad geométrica, con ángulos más variables. Las distancias  $H\cdots B$  son más largas, usualmente entre 2.0 y 3.0 Å. Su naturaleza es principalmente electrostática y sus energías de enlace son bajas ( $< 4$  kcal/mol) (Steiner, 2002).

### 3.3 Método de predicción de la estructura del ARN: FARFAR2

Actualmente, se han desarrollado diversos métodos computacionales para la predicción estructural de proteínas y ácidos nucleicos. Estos métodos se clasifican principalmente en tres tipos de acuerdo a su enfoque: los basados en algoritmos evolutivos, los que emplean conocimiento previo y los que utilizan aprendizaje automático (métodos de machine learning (ML) y deep learning (DL))(Zhang et al., 2022; Zhao et al., 2021). A lo largo de la última década, estos métodos han experimentado un avance considerable, en gran medida gracias al progreso en el desarrollo de modelos computacionales y a la creciente disponibilidad de datos estructurales experimentales (Laing & Schlick, 2011), lo que ha mejorado significativamente la precisión y capacidad de predicción.

En el presente trabajo de grado se emplea el software FARFAR2 (Fragment Assembly of RNA with Full-Atom Refinement) (Watkins et al., 2020), método basado en algoritmos evolutivos para la predicción estructural de moléculas de ARN. Este método combina 3 etapas para la construcción de la estructura tridimensional del ARN en formato PDB. En la primera etapa de baja resolución, se construyen posibles conformaciones de la cadena a partir del ensamblaje de fragmentos de 1-3 nucleótidos de longitud a partir de la secuencia nucleotídica con base en el modelo de Montecarlo. Después, se optimiza la geometría utilizando una función de puntuación geométrica mediante un modelo de DL entrenado a partir de estructuras caracterizadas experimental y teóricamente. A continuación, se aplica una minimización de energía, obteniendo predicciones de alta resolución y estabilidad. Por último, se evalúan las energías y se seleccionan los mejores modelos estructurales predichos, la metodología de predicción descrita observa en la *Figura 10*.

**Figura 10.** Algoritmo de FARFAR2 de predicción estructural de ARN.



Nota. Tomado de Watkins, A. M., Rangan, R., & Das, R. (2020). FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. Structure, CellPress (Watkins et al., 2020).

### 3.4 Dinámica molecular

La dinámica molecular es una técnica computacional que permite modelar las propiedades estructurales e interacciones de macromoléculas y moléculas pequeñas mediante simulaciones con base en la mecánica clásica (i.e. Las ecuaciones de movimiento de Newton), por lo que permite estudiar la evolución de sistemas físicos, químicos o biológicos en condiciones específicas de temperatura y presión. Para ello se implementan parámetros, algoritmos y campo de fuerza pertinentes al sistema estudiado, para reproducir las posiciones atómicas, dinámica de los enlaces, interacciones no covalentes de largo y corto alcance, y energía potencial del sistema (Leach, 2001).

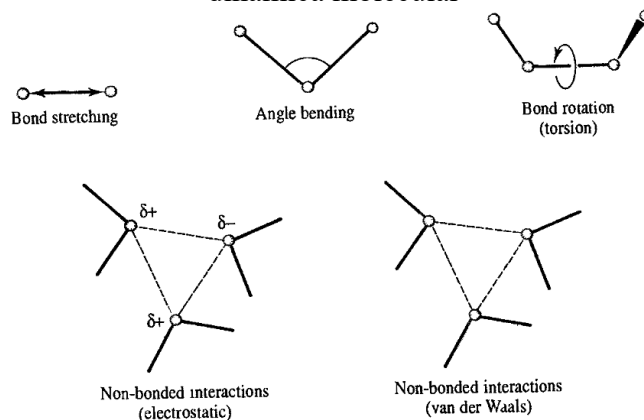
### 3.5 Campo de fuerza

El campo de fuerza es una aproximación empírica que permite simular el comportamiento de un sistema macromolecular, describiendo sus propiedades estructurales, las interacciones inter e intramoleculares y las energías del sistema mediante aproximaciones de la mecánica clásica y las ecuaciones de movimiento de Newton. Para ello, se calculan las fuerzas inter e intramoleculares y la energía potencial del sistema. En una ecuación de energía potencial, las interacciones inter e intramoleculares se describen como potenciales individuales mediante modelos de la mecánica clásica y funciones matemáticas que describen el movimiento de enlaces como resortes (i.e. La ley de Hooke) y las interacciones presentes como puntos cargados (i.e. mediante el potencial de Lennard-Jones y la ley de Coulomb). Esto se logra sumando las energías asociadas a todas las interacciones presentes en el sistema y determinando su estabilidad, dicha energía potencial se representa en las *ec. 1 y 2*. Los tipos de interacciones inter e intramoleculares presentes en estos sistemas se pueden observar en la Figura 11.

$$U = U_{inter} + U_{intra} \quad (\text{ec. 1})$$

$$U(r^n) = \sum_{stretching} \frac{K_i}{2} (l_i - l_{i,0})^2 + \sum_{angles} \frac{K_i}{2} (\theta_i - \theta_{i,0})^2 + \sum_{torsions} \frac{V_n}{2} (1 - \cos(n\omega - \gamma)) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left( 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \quad (ec. 2)$$

**Figura 11.** Esquema ilustrativo de las principales contribuciones de un campo de fuerza de dinámica molecular



Nota. Tomado de Leach, A. R. (2001). *Molecular Modelling: Principles and applications* (pp. 180). Prentice Hall(Leach, 2001).

En la *ec 2*, se observa la forma ampliada de la ecuación general que describe la energía potencial del sistema, donde se incluyen términos para el estiramiento de enlaces, cambios en los ángulos de enlace, torsiones e interacciones intermoleculares de tipo electrostáticas y de Van der Waals. Las interacciones enlazantes son representadas por la ley de Hooke donde los enlaces se consideran como resortes, esta aproximación permite describir los estiramientos, deformación de ángulos y torsiones de enlaces. Por otro lado, para el cálculo de las interacciones no enlazantes como las interacciones electrostáticas y a las fuerzas de dispersión de London, fuerzas encargadas de la atracción y repulsión entre átomos y moléculas, se usa la función analítica de Lennard-Jones con potencias de 12-6, la cual se observa en la *ec. 2*.

En el presente trabajo de grado se utiliza el campo de fuerza CHARMM36 (Chemistry at Harvard Macromolecular Mechanics) (Huang & MacKerell Jr, 2013), ampliamente destacado por su uso en el modelado de sistemas biomoleculares (i.e. proteínas, lípidos y ácidos nucleicos), ya que está parametrizado para el estudio específico de estos sistemas. La ecuación de energía potencial de CHARMM36 se presenta en la ec. 3.

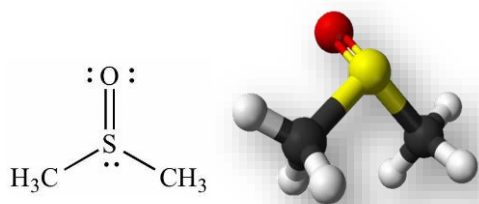
$$\begin{aligned}
 V(\vec{R}) = & \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi(1 + \cos(n\chi - \delta)) \\
 & + \sum_{UB} (S - S_0)^2 + \sum_{impropers} K_{imp}(\phi - \phi_0)^2 \\
 & + \sum_{non-bond} \epsilon \left[ \left( \frac{R_{minij}}{r_{ij}} \right)^{12} - \left( \frac{R_{minij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\epsilon_1 r_{ij}}
 \end{aligned} \tag{ec. 3}$$

Además de la ecuación de energía potencial del campo de fuerza, los paquetes de dinámica molecular implementan algoritmos matemáticos calcular los pasos de la simulación, la minimización de energía del sistema, para restringir enlaces de alta frecuencia, para descomponer y calcular las interacciones electrostáticas de larga y corta distancia, y para equilibrar la temperatura y presión del sistema. Para cada estudio se selecciona el algoritmo pertinente de acuerdo con el sistema y el rendimiento que se desee. Se describen con mayor detalle los algoritmos usados en el presente estudio en la sección del Apéndice.

### 3.6 Dimetil sulfóxido (DMSO)

El dimetil sulfóxido (DMSO) es un solvente orgánico de naturaleza anfipática, utilizado en síntesis orgánica y en aplicaciones médicas (Ramírez, 2014). Es un solvente aprótico, altamente polar debido a la presencia del grupo sulfóxido (-SO). El DMSO actúa como un excelente aceptor de puentes de hidrógeno debido al par de electrones libres del átomo de oxígeno, que facilita la formación de enlaces de hidrógeno con grupos donadores (como grupos -OH o -NH). A pesar de que no es buen donador, su elevada polaridad y naturaleza aprótica lo hacen especialmente

eficiente para estabilizar moléculas y disolver numerosas sustancias polares a través de interacción por puentes de hidrógeno, además, puede formar enlaces de hidrógeno en proporción de 1:2 (DMSO:2H<sub>2</sub>O) con el agua (Oh et al., 2017). Por otra parte, el DMSO tiene una simetría C<sub>s</sub> idealizada, con una geometría piramidal trigonal característica de compuestos S(IV) con tres enlaces coordinados. El átomo de azufre presenta un par de electrones no enlazantes y el agua dos pares, por lo que adopta una geometría aproximadamente tetraédrica



**Figura 12.** Estructura del DMSO

el DMSO puede atravesar fácilmente las membranas biológicas debido a su naturaleza polar y anfipática, lo que le permite difundir a través de la bicapa lipídica y actuar como método de transporte de otras moléculas transmembranalmente (Ramírez, 2014.). Un uso relevante es en técnicas como la reacción en cadena de la polimerasa (PCR, RT-PCR, RT-qPCR), ya que el DMSO afecta la estructura y estabilidad de los ácidos nucleicos ayudando a desnaturalizar sus estructuras secundarias y mejorar la eficiencia de lectura y la precisión de resultados de estas técnicas (Cadena-Caballero et al., 2022; Hardjasa et al., 2010; Tunçer et al., 2018; Wilcox et al., 2019). Las concentraciones típicas utilizadas para este propósito varían en el rango entre 5-10% v/v. Sin embargo, también se ha documentado que el DMSO también puede tener el efecto de estabilizar y prevenir la degradación de algunos motivos estructurales de ácido nucleicos como los G-quadruplexes (Lee et al., 2013; Tariq et al., 2022).

## 4. Metodología

### 4.1 Selección del fragmento de ARN

Se usó la secuencia nucleotídica de un fragmento de ARN de 30 nucleótidos perteneciente al gen humano DRD4 (Human Dopamine receptor D4, NM\_000797.4), el cual fue secuenciado experimentalmente por el Grupo de Investigación de Microbiología y Genética (GIMG), y sugerido por el profesor de la Escuela de Biología, Francisco M. Pérez (Francisco Jose Martinez Perez-Es, 2025.).

<i>Gen</i>	<i>Secuencia</i>
<i>Human Dopamine Receptor D4, (DRD4), NM_000797.4</i>	GCUCAUCGGCGCCACGUGGCUGCUGUCCGC

### 4.2 Determinación de la estructura secundaria del fragmento de ARN

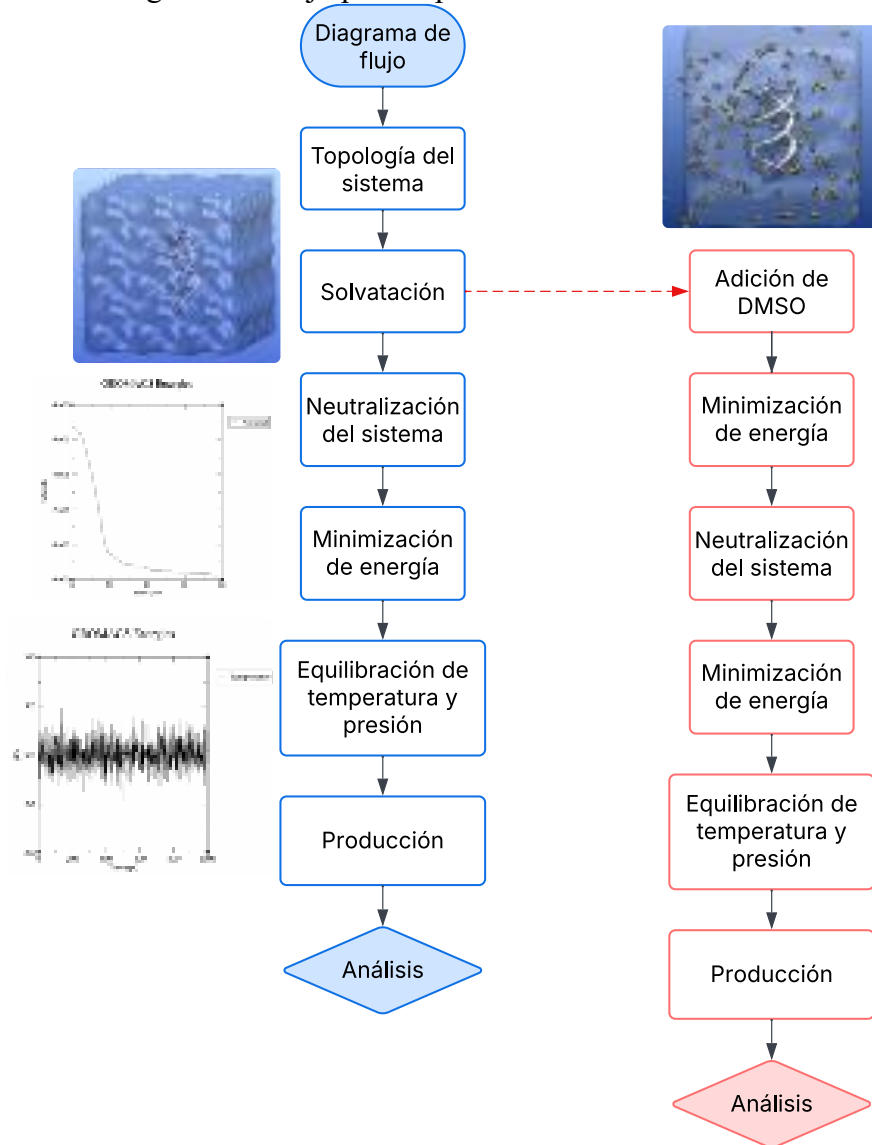
Se predijo la estructura secundaria y la estructura tridimensional del fragmento de ARN mediante el software FARFAR2 (Watkins et al., 2020). Para obtener la conformación más estable, se modelaron un total de 10000 estructuras con 100000 ciclos de monte Carlo. Posteriormente, la estructura secundaria del modelo 3D obtenida con FARFAR2 fue caracterizada usando las siguientes herramientas complementarias: rna-tools.online (Magnus, 2022), para obtener la estructura secundaria en formato dot-bracket a partir del archivo PDB; VARNA (Darty et al., 2009), para generar una representación 2D de la estructura secundaria a partir la notación dot-bracket; y DSSR (Lu et al., 2015), para un análisis detallado de la estructura secundaria del fragmento de ARN. Las representaciones estructurales 2D y 3D del fragmento de ARN se ilustran en las *Figuras 14 y 16*.

### 4.3 Dinámicas moleculares

Se realizaron seis simulaciones de dinámica molecular de 2 $\mu$ s cada una: tres réplicas del fragmento de ARN en fase acuosa y tres réplicas del fragmento de ARN en solución de DMSO al

5%v/v. A continuación, se ilustra el flujo de trabajo usado para la producción de las dinámicas moleculares en el esquema 1.

**Esquema 1.** Diagrama de flujo para la producción de las dinámicas moleculares



*Nota.* Se utilizó la semilla del llenado de moléculas de agua del sistema de ARN en fase acuosa para las simulaciones de ARN en presencia de DMSO, para tener la misma disposición inicial de moléculas de agua en las réplicas de ambos sistemas. Las moléculas de DMSO se adicionaron aleatoriamente y con relación al cálculo de la *sección 4.3.2*.

### 4.3.1 *Dinámicas de ARN en agua*

Para la producción de las dinámicas moleculares de ARN en agua, se corrieron tres réplicas, cada una de 2  $\mu$ s de duración, para cada réplica se siguió la ruta de trabajo ilustrada en el *esquema 1*, de esta forma se tuvieron las mismas configuraciones iniciales pero diferente semilla de distribución de moléculas del medio y velocidades iniciales. Todas las réplicas compartieron los siguientes parámetros: las simulaciones se generaron utilizando el software **GROMACS 2023-1 y 2024-5** (*GROMACS 2024.4 documentation*, 2024) con el campo de fuerza **CHARMM36-Jul 2022** (Huang & MacKerell Jr, 2013) y el modelo de agua **TIP3P** (Mark & Nilsson, 2001). Inicialmente, se minimizó la energía del sistema mediante el **Algoritmo de gradiente conjugado (GC)** (Nazareth, 2024), lo que permite optimizar la estructura inicial y eliminar cualquier solapamiento no deseado. A continuación, se colocó el fragmento de ARN en una caja cúbica, de manera que la distancia entre la superficie de la molécula y el límite de la caja fuese de 1.5 nm (Šponer et al., 2018). Posteriormente, se llenó la caja de la simulación con moléculas de agua. Para neutralizar la carga total del sistema, se añadieron contraiones de sodio, ajustando su concentración para igualar la carga neta del sistema. En el presente sistema, la carga es de -30, de acuerdo con el número de ribonucleótidos del fragmento de ARN seleccionado. Luego de preparar el sistema, este se sometió a un proceso de equilibración que consta de dos fases. Primero, se realizó una equilibración de 1 ns bajo el **conjunto NVT** a una temperatura de 300 K. Para controlar la temperatura, se utilizó el termostato modificado de Berendsen **reescalado de velocidad (V-rescale)** (Bussi et al., 2007) con un tiempo constante de  $\tau = 0.1$  ps. Una vez finalizada esta etapa, se realizó una segunda etapa de equilibración, también de 1 ns bajo el **conjunto NPT** a 1 bar de presión, empleando el acoplamiento de presión de **Parrinello-Rahman** (Parrinello & Rahman,

1981) de forma isotrópica, con un tiempo constante de 2.0 ps y una compresibilidad isotérmica de  $4,5 \cdot 10^{-5}$ . En ambos casos, se utilizó un paso de integración de 2 ps.

Así mismo, es importante señalar que algunos parámetros adicionales como las longitudes de enlace intermoleculares del agua se restringieron utilizando el algoritmo **LINCS** (Linear Constraint Solver) (Hess et al., 1997) con un orden de expansión de 4 para asegurar la estabilidad de la simulación. Además, para las interacciones electrostáticas de larga distancia se hizo uso del método **Particle Mesh Ewald (PME)** (Darden et al., 1993), configurado con un corte de espacio real de 10 Å y un espaciamiento de Fourier de 1.6 Å. Por otro lado, se adoptó el esquema de corte **Verlet** para la búsqueda de vecinos, utilizando un tipo de búsqueda por cuadrícula y un intervalo de actualización de la lista de vecinos de 100 pasos. Los cortes para las interacciones de Coulomb y de van der Waals fueron establecidos en 1.2 nm de acuerdo con las recomendaciones de GROMACS para trabajar con el campo de fuerza CHARMM, lo cual es adecuado para representar de manera eficiente las interacciones de corto alcance. En conclusión, estos parámetros en conjunto aseguran la estabilidad del sistema, garantizando que se modelen adecuadamente las interacciones de corto y largo alcance, así como el comportamiento termodinámico y estructural del ARN en un entorno acuoso.

#### **4.3.2 Dinámicas de ARN en solución de DMSO al 5% v/v**

Para la producción de las simulaciones del sistema de ARN con DMSO al 5% v/v, se generaron tres réplicas, cada una de 2  $\mu$ s de duración. Se usaron los mismos parámetros de las simulaciones de ARN en fase acuosa para mantener un modelado consistente del sistema, *sección 4.3.1*. La adición de DMSO requirió una parametrización adicional que se realizó utilizando Charmm Small Molecule Library (CSML), el **CGENFF** (Charmm General Force Field) y el script

cgenff\_charmm2gmx para generar los parámetros topológicos del DMSO. Esto aseguró un correcto modelado del DMSO y generar las simulaciones con adición de un cosolvente. Para la adición DMSO a una concentración de 5% v/v se realizó la siguiente relación:

Considerando un volumen de 100 mL solución 5% v/v de DMSO y asumiendo volúmenes aditivos dado que las densidades del DMSO  $\delta_{DMSO} = 1,004 \text{ g/cm}^3$  y el agua  $\delta_{H_2O} = 0,999 \text{ g/cm}^3$  son muy cercanas, se utiliza este valor de densidad del agua de acuerdo con el modelo TIP3P (Mark & Nilsson, 2001), entonces:

$$m_{DMSO} = 1,004 \frac{\text{g}}{\text{cm}^3} \cdot 5\text{mL} = 5,020\text{g} \quad m_{H_2O} = 0,999 \frac{\text{g}}{\text{cm}^3} \cdot 95\text{mL} = 94,905\text{g}$$

Donde se tiene,  $PM_{DMSO} = 78,13 \text{ g/mol}$  y  $PM_{H_2O} = 18,02 \text{ g/mol}$

$$n_{DMSO} = \frac{5,02\text{g}}{78,13 \frac{\text{g}}{\text{mol}}} = 0,064 \text{ mol} \rightarrow \text{moléculas} = 3,87 \times 10^{22}$$

$$n_{H_2O} = \frac{94,905\text{g}}{18,02 \frac{\text{g}}{\text{mol}}} = 5,268 \text{ mol} \rightarrow \text{moléculas} = 5,71 \times 10^{24}$$

Dónde,  $x_{H_2O}$  = moléculas de agua en el sistema y  $x_{DMSO}$  = moléculas de DMSO

$$x_{H_2O} \cdot 0,00678 = x_{DMSO}$$

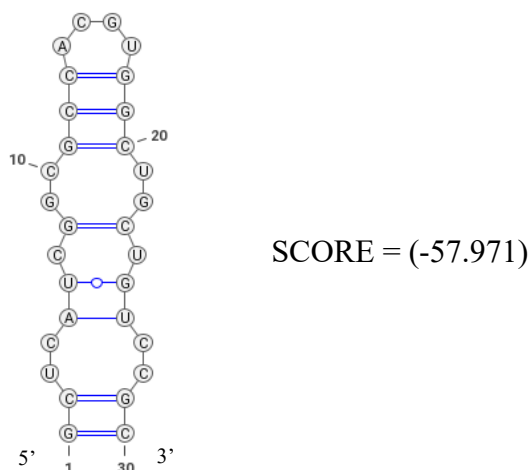
Posteriormente a la producción de las dinámicas moleculares, se realizó el análisis de resultados, para el cual se utilizaron diversos métodos de análisis para la comprensión del comportamiento del sistema y la elucidación de los mecanismos de interacción molecular. Los métodos de análisis empleados se discuten en la sección del *Apéndice*.

## 5. Resultados

### 5.1 Determinación y caracterización estructural del fragmento de ARN

A partir de la secuencia nucleotídica del fragmento de ARN de 30 nucleótidos del gen humano DRD4 se obtuvo la estructura inicial con el software FARFAR2 (Watkins et al., 2020). Se generaron más de 100,000 modelos tridimensionales mediante la modulación del número de ciclos de Montecarlo con el objetivo de obtener la mejor conformación. Una vez predicha y seleccionada la estructura más estable del fragmento de ARN, se caracterizaron los enlaces de hidrógeno entre bases nitrogenadas con rna-tools.online (Magnus, 2022), obteniendo la representación de la estructura secundaria en formato dot-bracket. Por último, se generó la representación 2D de la estructura secundaria con el software VARNA (Darty et al., 2009) para identificar los pares de bases y motivos de estructura secundaria del fragmento de ARN. En la figura 13 se muestra la representación 2D de la estructura secundaria más estable predicha.

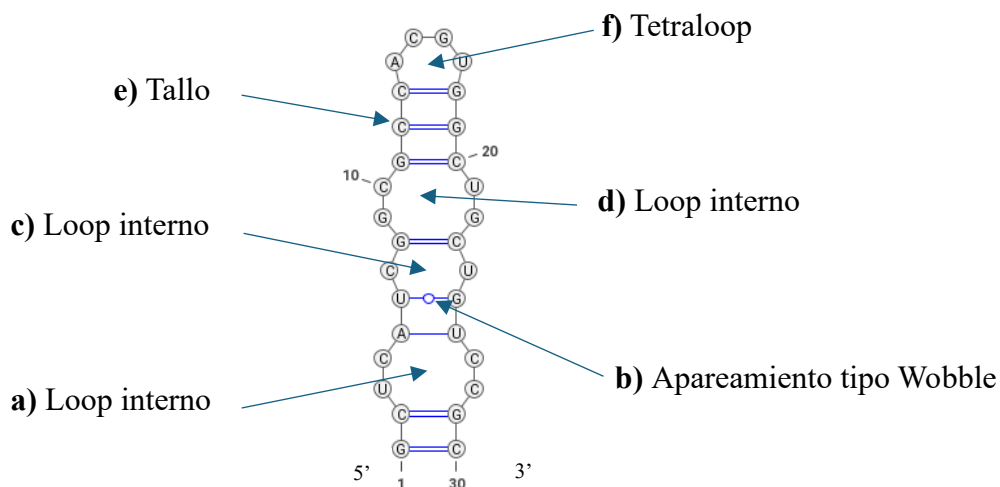
**Figura 13.** Estructura secundaria más estable del fragmento de ARN predicha con FARFAR2



La estructura secundaria seleccionada está definida por 8 pares de bases, con un total de 22 enlaces de hidrógeno, conformados por: 6 pares de bases G≡C (18 enlaces de hidrógeno), 1 par de bases A=U (2 enlaces de hidrógeno) y 1 par de bases G=U (2 enlaces de hidrógeno).

El análisis de la estructura secundaria predicha, *Figura 14*, con la herramienta DSSR reveló la presencia de múltiples motivos estructurales. La estructura se pliega simétricamente en forma de horquilla (hairpin) apareando sus extremos 5' y 3', y contiene los siguientes motivos estructurales: **a)** un primer loop interno (entre las bases 2-5 y 26-29), **b)** un apareamiento tipo Wobble (entre las bases U6=G25), **c)** un segundo loop interno (entre las bases 6-8 y las bases 23-25), **d)** un tercer loop interno (entre las bases 8-11 y las bases 20-23), **e)** un tallo (entre las bases 11-12 y las bases 13-20) y **f)** un tetraloop (entre las bases 14-17).

**Figura 14.** Motivos de estructura secundaria del fragmento de ARN de 30 nucleótidos del gen DRD4 humano.



Mediante DSSR (Lu et al., 2015) se evaluó la presencia de motivos de estructura terciaria, como variaciones en la orientación de los apareamientos de las bases nitrogenadas (i.e. apareamientos entre bases canónicas (Watson-Crick) y apareamientos no canónicos (Hoogsteen y sugar-edge) y se caracterizó la presencia de apilamiento entre bases nitrogenadas, *Tabla 1 a) y b)*.

**Tabla 1.** Caracterización de los pares de bases, tipos de apareamientos, motivos de estructura secundaria y apilamiento entre las bases nitrogenadas del fragmento de ARN

a)

Base pairs												
Pair ID	Pair	Chain1	ResID1	Base1	Base2	ResID2	Chain2	Name	Saenger	LW	DSSR	Hbond
1	G-C	R	1	GUA	CYT	30	R	WC	19-XIX	cWW	cW-W	3
2	C-G	R	2	CYT	GUA	29	R	WC	19-XIX	cWW	cW-W	3
3	U-C	R	3	URA	CYT	28	R	--	18-XVIII	cWW	cW-W	2
4	C-C	R	4	CYT	CYT	27	R	--	n/a	cWW	cW-W	2
5	A-U	R	5	ADE	URA	26	R	WC	20-XX	cWW	cW-W	2
6	U-G	R	6	URA	GUA	25	R	Wobble	28-XXVIII	cWW	cW-W	2
7	C-U	R	7	CYT	URA	24	R	--	18-XVIII	cWW	cW-W	2
8	G-C	R	8	GUA	CYT	23	R	WC	19-XIX	cWW	cW-W	3
9	G+G	R	9	GUA	GUA	22	R	--	06-VI	cHW	cM+W	2
10	C-U	R	10	CYT	URA	21	R	--	18-XVIII	cWW	cW-W	2
11	G-C	R	11	GUA	CYT	20	R	WC	19-XIX	cWW	cW-W	3
12	C-G	R	12	CYT	GUA	19	R	WC	19-XIX	cWW	cW-W	3
13	C-G	R	13	CYT	GUA	18	R	WC	19-XIX	cWW	cW-W	3

b)

Stack			
Serial	Num of nucleotides	Stacks	Nucleotides List
1	2	CU	R.CYT2,R.URA3
2	2	CA	R.CYT4,R.ADE5
3	2	CU	R.CYT20,R.URA21
4	4	UCGG	R.URA6,R.CYT7,R.GUA8,R.GUA9
5	4	UCCG	R.URA26,R.CYT27,R.CYT28,R.GUA29
6	5	ACGUG	R.ADE14,R.CYT15,R.GUA16,R.URA17,R.GUA18
7	6	GCGCUG	R.GUA11,R.CYT10,R.GUA22,R.CYT23,R.URA24,R.GUA25

*Nota.* **a)** Análisis del apareamiento entre bases nitrogenadas; número de pares de bases presentes, tipos de bases nitrogenadas apareadas, posición en la cadena, tipo y orientación del apareamiento entre las bases nitrogenadas, y el número de cada apareamiento entre bases. **b)** Análisis del apilamiento entre bases nitrogenadas; número de apilamiento, número y tipo de bases apiladas, y la posiciones en la cadena de cada apilamiento.

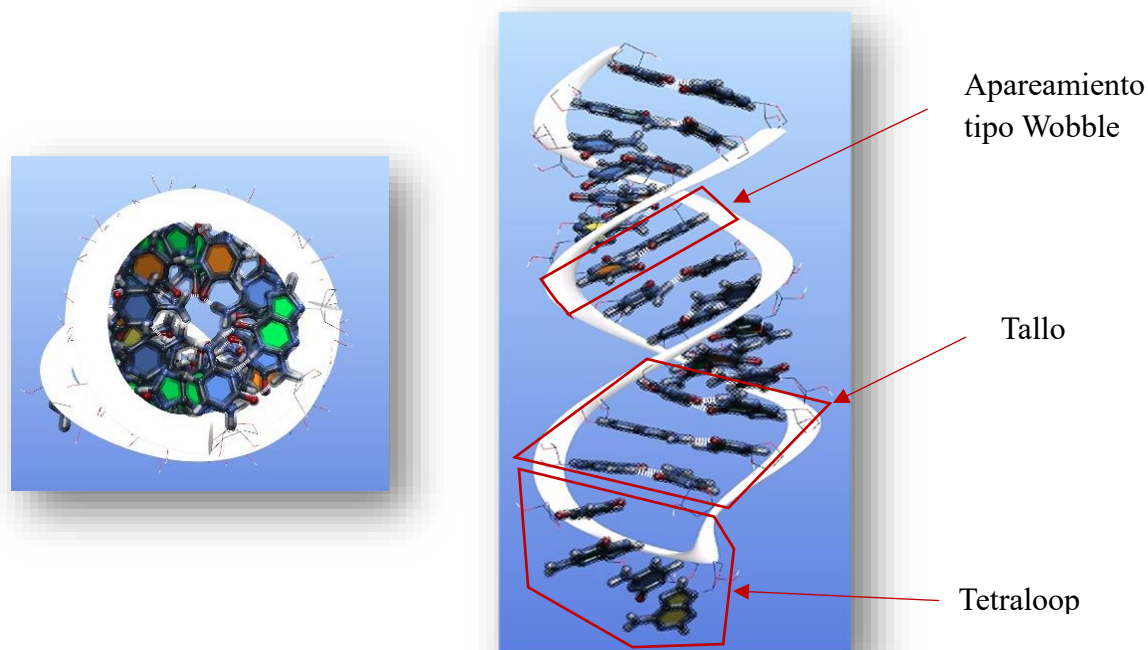
**Figura 15.** Representación 3D del fragmento de ARN para el análisis del apareamiento y apilamiento entre bases nitrogenadas



*Nota.* **a)** En rosado se señalan las zonas de la estructura en que se forman apareamientos entre bases nitrogenadas. **b)** En rosado se señalan las zonas de la estructura con apilamiento entre las bases nitrogenadas.

Mediante este análisis se identificó un único apareamiento de bases nitrogenadas no canónico, el apareamiento de bases tipo Wobble entre las bases (U6=G25). Los demás pares de bases se identificaron como apareamientos canónicos (i.e. apareamiento de tipos Watson-Crick). El análisis del apilamiento entre bases nitrogenadas mostró presencia de apilamiento de bases en el tetraloop (bases 14-17), en los bucles internos (bases 26-29), en las regiones helicoidales (i.e. tallos, las bases 11-13 y 18-20) y en otras regiones (bases 6-9).

Complementariamente, se elaboró una visualización 3D de la estructura con VMD (Humphrey et al., 1996), enfocada en facilitar la identificación de los residuos y átomos relevantes de la estructura, *Figura 16*.

**Figura 16.** Visualización 3D del fragmento de ARN

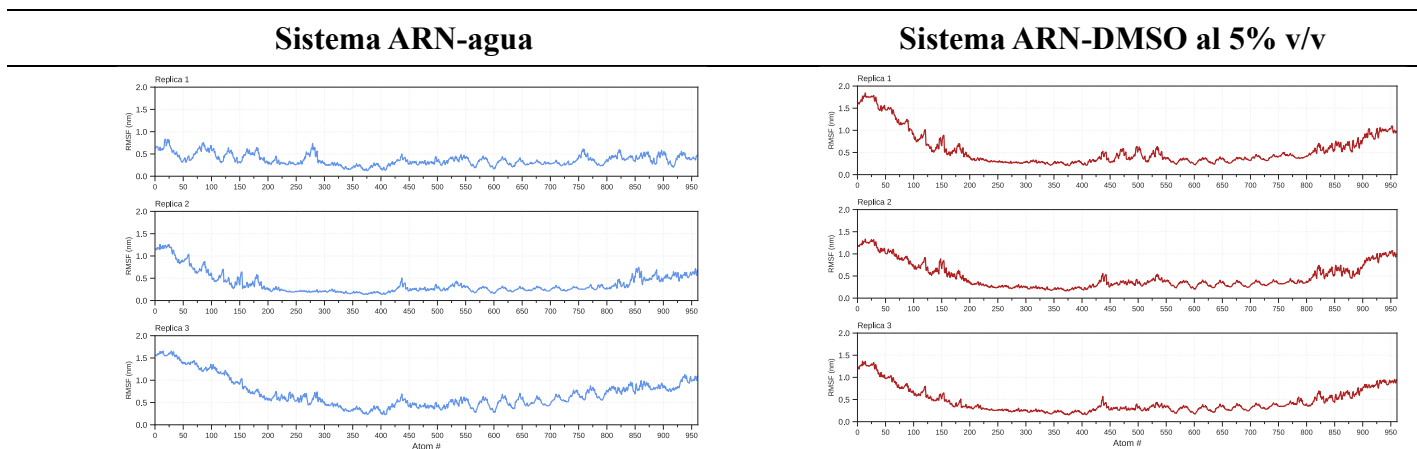
*Nota.* Se representan las bases nitrogenadas con colores específicos; Guanina (verde), Adenina (amarillo), Uracilo (naranja), Citosina (azul), los enlaces fosfodiéster del backbone como una cinta blanca y los azúcares ribosa del backbone con líneas. Se señalan algunos de los motivos de estructura secundaria como el tetraloop, el tallo y el apareamiento tipo Wobble.

## 5.2 Dinámica conformacional del fragmento de ARN en fase acuosa y en presencia de DMSO

Para evaluar la dinámica conformacional y estabilidad global, se calculó la fluctuación cuadrática media (RMSF) y la desviación cuadrática media (RMSD) del fragmento de ARN en fase acuosa y en presencia de DMSO, *Figuras 17 y 18*. El análisis de RMSF reveló que las regiones más flexibles del fragmento de ARN fueron los extremos terminales 5' y 3' y el tetraloop. Adicionalmente, el análisis de RMSD mostró una alta flexibilidad y una reiterada fluctuación conformacional en ambas condiciones simuladas, sin embargo, las simulaciones en presencia de DMSO se presentaron fluctuaciones conformacionales más recurrentes. Estos resultados sugieren

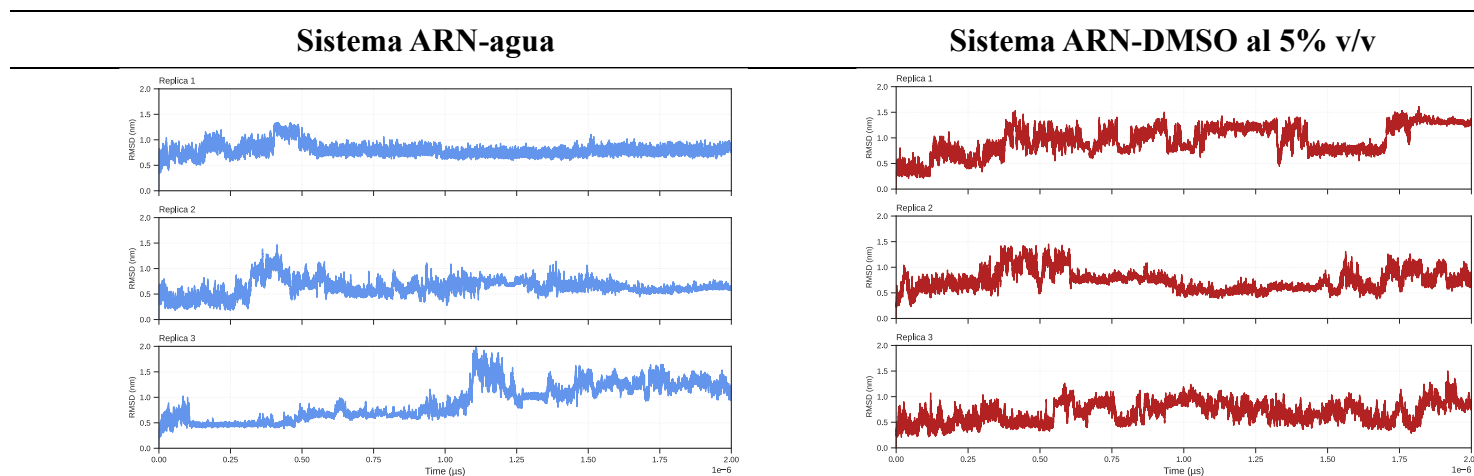
que el DMSO, aunque en la escala de tiempo simulada no induce un cambio estructural significativo, si incrementa ligeramente la fluctuación conformacional como se observa en las gráficas del RMSD.

**Figura 17.** Flexibilidad local del fragmento de ARN



Nota. Comparativa de las réplicas de la fluctuación cuadrática media (RMSF) entre los dos sistemas. Medición de la fluctuación promedio de la posición de cada átomo del fragmento de ARN a lo largo de los 2 $\mu$ s de simulación. Los picos elevados indican una alta flexibilidad local en estas regiones de la estructura. Las gráficas en azul corresponden al sistema de ARN en fase acuosa y en rojo al sistema de ARN con DMSO al 5% v/v.

Se observaron diferencias entre las réplicas debido a la dinámica conformacional y trayectoria particular del fragmento de ARN en cada simulación. A pesar de esto, se observó de forma consistente en todas las réplicas que los átomos pertenecientes a los extremos terminales 5' y 3' (nucleótidos 1-4 y 27-30) del fragmento de ARN evidenciaron mayor fluctuación cuadrática media (1.5 - 2.0 nm) en la totalidad de las simulaciones, debido a la ruptura de los dos primeros pares de bases de la estructura secundaria lo cual le otorga una alta flexibilidad a esta zona de la estructura. Además, El tetraloop (nucleótidos 14-17) presentó valores ligeramente altos de fluctuación cuadrática media (0.5 – 1.0 nm) ya que es un motivo estructural flexible.

**Figura 18.** Flexibilidad global del fragmento de ARN

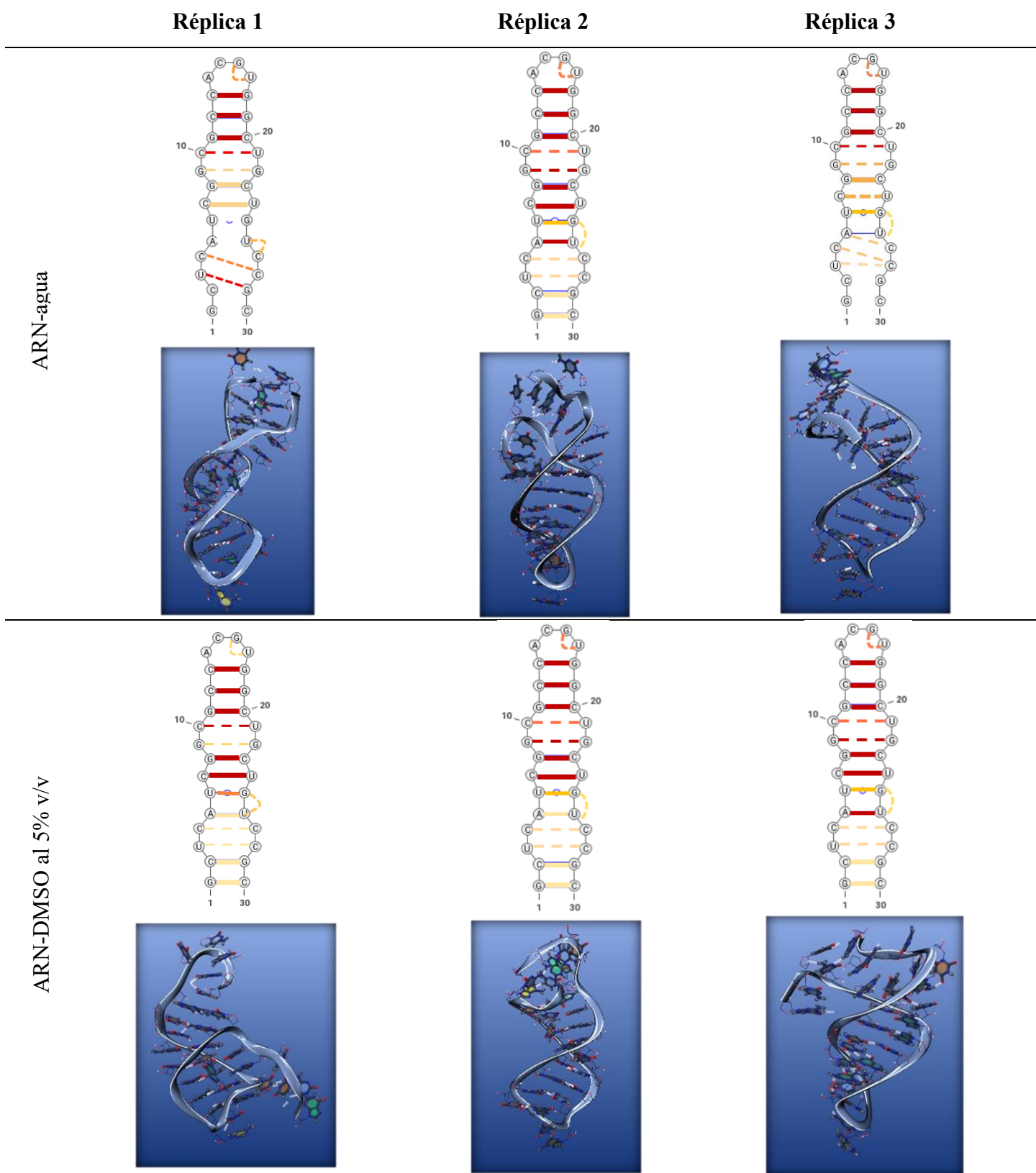
*Nota.* Comparativa de las réplicas de la desviación cuadrática media (RMSD) entre los dos sistemas. Medición de la desviación promedio de la conformación inicial del fragmento de ARN a lo largo de los 2  $\mu$ s de simulación. Los picos pronunciados indican transiciones conformacionales. Las gráficas en azul corresponden al sistema de ARN en fase acuosa y en rojo al sistema de ARN con DMSO al 5% v/v.

En la totalidad de las simulaciones el fragmento de ARN explora un amplio espacio conformacional (alcanzando un RMSD de hasta  $\sim 2$  nm). En las réplicas del sistema en fase acuosa los valores de RMSD fluctuaron, pero presentó periodos de estabilidad. En contraste, en las réplicas con presencia de DMSO se presentó mayor variabilidad de los valores de RMSD a lo largo de las trayectorias. Lo cual indica una mayor fluctuación conformacional y recurrente a lo largo de las trayectorias en comparación con el sistema en fase acuosa.

### ***5.2.1 Estabilidad de la estructura secundaria del ARN: Análisis de los enlaces de hidrógeno intramoleculares***

Se caracterizó la estabilidad de los enlaces de hidrógeno intramoleculares entre las bases nitrogenadas que definen y estabilizan la estructura secundaria del fragmento de ARN durante las dinámicas moleculares de ambos sistemas mediante un mapa de calor, *Figura 19 y A.2.*

**Figura 19.** Estabilidad de los enlaces intramoleculares del fragmento de ARN



Nota. En las estructuras 2D, las líneas en tonos cálidos oscuros representan los pares de bases más estables, y en tonos más claros se representan los pares de bases menos estables. Con línea sólida se representan pares de bases que se conservan desde la estructura inicial y en línea punteada los nuevos enlaces formados durante las simulaciones. Las estructuras 3D, representan la conformación estructural más dominante del cluster estructural de cada réplica obtenido mediante el plugin de VMD clustering, método que agrupa conformaciones similares con base en la desviación cuadrática media (RMSD).

En la totalidad de las réplicas se evidenció ruptura de los enlaces de hidrógeno entre los dos primeros pares bases de los extremos terminales 5' y 3', *Figura 19*. Se observó la formación de enlaces de hidrógeno transitorios en la región abierta de los extremos 5' y 3' (entre las bases 1-5 y 26-30), que se forman durante la exploración conformacional de la estructura en presencia del medio. En contraste, La región del tallo (bases 11-13 y 18 -20) fue la más estable de la estructura, con enlaces de hidrógeno que residen en más del 80% de la simulación, dicha estabilidad se atribuye a la presencia de tres pares de bases G≡C consecutivos cuyas bases están apiladas, por lo que es una zona con alta densidad de enlaces de hidrógeno e interacciones  $\pi$ - $\pi$  que contribuye a la estabilidad de la estructura secundaria.

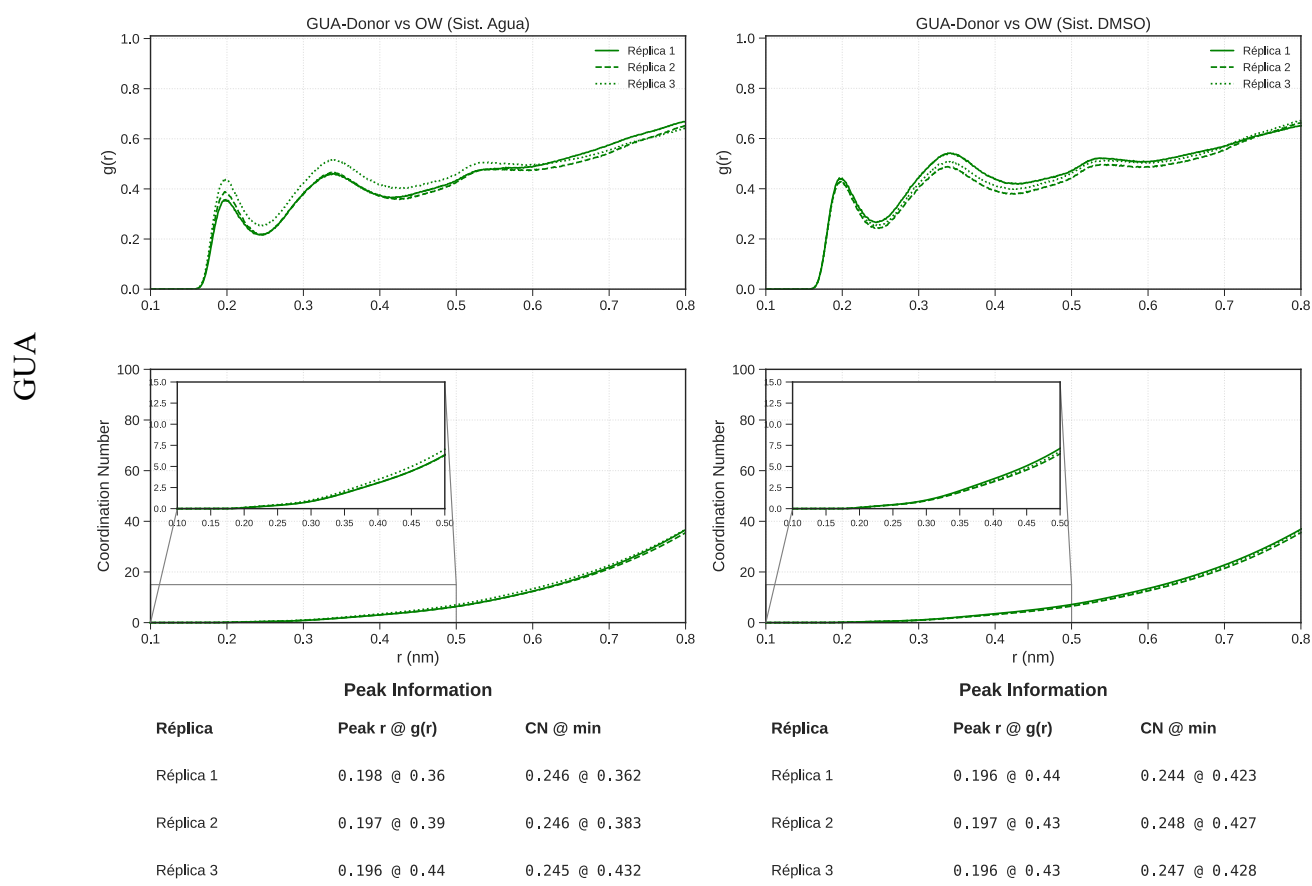
### **5.3 Caracterización de las interacciones intermoleculares no covalentes: ARN-agua y ARN-DMSO**

Se analizó la organización del agua y el DMSO alrededor del fragmento de ARN (con enfoque en los grupos de referencia de la *Figura A.3*, mediante funciones de distribución radial de las moléculas del medio (i.e. oxígeno del agua (OW), hidrógenos del agua (HW) y oxígeno del DMSO (O-DMSO)) en los sistemas ARN en fase acuosa y en presencia de DMSO al 5% v/v, *Figuras 20-22*. En ambos sistemas, el agua formó una red de hidratación bien definida y alrededor

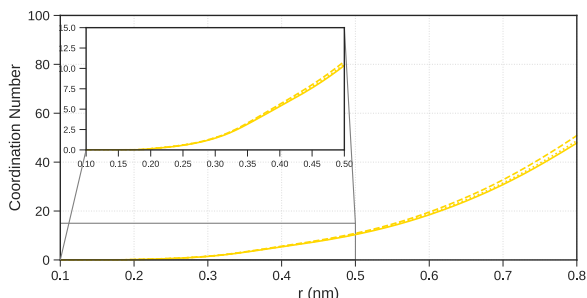
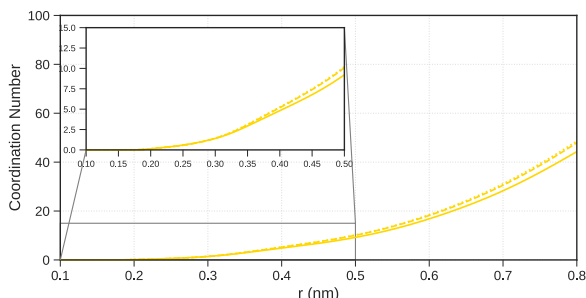
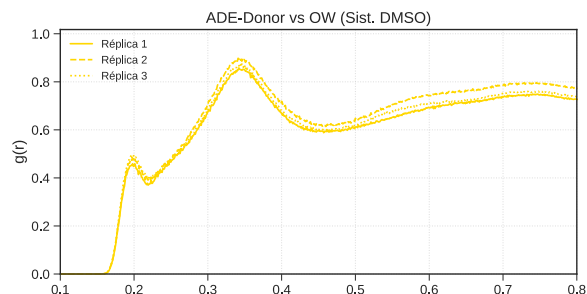
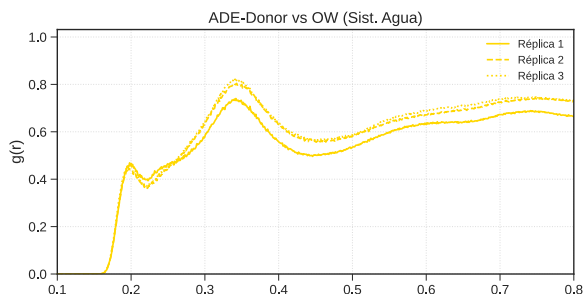
de los grupos polares de los residuos del fragmento de ARN, *Figura 20 y 21*. En contraste, la distribución del DMSO alrededor de los residuos del fragmento de ARN fue significativamente menos definida, con picos de baja intensidad, *Figura 22*. Esto indica que la superficie del fragmento de ARN interactúa preferencial con el agua debido a que se encuentra principalmente rodeada por una red de hidratación, y que el DMSO a pesar de interactuar ligeramente con el ARN se distribuye de manera difusa en el volumen del sistema.

### 5.3.1 Distribución radial de las moléculas de agua.

**Figura 20.** Distribución radial de las moléculas de agua (OW) alrededor de los grupos donadores de enlaces de hidrógeno de los residuos del fragmento de ARN



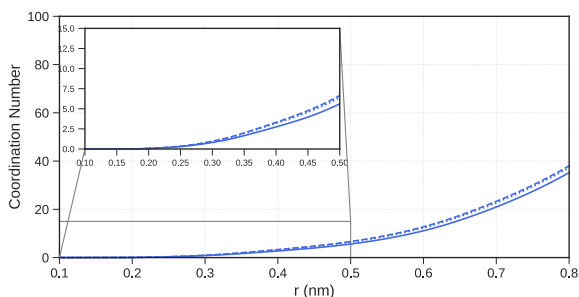
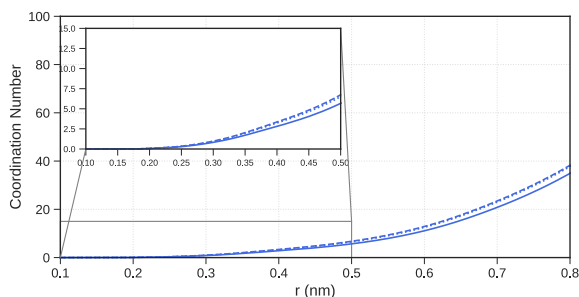
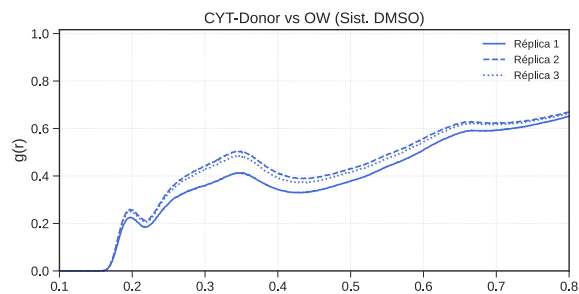
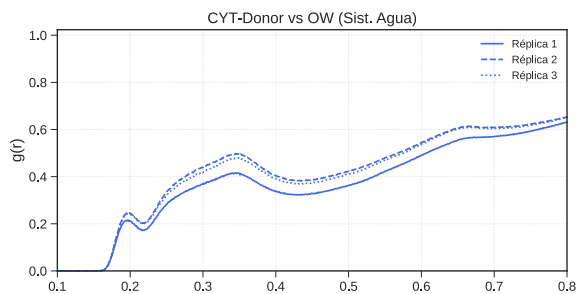
ADE



Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,198 @ 0,47	0,223 @ 0,316
Réplica 2	0,196 @ 0,45	0,222 @ 0,295
Réplica 3	0,200 @ 0,46	0,222 @ 0,306

Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,198 @ 0,46	0,220 @ 0,280
Réplica 2	0,198 @ 0,48	0,222 @ 0,310
Réplica 3	0,199 @ 0,50	0,222 @ 0,318

CYT

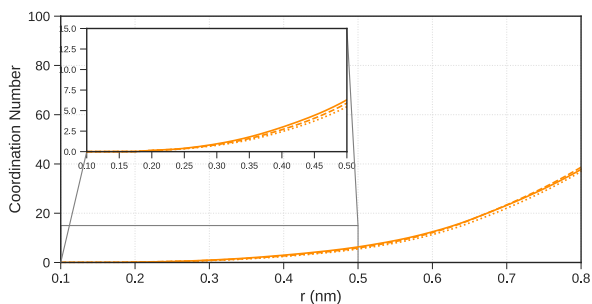
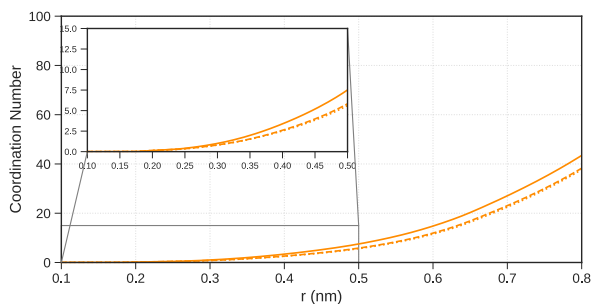
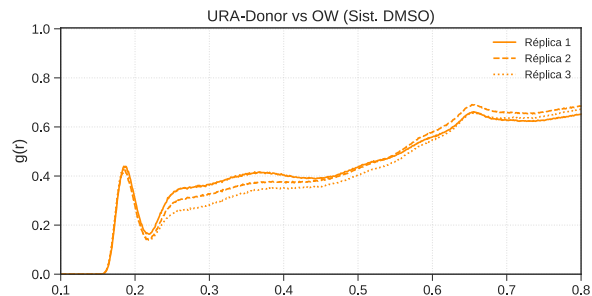
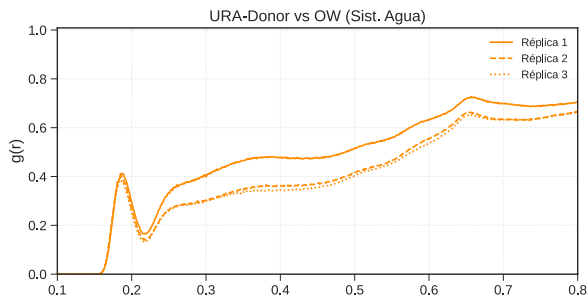


Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,196 @ 0,22	0,220 @ 0,137
Réplica 2	0,195 @ 0,25	0,220 @ 0,156
Réplica 3	0,198 @ 0,24	0,220 @ 0,154

Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,197 @ 0,23	0,220 @ 0,139
Réplica 2	0,197 @ 0,26	0,220 @ 0,159
Réplica 3	0,197 @ 0,25	0,220 @ 0,155

# ESTUDIO COMPUTACIONAL DE LAS INTERACCIONES ENTRE EL DMSO Y ARN

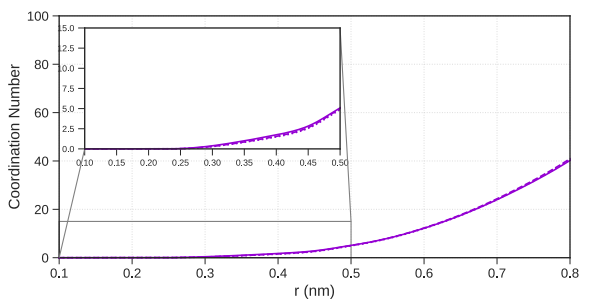
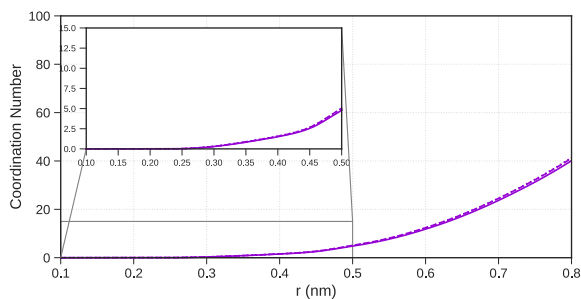
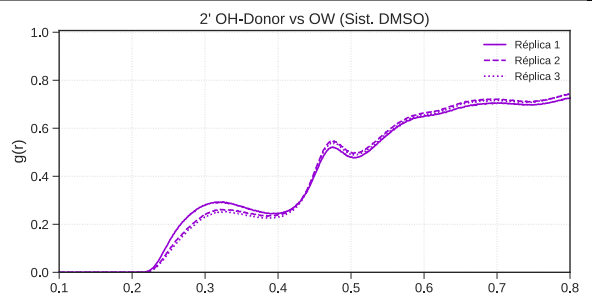
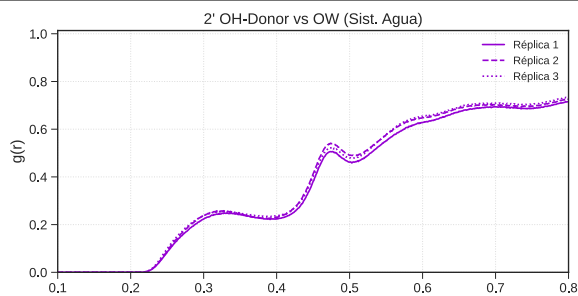
URA



Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0.188 @ 0.41	0.220 @ 0.226
Réplica 2	0.185 @ 0.40	0.220 @ 0.214
Réplica 3	0.185 @ 0.39	0.220 @ 0.204

Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0.185 @ 0.44	0.220 @ 0.230
Réplica 2	0.186 @ 0.42	0.220 @ 0.217
Réplica 3	0.185 @ 0.43	0.220 @ 0.220

2' OH



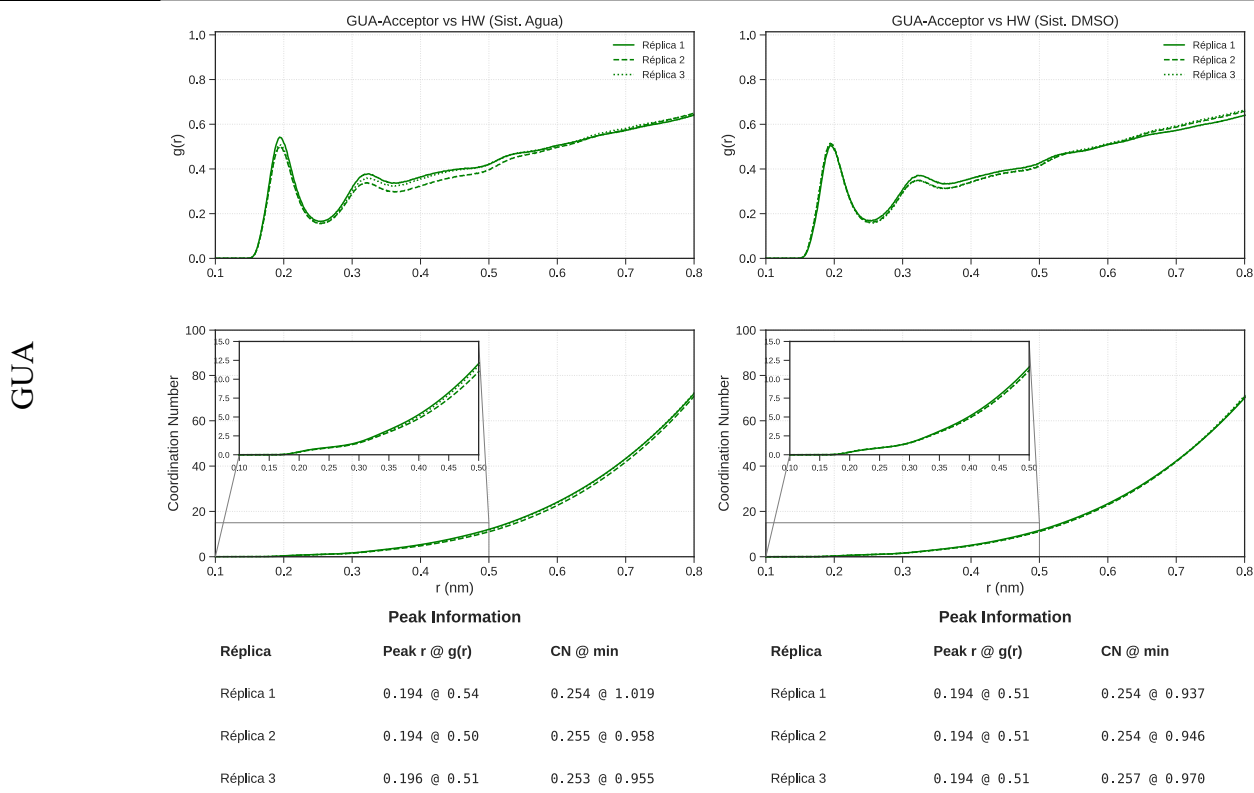
Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0.332 @ 0.25	0.220 @ 0.000
Réplica 2	0.328 @ 0.26	0.220 @ 0.000
Réplica 3	0.327 @ 0.26	0.220 @ 0.000

Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0.324 @ 0.29	0.220 @ 0.000
Réplica 2	0.321 @ 0.26	0.220 @ 0.000
Réplica 3	0.417 @ 0.25	0.220 @ 0.000

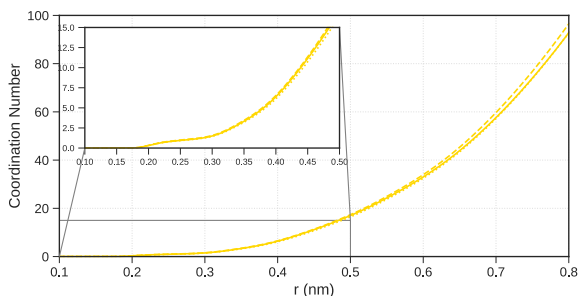
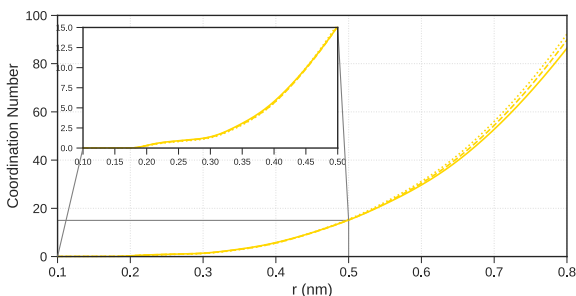
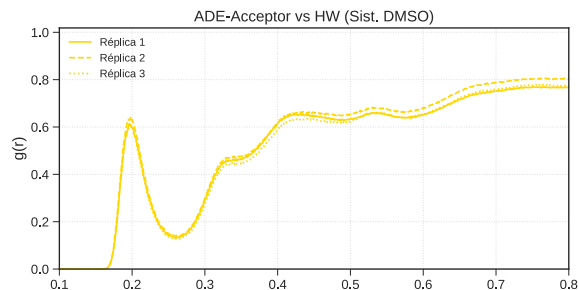
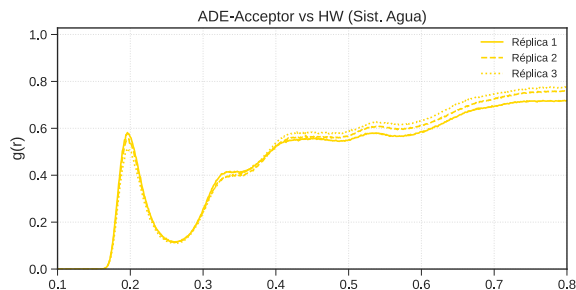
Nota. Comparativa entre las réplicas de los sistemas. A la izquierda el sistema de ARN-agua y a la derecha el sistema de ARN-DMSO al 5% v/v. Al pie de cada gráfica se observa el  $g(r)$  del primer pico y el valor correspondiente de moléculas coordinadas (CN). Se seleccionaron como grupos de referencia a los átomos donadores de enlaces de hidrógeno de los residuos del fragmento de ARN, *Figura A.3*.

Las guaninas, adeninas y uracilos presentaron una mayor distribución de moléculas de agua (orientadas por el oxígeno OW) a distancia entre 1.8-1.9 Å, con un primer pico bien definido, indicativo de una distribución consistente de la primera capa de moléculas de agua. El grupo 2'OH del backbone evidenció un primer pico menos intenso y ancho de 1.7 hasta 4 Å, lo cual indica mayor difusión, ya que estos grupos están más expuestos al solvente y tienen más aguas coordinadas y una red de moléculas de agua más amplia.

**Figura 21.** Distribución radial de las moléculas de agua (HW) alrededor los grupos aceptores de enlaces de hidrógeno de los residuos del fragmento de ARN



ADE



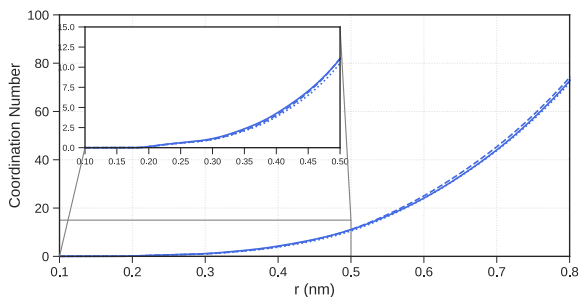
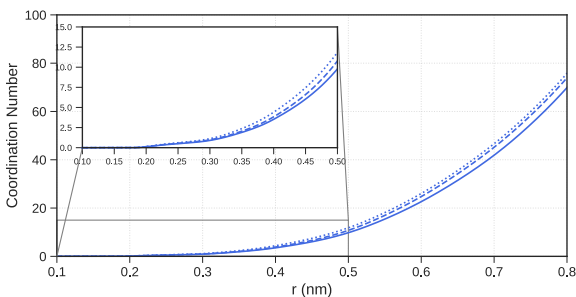
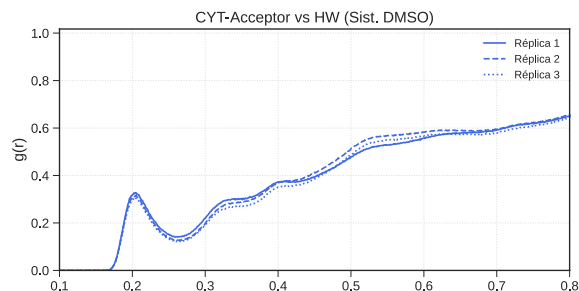
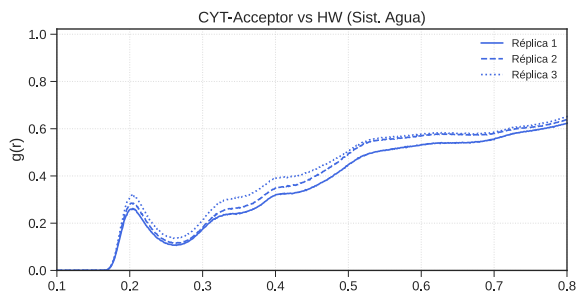
Peak Information

Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,196 @ 0,58	0,260 @ 0,967
Réplica 2	0,198 @ 0,56	0,262 @ 0,949
Réplica 3	0,195 @ 0,51	0,259 @ 0,863

Peak Information

Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,196 @ 0,61	0,259 @ 1,012
Réplica 2	0,198 @ 0,64	0,262 @ 1,082
Réplica 3	0,196 @ 0,61	0,262 @ 1,013

CYT



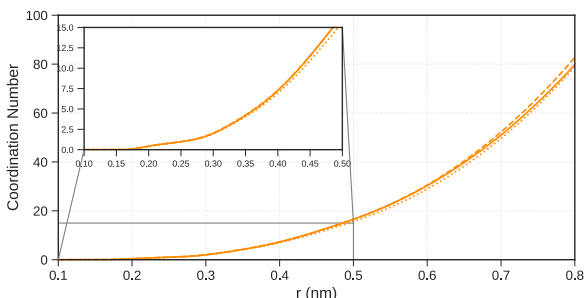
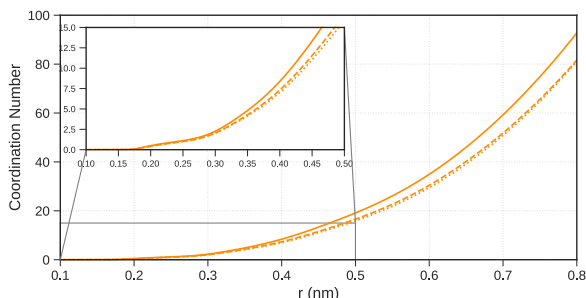
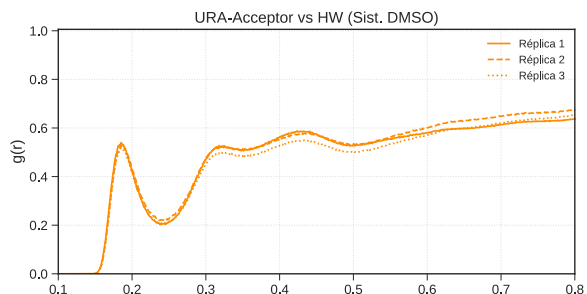
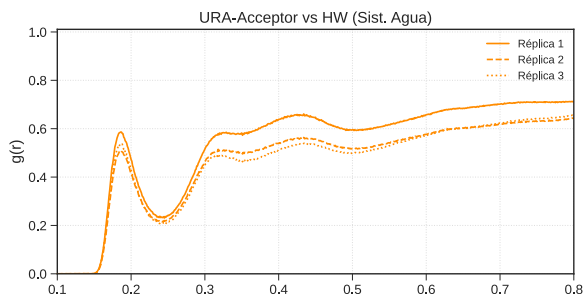
Peak Information

Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,204 @ 0,26	0,264 @ 0,583
Réplica 2	0,203 @ 0,28	0,263 @ 0,629
Réplica 3	0,203 @ 0,32	0,264 @ 0,722

Peak Information

Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,205 @ 0,33	0,259 @ 0,685
Réplica 2	0,205 @ 0,32	0,261 @ 0,661
Réplica 3	0,205 @ 0,30	0,264 @ 0,656

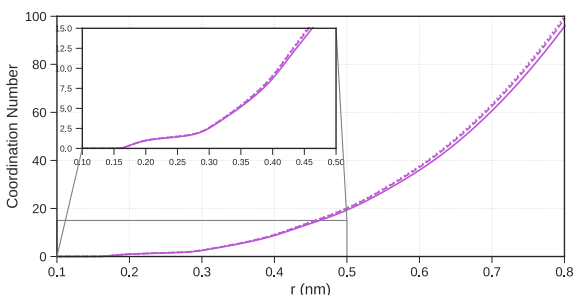
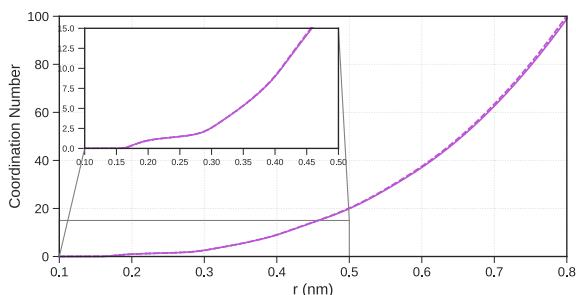
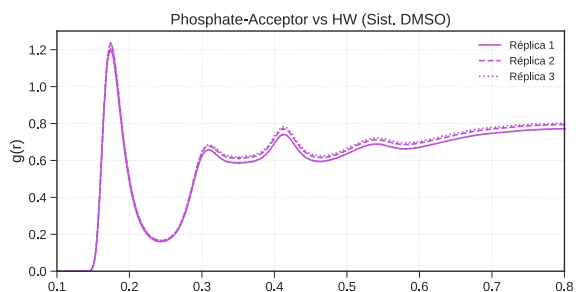
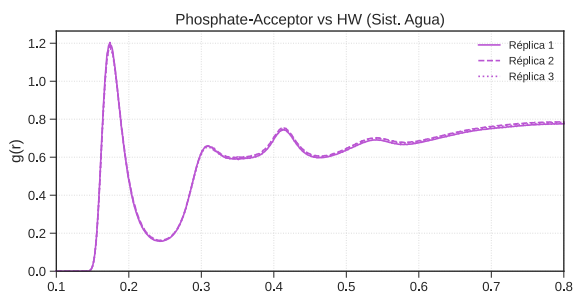
URA



Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,186 @ 0,59	0,251 @ 1,123
Réplica 2	0,186 @ 0,51	0,250 @ 0,989
Réplica 3	0,185 @ 0,54	0,251 @ 1,017

Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,185 @ 0,54	0,252 @ 0,993
Réplica 2	0,187 @ 0,53	0,250 @ 0,993
Réplica 3	0,185 @ 0,52	0,250 @ 0,952

Fosfato (PO<sub>4</sub>)



Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,174 @ 1,20	0,250 @ 1,473
Réplica 2	0,174 @ 1,19	0,250 @ 1,483
Réplica 3	0,174 @ 1,19	0,250 @ 1,467

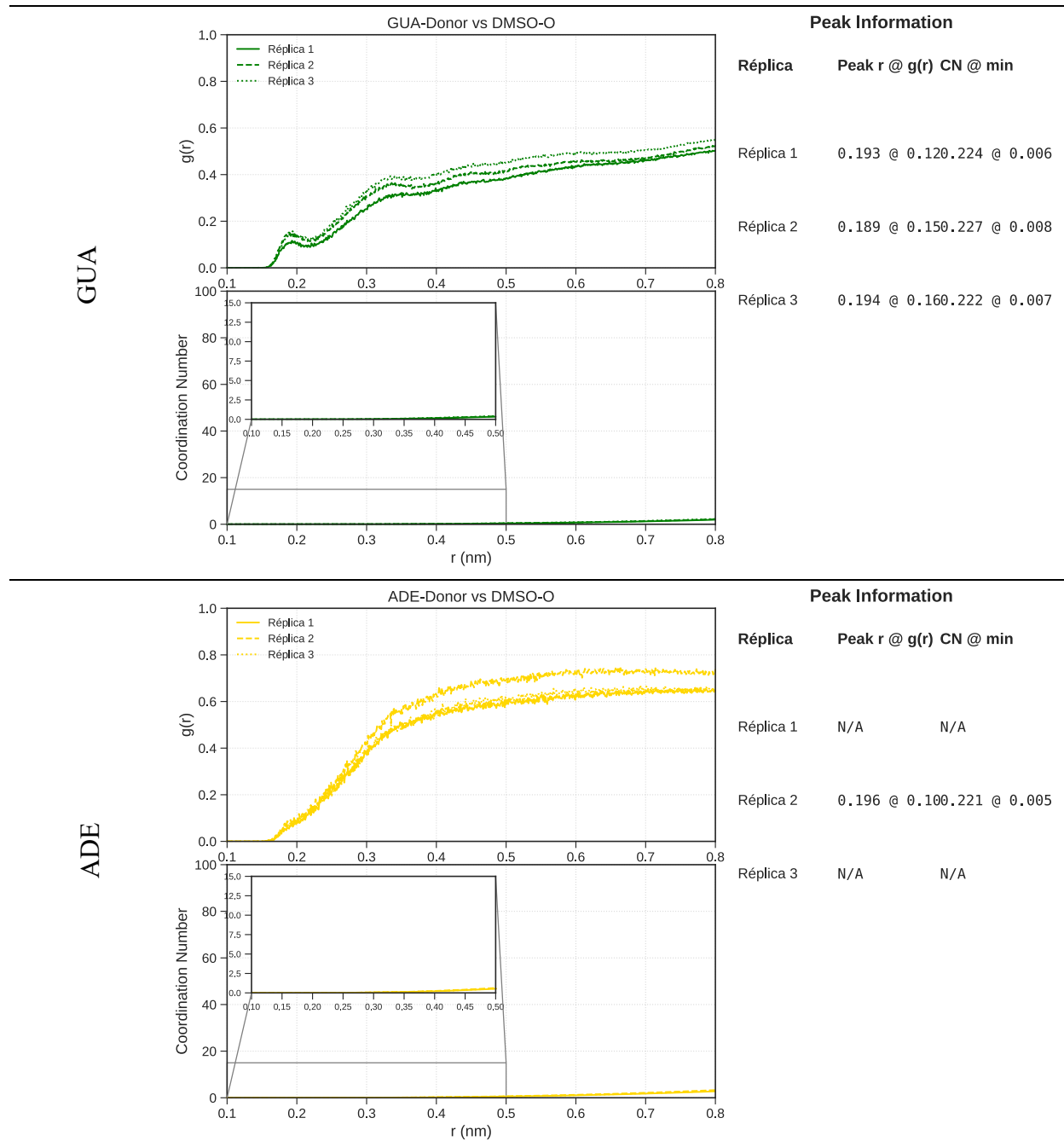
Peak Information		
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	0,174 @ 1,20	0,250 @ 1,442
Réplica 2	0,174 @ 1,23	0,250 @ 1,484
Réplica 3	0,174 @ 1,24	0,250 @ 1,495

*Nota.* Comparativa entre las réplicas de los sistemas. A la izquierda el sistema de ARN-agua y a la derecha el sistema de ARN-DMSO al 5% v/v. Al pie de cada gráfica se observa el  $g(r)$  del primer pico y el valor correspondiente de moléculas coordinadas (CN). Se seleccionaron como grupos de referencia a los átomos aceptores de enlaces de hidrógeno de los residuos del fragmento de ARN, *Figura A.3.*

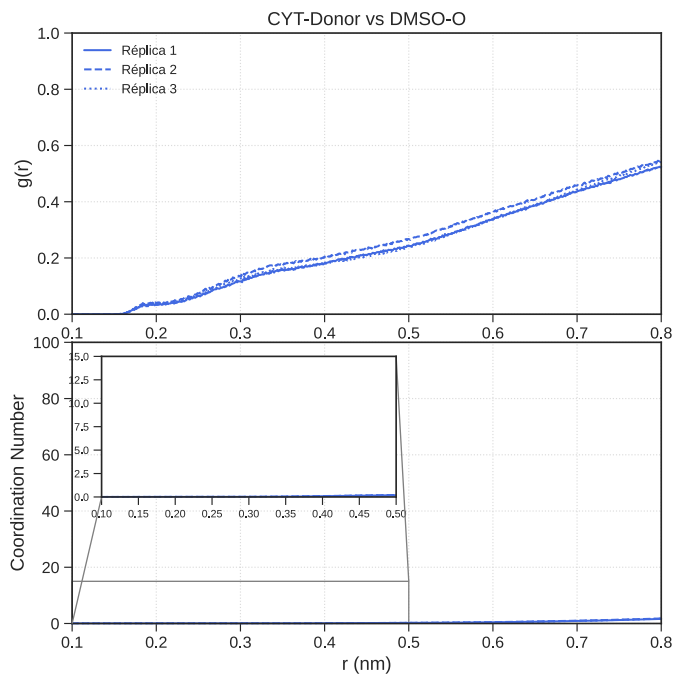
Las guaninas, adeninas y uracilos presentaron una mayor distribución de moléculas de agua (orientadas por los hidrógenos HW) a distancia entre 1.7-1.9 Å, con un primer pico bien definido, indicativo de una distribución consistente de la primera capa de moléculas de agua. Resalta la intensidad de la distribución del agua alrededor de los grupos fosfato ( $\text{PO}_4$ ) con un  $g(r) > 1$  a 1.74 Å, esto se ocurre dado a que cada grupo fosfato tienen cuatro oxígenos aceptores de enlaces de hidrógeno y cada agua dos HW, por lo que la probabilidad de encontrar moléculas de agua orientadas por cualquiera de sus dos hidrógenos es previsiblemente alta alrededor del backbone.

5.3.2 Distribución radial de las moléculas de DMSO

**Figura 22.** Distribución radial de las moléculas de DMSO (O-DMSO) alrededor los grupos donadores de enlaces de hidrógeno de los residuos del fragmento de ARN



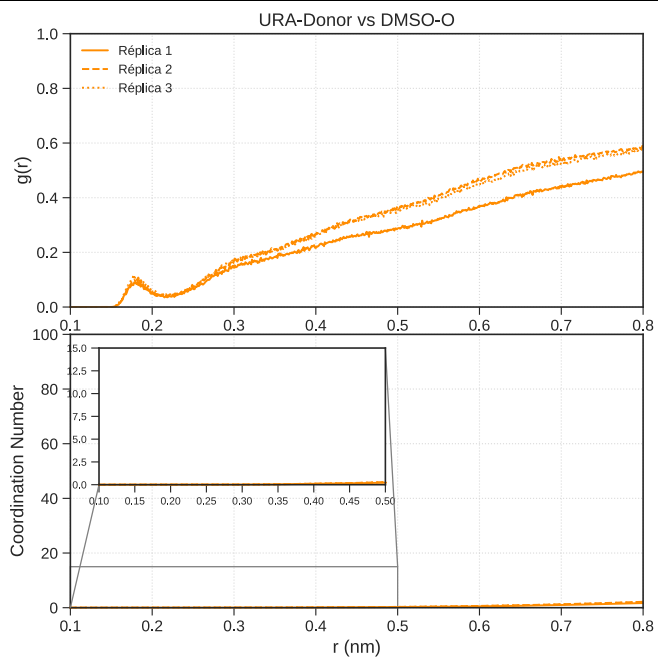
CYT



**Peak Information**

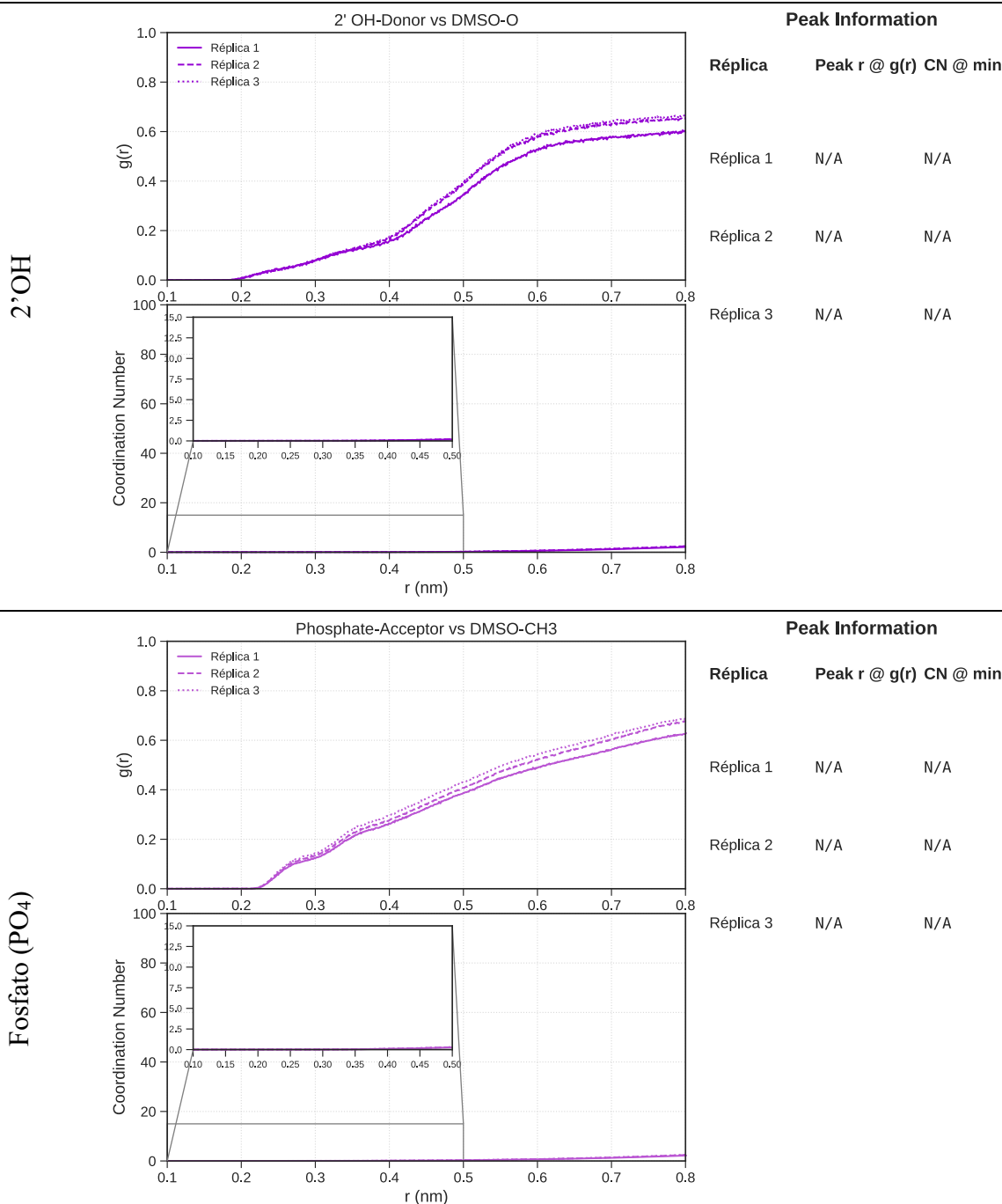
Réplica	Peak r @ g(r)	CN @ min
Réplica 1	N/A	N/A
Réplica 2	N/A	N/A
Réplica 3	N/A	N/A

URA



**Peak Information**

Réplica	Peak r @ g(r)	CN @ min
Réplica 1	N/A	N/A
Réplica 2	0.183 @ 0.110.224 @ 0.004	
Réplica 3	0.177 @ 0.120.225 @ 0.005	



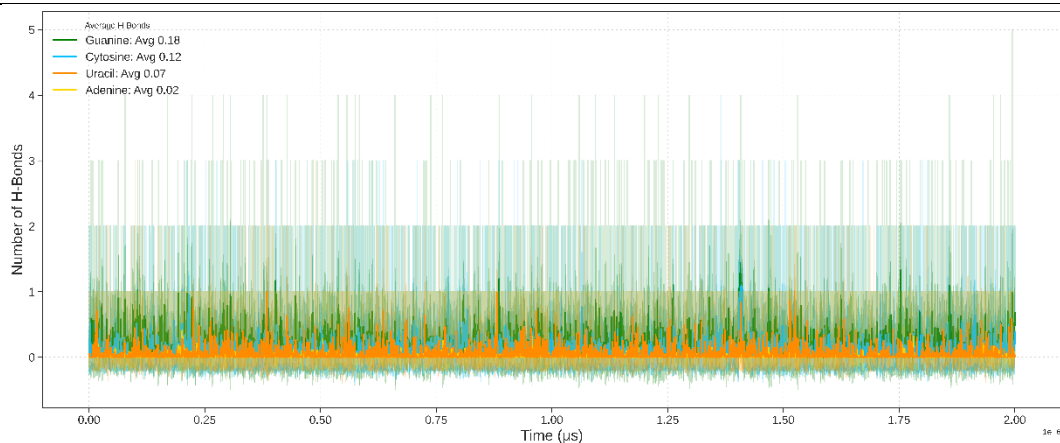
Se observó que la distribución de moléculas de DMSO orientadas por su oxígeno (O-DMSO) alrededor del ARN fue muy baja, con picos poco intensos. A pesar de esto, se observó ligeramente una mayor distribución de moléculas de DMSO alrededor de las guaninas y uracilos a distancias de 1.67 y 1.75 Å, indicando ligeramente mayor interacción del DMSO con las guaninas y uracilos, y una distribución difusa del DMSO en el volumen del sistema.

### ***5.3.3 Enlaces de hidrógeno intermoleculares ARN-DMSO***

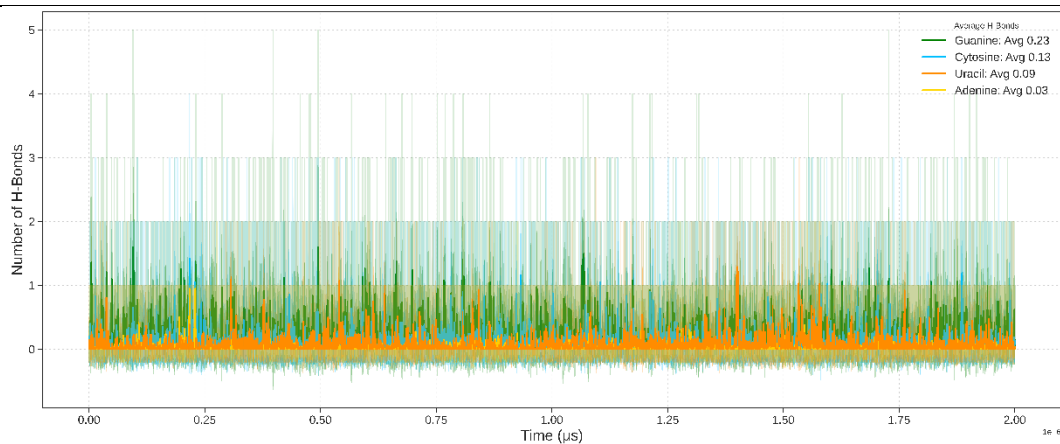
Mediante el análisis del número de enlaces de hidrógeno, distribución de las distancias y las ocupancias de los enlaces de hidrógeno ARN-DMSO, *Figuras 23, 24 y 25*. Se identificó que el número de enlaces ARN-DMSO en las trayectorias es muy bajo y con tiempos de residencia muy cortos. Esto indica que no existe preferencia del DMSO por interactuar con el ARN, que el DMSO se difunde homogéneamente en el medio y que las interacciones observadas en la escala de tiempo simuladas son transitorias.

**Figura 23.** Conteo de enlaces de hidrógeno ARN-DMSO**Réplica****Conteo de enlaces de hidrógeno ARN-DMSO**

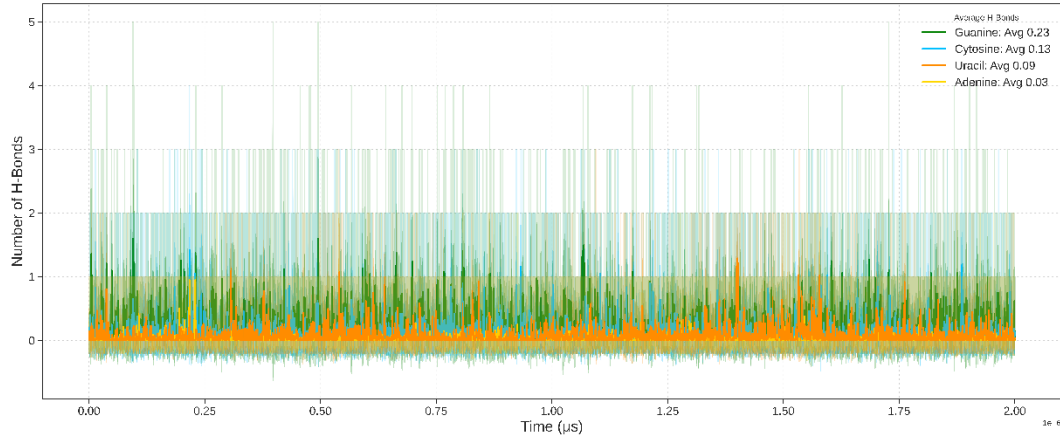
1



2



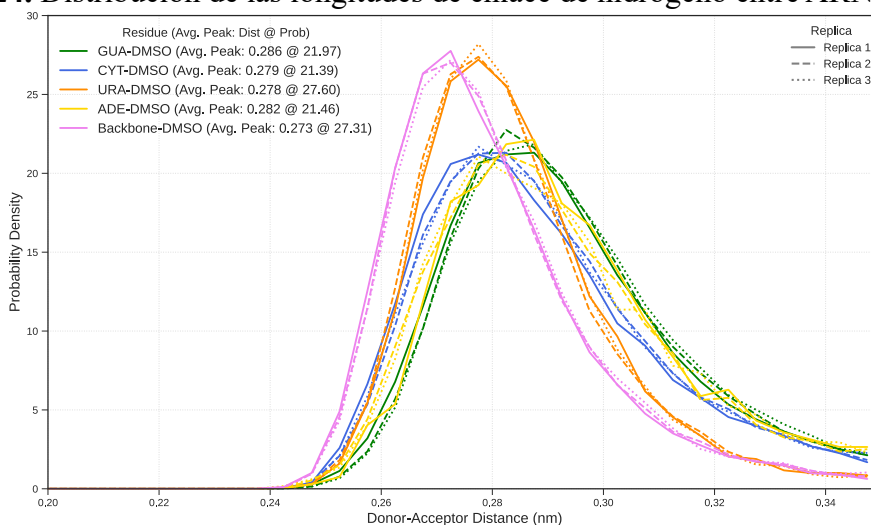
3



Nota. En verde, amarillo, azul y naranja se representa el conteo de enlaces de hidrógeno entre el DMSO y guanina, adenina, citosina y uracilo. El valor promedio de enlaces de hidrógeno nucleótido-DMSO es  $< 1$  enlace por fotograma de simulación.

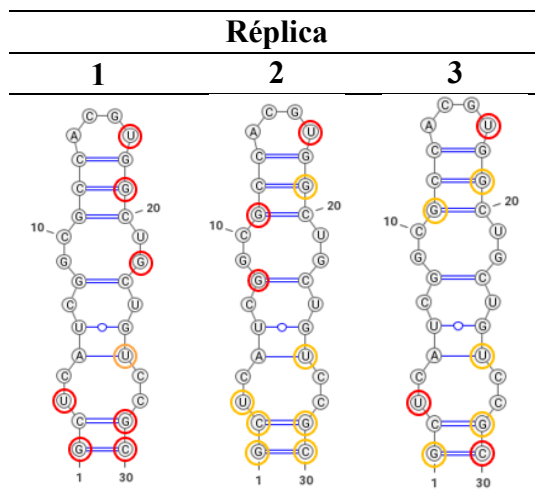
La distribución de la longitud de enlace de los enlaces de hidrógeno entre el DMSO con los residuos del fragmento de ARN estuvo en el rango de 2.7 a 2.9 Å, distancias en el orden de los enlaces de hidrógeno débiles. El uracilo y el backbone fueron los residuos con mayor distribución de enlaces de hidrógeno a distancias de 2.8 y 2.7 Å, *Figura 24*

**Figura 24.** Distribución de las longitudes de enlace de hidrógeno entre ARN y DMSO.



Nota. Distribución de los enlaces de hidrógeno de la guanina, adenina, citosina, uracilo y el backbone en verde, amarillo, azul, naranja y rosado respectivamente.

**Figura 25.** Nucleótidos con mayores tiempos de residencia de enlace de hidrógeno con el DMSO



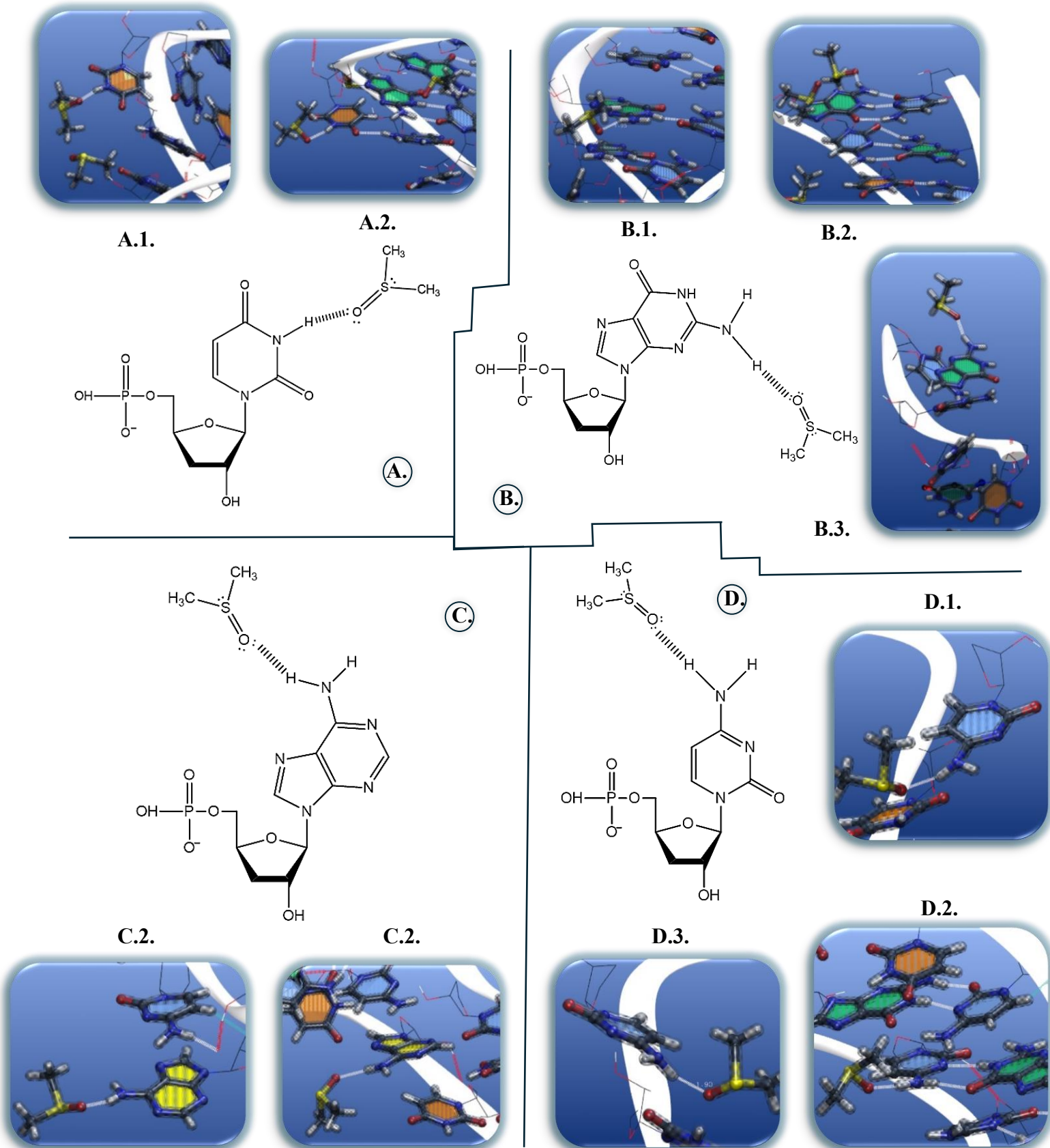
Nota. En rojo se señalan las bases que tuvieron ocupancias de enlaces de hidrógeno > 1% con el DMSO y en naranja se señalan las bases que tuvieron ocupancias entre  $0.6\% \leq 1\%$ . La ocupancia

se refiere al porcentaje de tiempo de la trayectoria en que se forman enlaces de hidrógenos nucleótido-DMSO, *Figura A.5*.

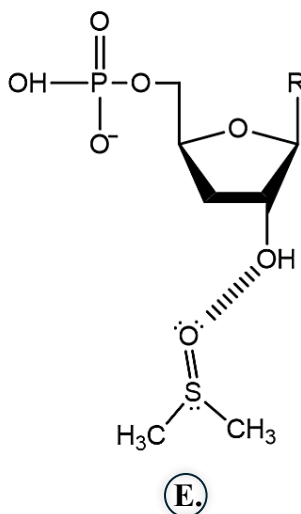
Mediante el análisis de ocupancias se identificó que los nucleótidos con mayor interacción con el DMSO se localizan consistentemente en las zonas de alta flexibilidad estructural, en los bucles y extremos terminales 5' y 3'. La ocupancia de los enlaces de hidrógeno ARN-DMSO tuvieron tiempos de residencia muy cortos, presentes en menos del 2% de la trayectoria, lo cual que los enlaces ARN-DMSO fueron. Se identificó, que consistentemente los nucleótidos con mayor frecuencia de formación de enlaces de hidrógeno con el DMSO fueron las guaninas y uracilos localizados en los bucles y los extremos abiertos 5' y 3', *Figuras 25 y A.5*. Sin embargo, no se observó ninguna interacción estable o acoplamiento del DMSO con las bases nitrogenadas, ni que el DMSO compitiera directamente por los enlaces de hidrógeno intramoleculares de las bases nitrogenadas apareadas. Por lo que únicamente se evidenció la interacción del DMSO con las zonas expuestas y flexibles de la estructura, pero no se observó que el DMSO promoviera la desestabilización directa de la estructura secundaria del fragmento de ARN.

A continuación, en el *Esquema 2* ilustra las geometrías de interacción más representativas entre el oxígeno aceptor del DMSO y los grupos donadores de enlaces de hidrógeno de cada tipo de nucleótido, principalmente los grupos amino (-NH<sub>2</sub>) y amida (-NH) de las bases nitrogenadas y el grupo hidroxilo (-2'OH) de la ribosa. Adicionalmente, se caracterizaron interacciones menos frecuentes pero relevantes, como aquellas que involucran al átomo de azufre del DMSO y la formación de puentes de hidrógeno en proporción 2:1 (DMSO:base).

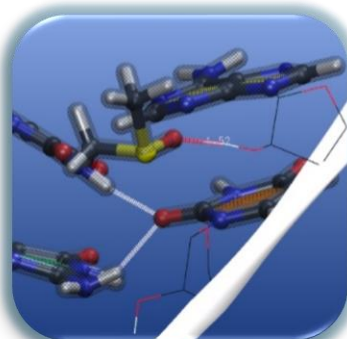
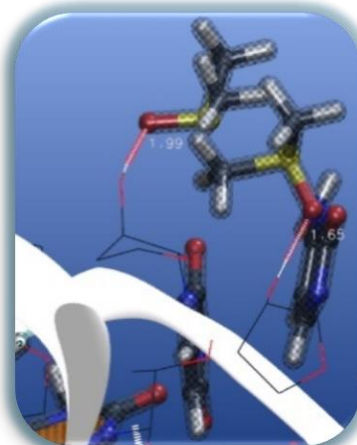
**Esquema 2.** Visualización de los enlaces de hidrógeno formados entre el DMSO y el fragmento de ARN



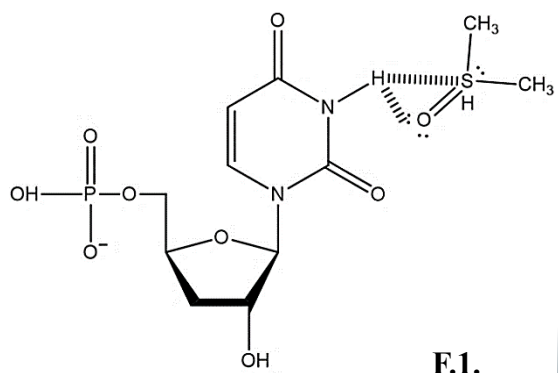
**E.1.**



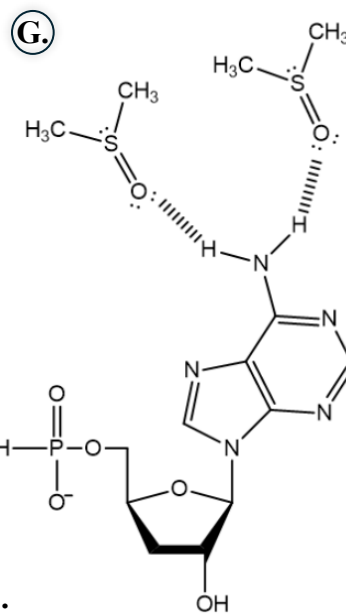
**E.3.**



**E.2.**

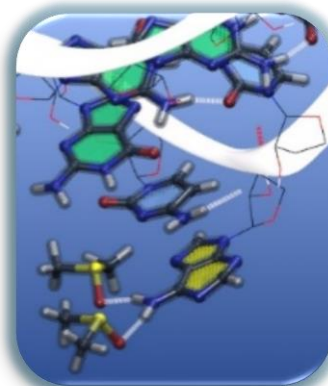


**F.1.**

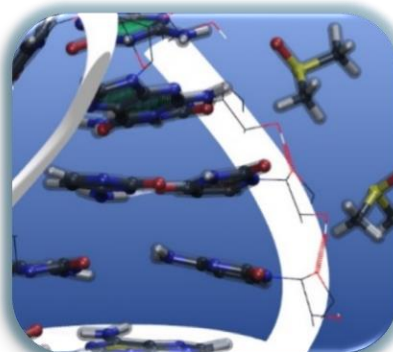
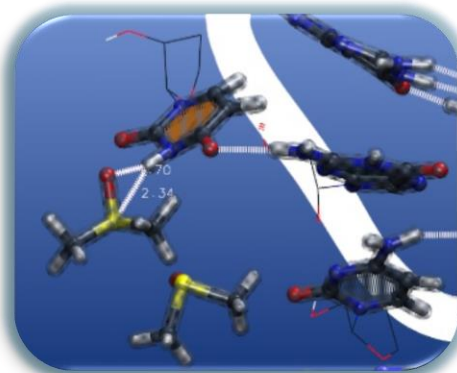


**G.3.**

**G.1.**



**G.2.**

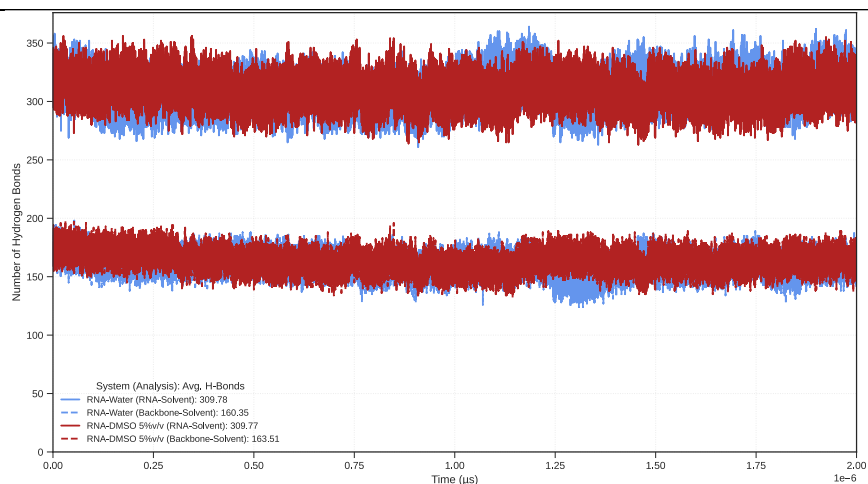


Nota. Caracterización de los enlaces de hidrógeno ARN-DMSO, orientación y átomos que interactúan, en representación de Lewis y 3D. **A.** Enlaces de hidrógeno entre el uracilo y el DMSO. **B.** Enlaces de hidrógeno entre la guanina y el DMSO. **C.** Enlaces de hidrógeno entre la adenina y el DMSO. **D.** Enlaces de hidrógeno entre la citosina y el DMSO. **E.** Enlaces de hidrógeno entre el backbone y el DMSO. **F. y G.** Enlaces de hidrógeno adicionales observados, enlaces de hidrógeno del azufre-base nitrogenada, doble enlace de hidrógeno 2 DMSO - base nitrogenada y enlaces de hidrógeno entre ribosas para estabilizar el backbone.

En el *Esquema 1* se visualizan las orientaciones más comunes en que el DMSO interactúa con cada residuo del fragmento de ARN. En la ilustración A.1 y A.2 se observa que el DMSO forma enlaces de hidrógeno con el grupo amino (-NH) del uracilo. En las ilustraciones B.1, B.2 y B.3 se identifica que el DMSO forma enlaces de hidrógeno con el grupo amino (-NH<sub>2</sub>) de la guanina. En las ilustraciones C.1 y C.2 se identifica que el DMSO forma enlaces de hidrógeno con el grupo amino (-NH<sub>2</sub>) de la adenina. En las ilustraciones D.1, D.2 y D.3 se identifica que el DMSO forma enlaces de hidrógeno con el grupo amino (-NH<sub>2</sub>) de la citosina. En las ilustraciones E.1, E.2 y E.3 se identifica que el DMSO forma enlaces de hidrógeno con el grupo hidroxilo 2' (-2'OH) de la ribosa del backbone. En la ilustración F.1 se identifica que el átomo de azufre (-S) del DMSO puede formar enlaces de hidrógeno con los grupos donadores de enlaces de hidrógeno de las bases nitrogenadas. En las ilustraciones G.1 y G.2 se identifica que el DMSO puede formar enlaces de hidrógeno en proporción 2:1 con los grupos donadores que tengan 2 hidrógenos como los grupos amino (-NH<sub>2</sub>) de las bases GUA, ADE y CYT, y en la ilustración G.3 se evidencia la presencia de enlaces de hidrógeno entre las ribosas del backbone que estabilizan estructural del backbone. La preferencia de enlace de hidrógeno de DMSO por la guanina en comparación la citosina se debe a la mayor presencia de grupos donadores de enlaces de hidrógeno (-NH y -NH<sub>2</sub>) en la guanina.

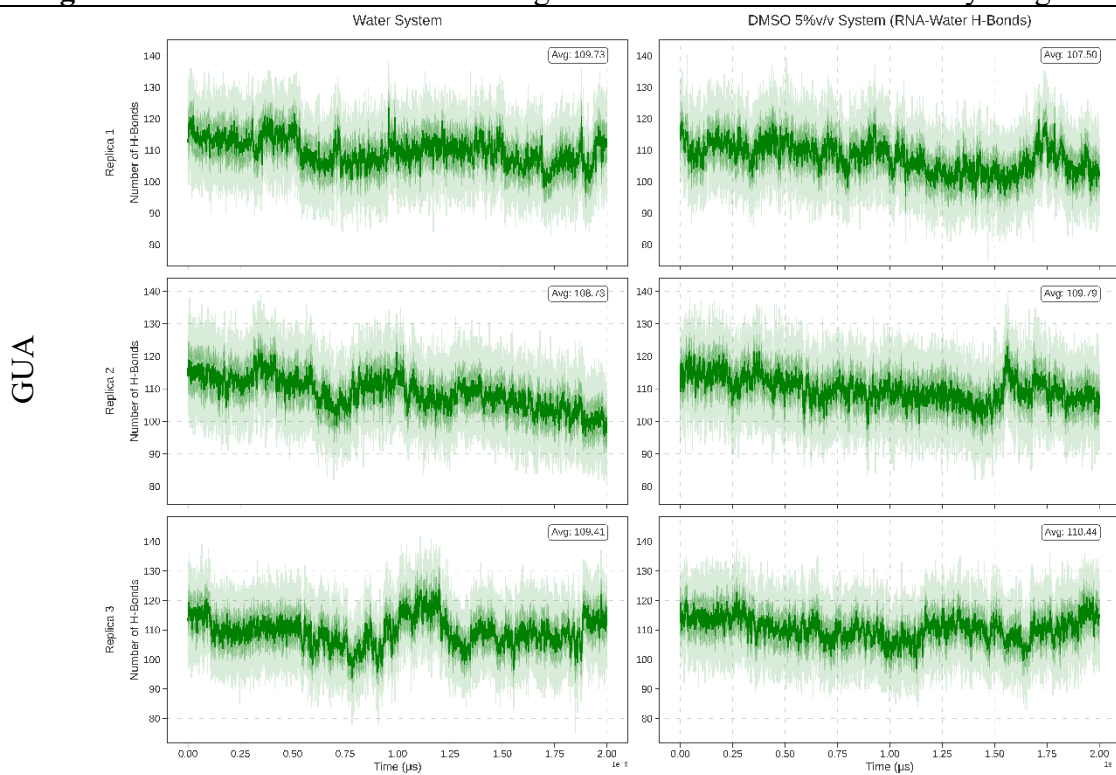


3

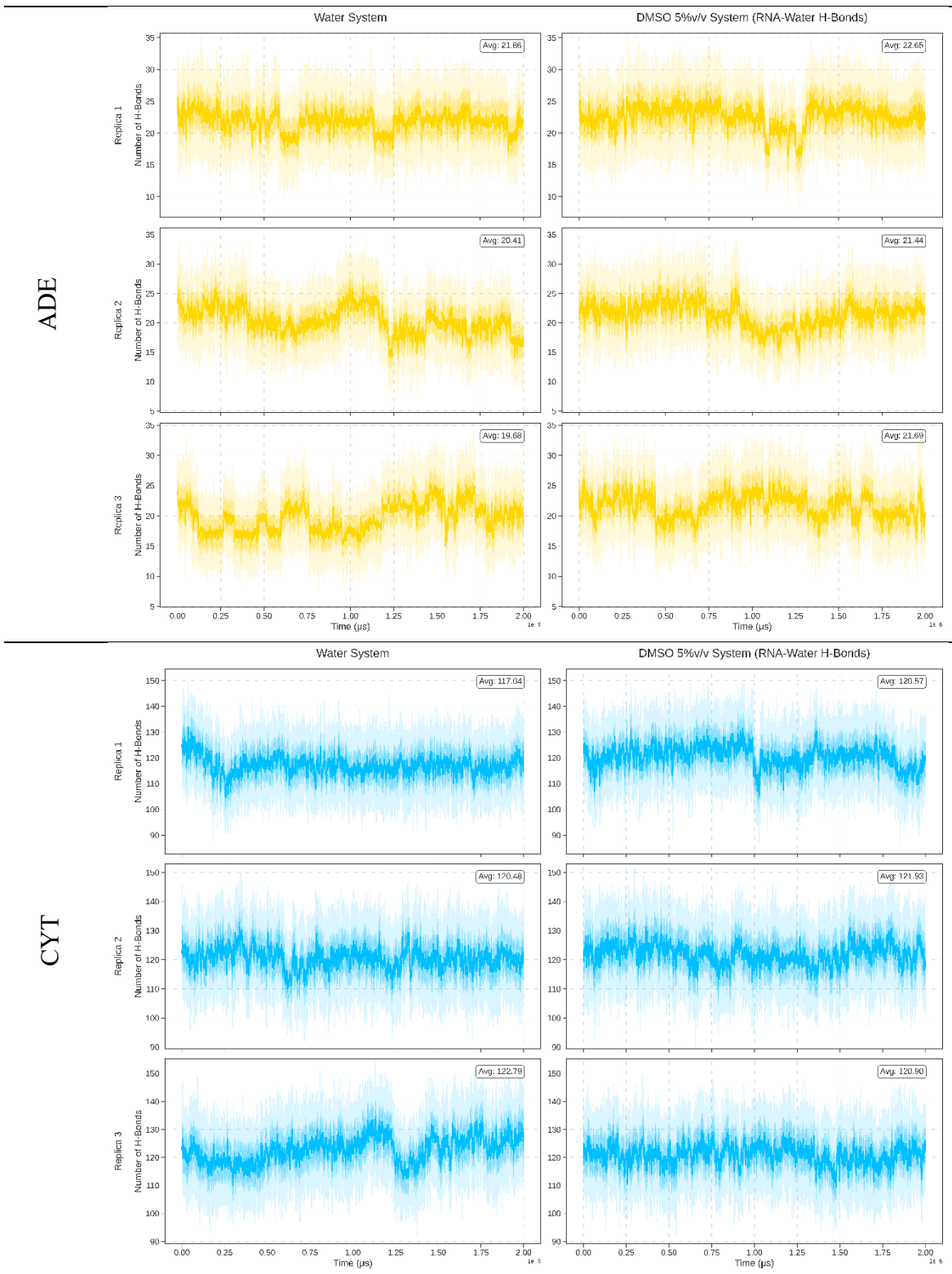


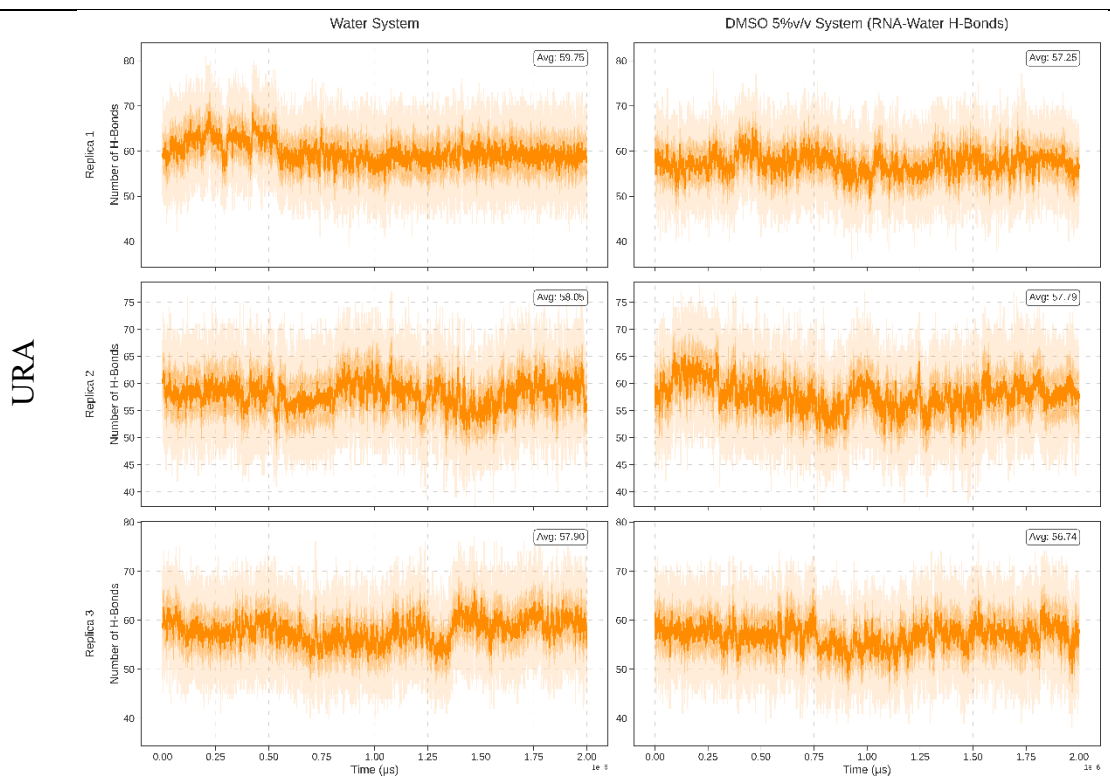
Nota. Comparativa entre las réplicas de cada sistema. En rojo se observa el número en el sistema ARN en DMSO al 5%v/v, y en azul el número de enlaces de hidrógeno entre el ARN y el agua. En la parte superior de la gráfica se observa el conteo de enlaces de hidrógeno entre el ARN-agua y en la parte inferior el conteo de enlaces de hidrógeno backbone-agua.

**Figura 27.** Conteo de enlaces de hidrógeno entre los cuatro nucleótidos y el agua



GUA

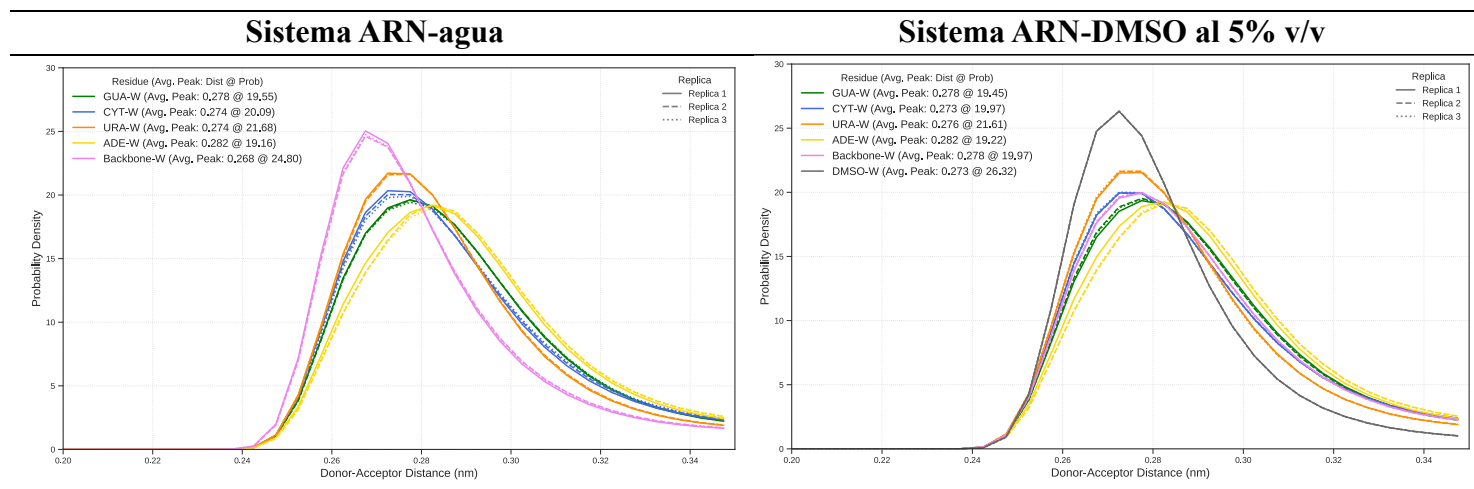




*Nota.* Comparativa entre las réplicas de cada sistema. Se etiqueta el valor promedio de los enlaces de hidrógenos con las moléculas de agua.

El número promedio de enlaces de hidrógeno de cada nucleótido con el agua no varió en presencia de DMSO. Se observa que el comportamiento coincide entre las réplicas de los sistemas de ARN en fase acuosa y ARN en presencia de DMSO al 5% v/v.

La distribución de los enlaces de hidrógeno entre el DMSO con los residuos del fragmento de ARN fue del rango de 2,6 a 2,9 Å, distancias en el orden de los enlaces de hidrógeno débil, lo que indica que las moléculas que hidratan al fragmento de ARN se difunden durante la trayectoria. El uracilo y el backbone fueron los residuos con mayor distribución de enlaces de hidrógeno a distancias de 2,8 y 2,7 Å, *Figura 28*.

**Figura 28.** Distribución de las distancias de enlace de hidrógeno ARN-agua

Nota. Comparativa entre las réplicas de cada sistema. Distribución de las distancias de enlaces de hidrógeno de la guanina, adenina, citosina, uracilo, el backbone y el DMSO con el agua en verde, amarillo, azul, naranja, rosado y negro respectivamente.

## 6. Conclusiones

En el presente estudio, mediante simulaciones de dinámica molecular, cada una de 2  $\mu$ s, se reveló que el DMSO a una concentración del 5% v/v no promueve la desestabilización de la estructura secundaria del ARN mediante la ruptura directa de los enlaces de hidrógeno intramoleculares. Si bien el DMSO muestra una ligera preferencia para interactuar con los nucleótidos de G y de U de regiones flexibles de la estructura, no se observó la formación de enlaces de hidrógeno estables entre el ARN y el DMSO, por el contrario, los enlaces de hidrógeno formados fueron transitorios y de corto tiempo de residencia.

Se resalta el reto del modelado computacional del ARN. A pesar de que en la escala de tiempo simulada (2 $\mu$ s) no se observó la desnaturalización completa del fragmento de ARN, si se observó consistentemente una apertura parcial de los extremos 5' y 3' en todas las simulaciones, debida a

la inestabilidad y flexibilidad intrínseca del fragmento de ARN modelado. Además, el fragmento de ARN exploró una vasta dinámica conformacional que varió inclusive entre réplicas de una misma simulación. Además, la escala de tiempo simulada es una variable crítica. Si bien en 2  $\mu$ s no se observa un cambio estructural significativo, los cambios estructurales observados podrían ser los precursores de eventos como la desnaturalización del ARN a escalas de tiempo más largas.

En conjunto, este trabajo proporciona la primera visión a escala atómica de las interacciones entre el DMSO y el ARN, sentando las bases para una comprensión del efecto del DMSO sobre la estabilidad del fragmento de ARN de 30 ribonucleótido del gen DRD4 humano.

### **Recomendaciones**

Los resultados de este trabajo sugieren que el mecanismo de desestabilización de la estructura secundaria del ARN podría ser indirecto y futuras investigaciones podrían explorar el efecto del DMSO sobre la red de las capas de hidratación del ARN, usar concentraciones más elevadas de DMSO (e.g. 10-15% v/v), simular mayores escalas de tiempo de simulación (e.g.  $\mu$ s) y técnicas mejoradas como replica exchange o un campo de fuerza polarizable para lograr una comprensión más amplia del mecanismo de interacción ARN-DMSO.

### Referencias bibliográficas

- Bussi, G., Donadio, D., & Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, *126*(1), 014101. <https://doi.org/10.1063/1.2408420>
- Cadena-Caballero, C., Vera, L., Barrios Hernandez, C., Rueda-Plata, D., Forero-Buitrago, L., Torres-Jimenez, C., Lizarazo-Gutierrez, E., Agudelo-Rodriguez, M., & Martinez, F. (2022). Denaturing and dNTPs reagents improve SARS-CoV-2 detection via single and multiplex RT-qPCR. *F1000Research*, *11*, 331. <https://doi.org/10.12688/f1000research.109673.1>
- Chakrabarti, R., & Schutt, C. E. (2001). The enhancement of PCR amplification by low molecular-weight sulfones. *Gene*, *274*(1-2), 293-298. [https://doi.org/10.1016/s0378-1119\(01\)00621-7](https://doi.org/10.1016/s0378-1119(01)00621-7)
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *The Journal of Chemical Physics*, *98*(12), 10089-10092. <https://doi.org/10.1063/1.464397>
- Darty, K., Denise, A., & Ponty, Y. (2009). VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics (Oxford, England)*, *25*(15), 1974-1975. <https://doi.org/10.1093/bioinformatics/btp250>
- Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. *The Computer Journal*, *7*(2), 149-154. <https://doi.org/10.1093/comjnl/7.2.149>
- Francisco Jose Martinez Perez-Es. (2025.). *Profesores UIS*. Recuperado 12 de septiembre de 2025, de <https://profesores.uis.edu.co/francisco-jose-martinez-perez-es/>
- George, A., Mondal, S., Purnaprajna, M., & Athri, P. (2022). Review of Electrostatic Force Calculation Methods and Their Acceleration in Molecular Dynamics Packages Using

Graphics Processors. *ACS Omega*, 7(37), 32877-32896.

<https://doi.org/10.1021/acsomega.2c03189>

*GROMACS 2024.4 documentation*. (2024.). Recuperado 10 de diciembre de 2024, de

<https://manual.gromacs.org/documentation/2024.4/index.html>

Hardjasa, A., Ling, M., Ma, K., & Yu, H. (2010). *Investigating the Effects of DMSO on PCR Fidelity Using a Restriction Digest-Based Method*. 14.

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12), 1463-1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12%253C1463::AID-JCC4%253E3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12%253C1463::AID-JCC4%253E3.0.CO;2-H)

Hoehndorf, R., Batchelor, C., Bittner, T., Dumontier, M., Eilbeck, K., Knight, R., Mungall, C., Richardson, J., Stombaugh, J., Westhof, E., Zirbel, C., & Leontis, N. (2011). The RNA Ontology (RNAO): An ontology for integrating RNA sequence and structure data. *Applied Ontology*, 6, 53-89. <https://doi.org/10.3233/AO-2011-0082>

Huang, J., & MacKerell Jr, A. D. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry*, 34(25), 2135-2145. <https://doi.org/10.1002/jcc.23354>

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33-38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)

Ingold, C. K. (1940). The Nature of the Chemical Bond and the Structure of Molecules and Crystals. *Nature*, 145(3678), 644-645. <https://doi.org/10.1038/145644b0>

Laing, C., & Schlick, T. (2011). Computational approaches to RNA structure prediction, analysis, and design. *Current Opinion in Structural Biology*, 21(3), 306-318.

<https://doi.org/10.1016/j.sbi.2011.03.015>

Leach, A. R. (2001). *Molecular Modelling: Principles and Applications*. Prentice Hall.

Lee, J., Vogt, C. E., McBairty, M., & Al-Hashimi, H. M. (2013). Influence of Dimethylsulfoxide on RNA Structure and Ligand Binding. *Analytical Chemistry*, 85(20),

10.1021/ac402038t. <https://doi.org/10.1021/ac402038t>

*Lehninger Principles of Biochemistry, 8th Edition | Macmillan Learning US*. (2021.).

Recuperado 15 de diciembre de 2024, de

<https://www.macmillanlearning.com/college/us/product/Lehninger-Principles-of-Biochemistry/p/1319228003>

Lestarini, I. A., Suryani, D., & Sabrina, Y. (2019). An Efficient Polymerase Chain Reaction (PCR) Enhancer for Highly Guanine-Cytosine (GC)-Rich DNA Sequences. *Bali Medical Journal*, 8(2), Article 2. <https://doi.org/10.15562/bmj.v8i2.1357>

Lu, X.-J., Bussemaker, H. J., & Olson, W. K. (2015). DSSR: An integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Research*, 43(21), e142.

<https://doi.org/10.1093/nar/gkv716>

Magnus, M. (2022). rna-tools.online: A Swiss army knife for RNA 3D structure modeling workflow. *Nucleic Acids Research*, 50(W1), W657-W662.

<https://doi.org/10.1093/nar/gkac372>

Mark, P., & Nilsson, L. (2001). Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *The Journal of Physical Chemistry A*, 105(43), 9954-9960.

<https://doi.org/10.1021/jp003020w>

- Nazareth, J. L. (2024). Conjugate-Gradient Methods. En *Encyclopedia of Optimization* (pp. 1-7). Springer, Cham. [https://doi.org/10.1007/978-3-030-54621-2\\_85-1](https://doi.org/10.1007/978-3-030-54621-2_85-1)
- Nöthen, M. M., Cichon, S., Hemmer, S., Hebebrand, J., Remschmidt, H., Lehmkuhl, G., Poustka, F., Schmidt, M., Catalano, M., & Fimmers, R. (1994). Human dopamine D4 receptor gene: Frequent occurrence of a null allele and observation of homozygosity. *Human Molecular Genetics*, 3(12), 2207-2212. <https://doi.org/10.1093/hmg/3.12.2207>
- Oh, K.-I., Rajesh, K., Stanton, J. F., & Baiz, C. R. (2017). Quantifying Hydrogen-Bond Populations in Dimethyl Sulfoxide/Water Mixtures. *Angewandte Chemie International Edition*, 56(38), 11375-11379. <https://doi.org/10.1002/anie.201704162>
- Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12), 7182-7190. <https://doi.org/10.1063/1.328693>
- (PDF) Pseudoknots: RNA Structures with Diverse Functions. (2024). *ResearchGate*. <https://doi.org/10.1371/journal.pbio.0030213>
- Ramírez, J. L. R. (2014.). *TESIS: ESTUDIO DE LA PERMEABILIDAD TRANSDÉRMICA DE NANOPARTÍCULAS DE PLATA*.
- Schiller, U. D. (2017). *An overview of integration schemes for molecular dynamics simulations*.
- Šponer, J., Bussi, G., Krepl, M., Banáš, P., Bottaro, S., Cunha, R. A., Gil-Ley, A., Pinamonti, G., Poblete, S., Jurečka, P., Walter, N. G., & Otyepka, M. (2018). RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chemical Reviews*, 118(8), 4177-4338. <https://doi.org/10.1021/acs.chemrev.7b00427>

- Stachura, S. S., Malajczuk, C. J., & Mancera, R. L. (2018). Molecular dynamics simulations of a DMSO/water mixture using the AMBER force field. *Journal of Molecular Modeling*, 24(7), 174. <https://doi.org/10.1007/s00894-018-3720-x>
- Steiner, T. (2002). The hydrogen bond in the solid state. *Angewandte Chemie (International Ed. in English)*, 41(1), 49-76. [https://doi.org/10.1002/1521-3773\(20020104\)41:1%253C48::aid-anie48%253E3.0.co;2-u](https://doi.org/10.1002/1521-3773(20020104)41:1%253C48::aid-anie48%253E3.0.co;2-u)
- Tariq, N., Kume, T., Luo, L., Cai, Z., Dong, S., & Macgregor, R. B. (2022). Dimethyl sulfoxide (DMSO) is a stabilizing co-solvent for G-quadruplex DNA. *Biophysical Chemistry*, 282, 106741. <https://doi.org/10.1016/j.bpc.2021.106741>
- Tunçer, S., Gurbanov, R., Sheraj, I., Solel, E., Esenturk, O., & Banerjee, S. (2018). Low dose dimethyl sulfoxide driven gross molecular changes have the potential to interfere with various cellular processes. *Scientific Reports*, 8(1), 14828. <https://doi.org/10.1038/s41598-018-33234-z>
- Ulz, M. H. (2012). Invariance of Parrinello-Rahman molecular dynamics and its relation to continuum mechanics. *PAMM*, 12(1), 419-420. <https://doi.org/10.1002/pamm.201210198>
- Varani, G., & McClain, W. H. (2000). The G·U wobble base pair. *EMBO reports*, 1(1), 18-23. <https://doi.org/10.1093/embo-reports/kvd001>
- Walter, N. G., Woodson, S. A., & Batey, R. T. (Eds.). (2009). *Non-protein coding RNAs*. Springer.
- Watkins, A. M., Rangan, R., & Das, R. (2020). FARFAR2: Improved De Novo Rosetta Prediction of Complex Global RNA Folds. *Structure*, 28(8), 963-976.e6. <https://doi.org/10.1016/j.str.2020.05.011>

Wilcox, A. H., Delwart, E., & Diaz-Munoz, S. L. (2019). Next-generation sequencing of dsRNA is greatly improved by treatment with the inexpensive denaturing reagent DMSO.

*MICROBIAL GENOMICS*, 5(11), 000315. <https://doi.org/10.1099/mgen.0.000315>

Zhang, J., Fei, Y., Sun, L., & Zhang, Q. C. (2022). Advances and opportunities in RNA structure experimental determination and computational modeling. *Nature Methods*, 19(10),

Article 10. <https://doi.org/10.1038/s41592-022-01623-y>

Zhao, Q., Zhao, Z., Fan, X., Yuan, Z., Mao, Q., & Yao, Y. (2021). Review of machine learning methods for RNA secondary structure prediction. *PLoS Computational Biology*, 17(8),

e1009291. <https://doi.org/10.1371/journal.pcbi.1009291>

## Apendice

### ▪ Algoritmos y parámetros de simulación

#### Integrador:

Es el algoritmo usado para integrar las ecuaciones de movimiento de newton para propagar las coordenadas, velocidades y aceleraciones del sistema a lo largo del tiempo. Se empleó algoritmo de **Verlet (md-vv)**(Schiller, 2008.), algoritmo usado ampliamente por su robustez, reversibilidad temporal y eficiencia computacional en la conservación de la energía.

1) Determinación de nuevas posiciones a partir de  $r(t)$  y  $v(t)$  para un paso de tiempo  $\Delta t$

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 \quad (\text{ec. 4})$$

2) Cálculo de la nueva aceleración  $a(t + \Delta t)$  con base en la posición  $r(t + \Delta t)$

$$a(t + \Delta t) = \frac{F(t + \Delta t)}{m} \quad (\text{ec. 5})$$

3) Actualiza velocidades  $v(t + \Delta t)$

$$v(t + \Delta t) = v(t) + \frac{1}{2}[a(t) + a(t + \Delta t)]\Delta t \quad (\text{ec. 6})$$

4) Loop de los anteriores pasos en el tiempo

### Minimización de energía:

Proceso previo a la dinámica molecular, implementado para que la estructura del sistema adopte la configuración espacial con menor energía potencial, eliminando tensiones estéricas y posibles solapamientos atómicos. Se usó el algoritmo de **gradiente conjugado (CG)** (Fletcher & Reeves, 1964), un algoritmo iterativo que ajusta las posiciones atómicas en la dirección opuesta al gradiente de la energía potencial hasta alcanzar el mínimo global de energía del sistema.

$$F(t + \Delta t) = -\nabla U(r_1, r_2 \dots r_N) \quad (\text{ec. 7})$$

### Restricciones:

Algoritmo implementado para restringir los enlaces covalentes que involucran átomos de hidrógeno (ej., C-H, O-H). Esto elimina las vibraciones de alta frecuencia, que no son relevantes para los movimientos macromoleculares, permitiendo así el uso de un paso de tiempo de integración mayor (e.g, 2 fs) sin perder estabilidad numérica. Se empleó el algoritmo **LINCS** (Hess et al., 1997), método usado ampliamente en dinámica molecular ya que aplica una corrección a las posiciones atómicas en cada paso para satisfacer las restricciones de distancia ( $d_{ij}$ ) definidas.

$$|r_i - r_j|^2 - d_{ij}^2 = 0 \quad (\text{ec. 8})$$

Dónde:  $r_i, r_j$  son posiciones atómicas y  $d_{ij}$  es la distancia restringida.

**Interacciones electrostáticas:**

Algoritmo empleado para el cálculo eficiente de las interacciones electrostáticas entre todas las partículas del sistema de forma periódica. Se empleó el algoritmo de **Particle Mesh Ewald (PME)** (George et al., 2022), método que divide el cálculo en dos componentes; Las interacciones electrostáticas de corto alcance usando el potencial de Coulomb con un componente directo en el espacio real; y las interacciones electrostáticas de largo alcance, con un componente recíproco en el espacio de Fourier.

$$E_{direct} = \sum_{r_{ij} < r_c} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (\text{ec. 9})$$

$$E_{recip} = \frac{1}{2V} \sum_{k \neq 0} \frac{4\pi}{k^2} e^{-|k|^2/4\alpha^2} \cdot |\rho(k)|^2 \quad (\text{ec. 10})$$

Donde:  $\mathbf{k}$ : Son los vectores recíprocos,  $\rho(\mathbf{k})$ : Es la densidad de carga transformada de Fourier y,  $V$ : es el volumen de la caja de simulación.

**Equilibración de presión y temperatura:**

Algoritmos con los que se mantiene la temperatura (NVT) y presión (NPT) del sistema constante durante las simulaciones. Para la equilibración de la presión se implementó el algoritmo de **Parrinello-Rahman** (Ulz, 2012), método que mantiene la presión de referencia constante ajustando el volumen y la forma de la caja de simulación, permitiendo cambios anisotrópicos. El algoritmo de Parrinello-Rahman actualiza estos vectores de caja a lo largo del tiempo introduciendo un lagrangiano entendido en el que acopla ecuaciones de movimiento para las partículas y la caja de simulación, y adicionalmente calcula un tensor de presión completo para regular la deformación anisotrópica de la caja. Por otro lado, para la equilibración de la temperatura

se implementó el algoritmo **Velocity-rescale (V-rescale)** (Bussi et al., 2007), modificación del termostato de Berendsen que mantiene la temperatura de referencia mediante un reescalado estocástico de las velocidades de las partículas, asegurando mantener una temperatura cercana a la de referencia mediante una correcta distribución canónica de la energía cinética.

Halla el producto de las velocidades de todas las partículas por un factor que depende de la temperatura actual y la temperatura objetivo.

$$dK = (K_0 - K) \frac{dt}{\tau} + 2 \sqrt{\frac{KK_0}{N_f}} \frac{dW}{\sqrt{\tau}} \quad (\text{ec. 11})$$

Donde:  $K_0$  es la energía cinética objetivo correspondiente a la temperatura deseada,  $\tau$  es la constante de tiempo de acoplamiento,  $N_f$  es el número de grados de libertad y  $dW$  es un proceso de Wiener (ruido aleatorio).

#### ▪ Métodos de análisis

##### **Función de distribución radial (RDF)**

La función de distribución radial es una herramienta estadística que permite medir la probabilidad de encontrar una partícula a una distancia ( $r$ ) con respecto a una partícula de referencia, la probabilidad se denota como  $g(r)$ . Mediante este método se logra caracterizar la estructura local del solvente alrededor de la macromolécula identificando las capas de solvatación.

La función de distribución radial se define como la densidad local a una distancia  $r$  entre la densidad total del volumen del sistema:

$$g(r) = \rho(r) / \rho_{bulk}$$

Y para un sistema con dos tipos de partículas, A (referencia) y B (objetivo) se calcula como:

$$g_{AB}(r) = \frac{V}{N_A N_B} \sum_i \sum_j \frac{\langle NB(r, r + \Delta r) \rangle}{4\pi r^2 \Delta r} \quad (\text{ec. 12})$$

Donde:  $\langle NB(r, r + \Delta r) \rangle$  es el número promedio de partículas B encontradas en una capa esférica de grosor  $\Delta r$  a una distancia  $r$  de una partícula A.  $V$  es el volumen total de la caja de simulación.  $N_A$  y  $N_B$  son el número total de partículas A y B en el sistema, respectivamente. Y el volumen de la capa esférica que normaliza el conteo es,  $4\pi r^2 \Delta r$ .

### **Análisis de enlace de hidrógeno**

Conteo de la formación de enlaces de hidrógeno entre los grupos donadores y aceptores de la molécula de interés con el solvente a lo largo de la simulación. La formación de enlaces de hidrógeno se determina con criterios geométricos mediante parámetros estandarizados experimental y teóricamente, como la distancia mínima donador-aceptor, y ángulo mínimo donador-H-aceptor.

### **Desviación cuadrática media (RMSD)**

Desviación promedio conformacional con respecto a la estructura inicial a lo largo de la simulación. Método mediante el cual se evalúa la estabilidad conformacional global. Valores bajos y estables indican que el sistema no se desvía significativamente de su estructura de partida

$$RMSD(t) = \sqrt{\frac{1}{M} \sum_{i=1}^M \|r_i(t) - r_i^{ref}\|^2} \quad (\text{ec. 13})$$

### **Fluctuación cuadrática media (RMSF)**

Fluctuación promedio de la posición de cada átomo de la macromolécula con respecto a su posición media a lo largo de la simulación. Indicando la flexibilidad local de las diferentes regiones de la estructura. Valores altos de RMSF corresponden con regiones altamente flexibles (e.g, bucles)

$$RMSF = \sqrt{\langle ||r_i||^2 \rangle} = \sqrt{\frac{1}{M} \sum_{i=1}^M ||r_i(t) - \langle r_i \rangle||^2} \quad (\text{ec. 14})$$

### Área superficial accesible al solvente (SASA)

Medición del área de la molécula que es accesible al solvente. Cuantifica la exposición global y local de la estructura. Permite analizar cambios conformacionales asociados a plegamientos e interacciones. Se aproximan la superficie de la molécula mediante puntos discretos sobre la superficie de Van der Waals de cada átomo y el solvente se describe por medio de esferas de radio definido y el SASA se calcula contando y sumando el número moléculas del solvente en el área de cada esfera de la superficie de Van der Waals.

### Matriz de covarianza, PCA y FEL

La matriz de covarianza es un método mediante el cual se extraen los modos de movimiento colectivo dominantes de la trayectoria, identificando las conformaciones con coordenadas atómicas correlacionadas. Con la matriz de covarianza contiene la variación de las posiciones atómicas ( $r_i$ ) con respecto a su posición promedio ( $\langle r_i \rangle$ ) a lo largo de la simulación.

$$C_{ij} = \frac{1}{1 - M} \sum_{t=1}^M (r_i(t) - \langle r_i \rangle)(r_j(t) - \langle r_j \rangle) \quad (\text{ec. 15})$$

A partir de la diagonalización de la matriz de covarianza se obtienen los valores y vectores propios que indican la varianza y las direcciones de los movimientos colectivos del sistema mediante el cual se realiza el análisis de las componentes principales.

$$(C - \lambda)v = 0 \quad (\text{ec. 16})$$

Finalmente, proyectando las dos primeras componentes principales (PC1 y PC2) se visualizan los cambios conformacionales dominantes de la trayectoria en un espacio de baja dimensionalidad.

$$p_i(t) = (\Delta r(t))^T v_i \quad (\text{ec. 17})$$

A partir del paisaje conformacional construido con la proyección de los componentes principales se puede construir el paisaje de energía libre (FEL). Mediante el conteo y normalización de las conformaciones de la proyección de los componentes principales se obtiene una distribución de probabilidad  $P(PC1, PC2)$  y utilizando la ecuación de Boltzmann se convierte esta probabilidad en un valor de energía libre de Gibbs ( $\Delta G$ ), y se añade como una tercera dimensión en la proyección.

$$\Delta G = -k_b T \ln(P) \quad (\text{ec. 18})$$

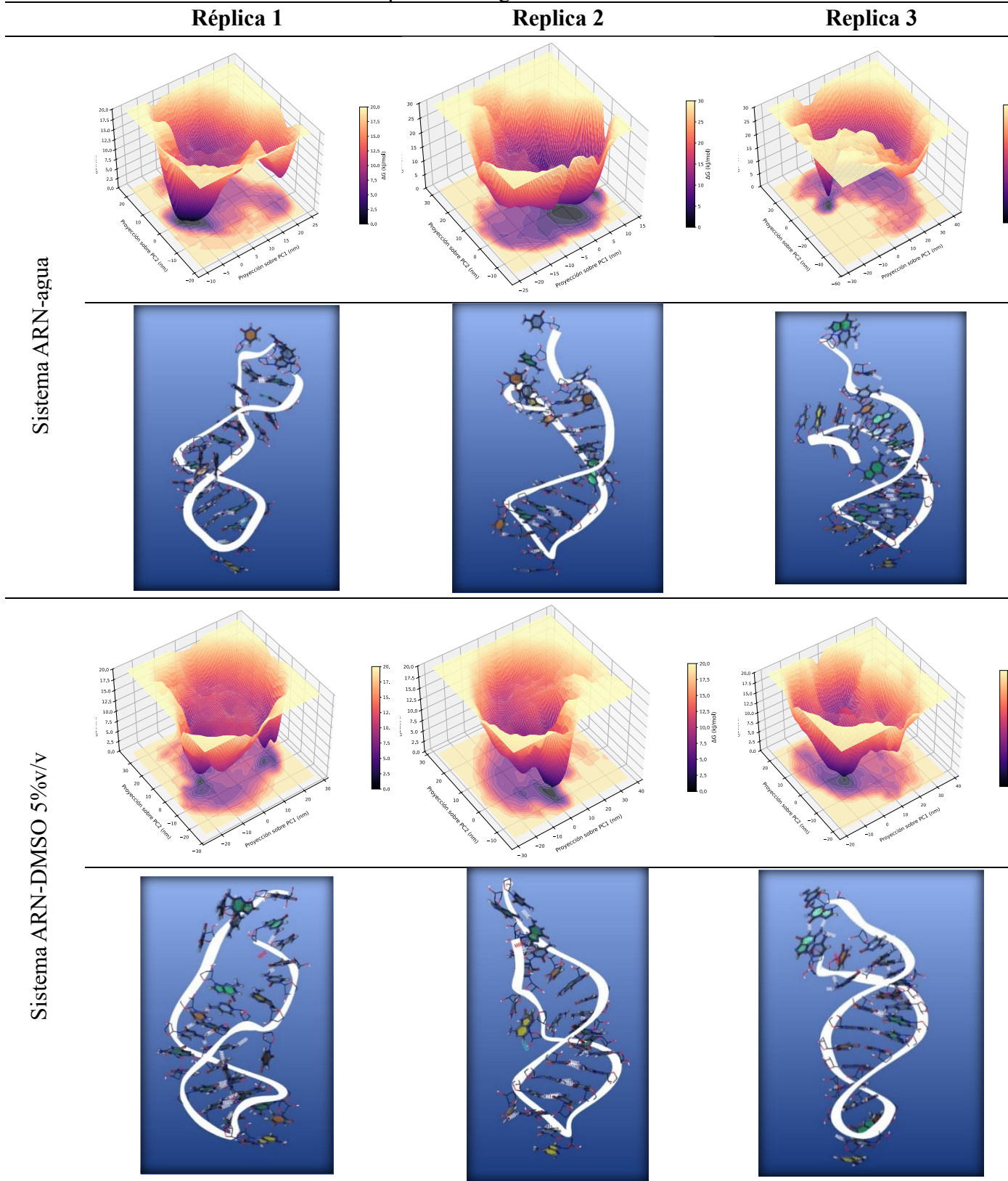
Donde:  $k_B$  es la constante de Boltzmann y  $T$  es la temperatura de la simulación en Kelvin.

#### ▪ **Análisis complementario de la dinámica conformacional del fragmento de ARN**

Se elaboró el análisis de componentes principales (PCA) de la dinámica conformacional del fragmento de ARN en las réplicas de cada sistema y se identificó la estructura más estable alcanzada en cada simulación, *Figura A.1*.

El paisaje energético permite analizar cualitativamente la basta adopción de conformaciones estructurales del fragmento de ARN. Además, permite observar que la dinámica conformacional en las réplicas de un mismo sistema difiere a pesar de mantener los mismos parámetros y tiempo de simulación, lo que evidencia la complejidad que se menciona en la literatura de trabajar con moléculas flexibles como el ARN

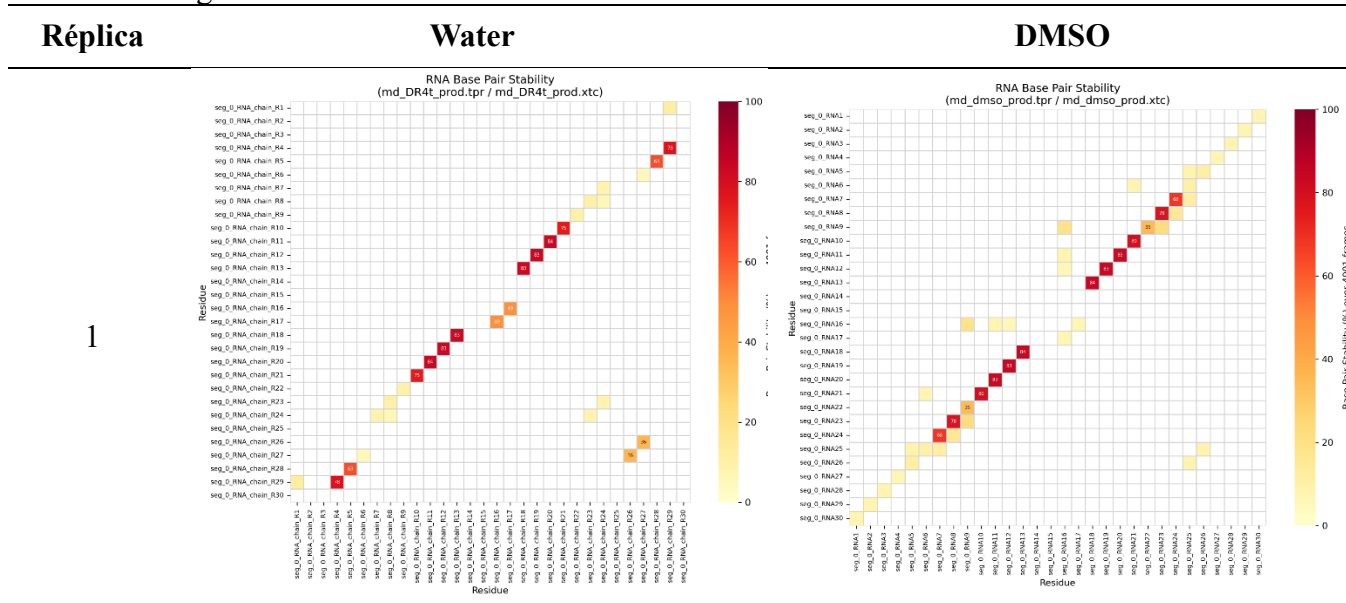
Figura A. 1. Paisaje de energía libre 3D y las conformaciones de menor energía libre en cada réplica del fragmento de ARN



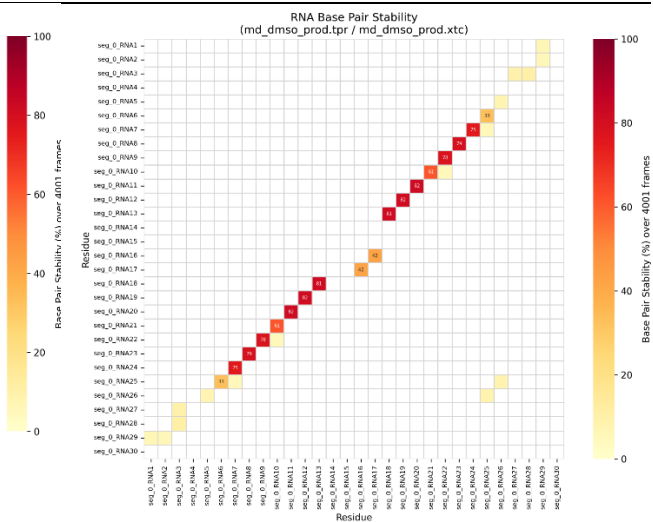
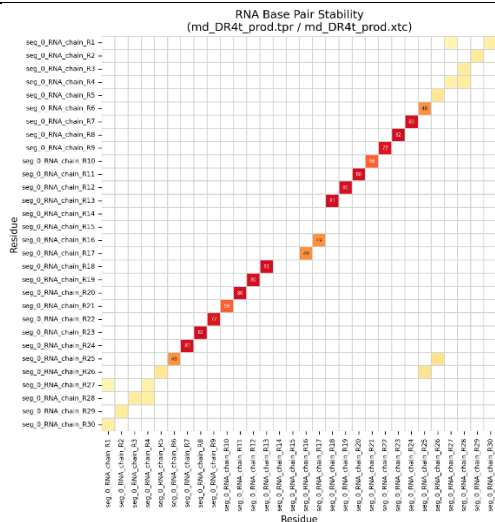
*Nota.* Proyección 3D de paisaje energético de las conformaciones estructurales del fragmento del ARN en cada simulación. En color morado oscuro se observa el mínimo global y los mínimos locales de las conformaciones estructurales con menor energía libre de Gibbs ( $\Delta G$ ) del fragmento de ARN y en color naranja conformaciones estructurales con mayor energía libre de Gibbs ( $\Delta G$ ). Las estructuras 3D pertenecen al mínimo global de cada paisaje energético.

Además, se presenta el mapa de calor, *Figura A.2*, con que se identificaron los apareamientos entre bases más estables del fragmento de ARN durante las simulaciones del análisis de la estabilidad estructural, *Figura 19*.

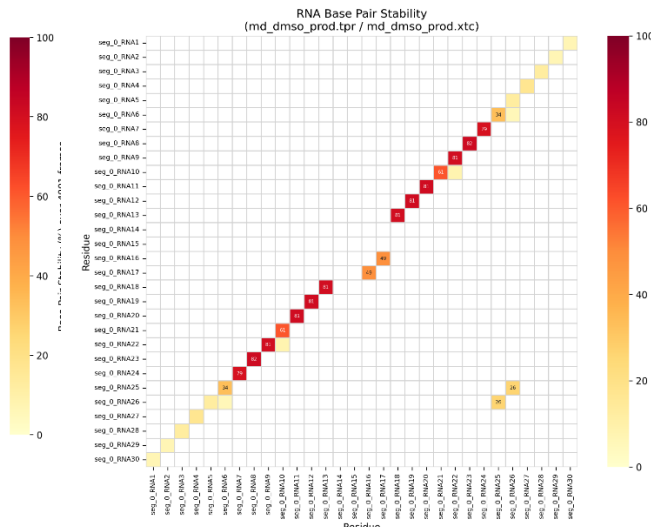
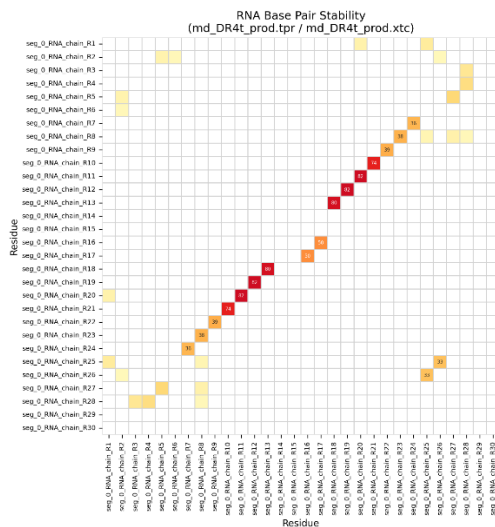
**Figura A. 2** Mapa de calor de la estabilidad de los enlaces de hidrógeno intramoleculares entre bases nitrogenadas



2



3

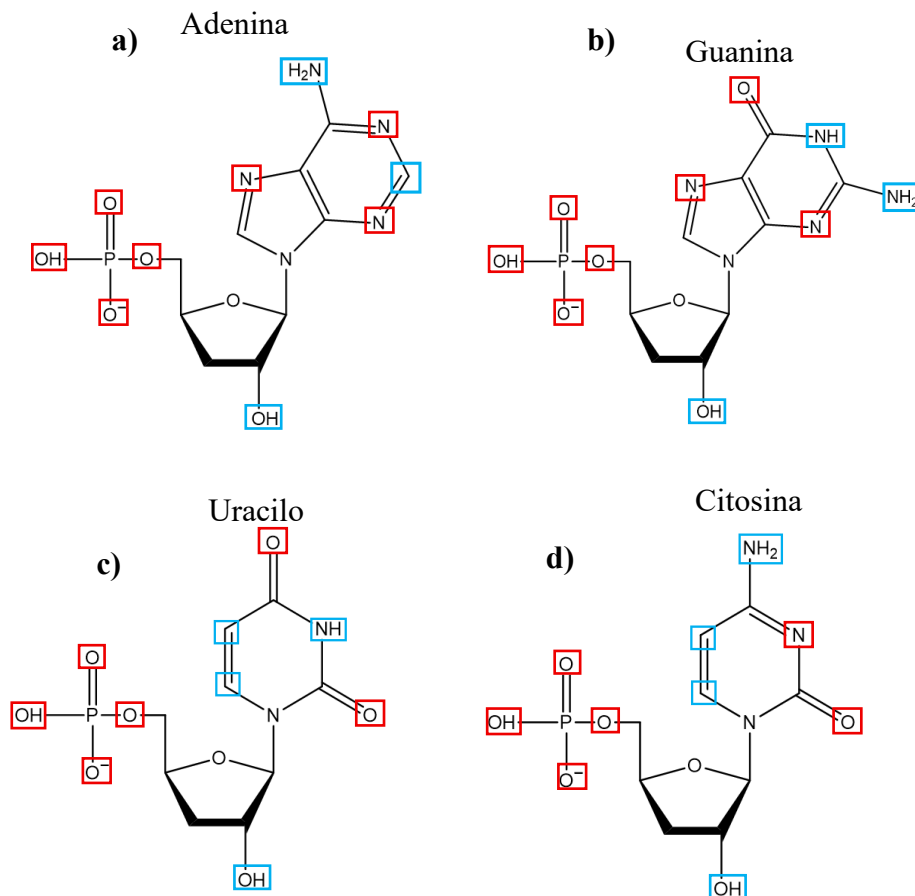


*Nota.* En tonos cálidos oscuros se representan los enlaces de hidrógenos presentes en la mayoría de frames de cada simulación, y en tonos más claros se representan los enlaces de hidrógeno menos presentes en la simulación. Análisis realizado con la librería de Python MD analysis y DSSR (Lu et al., 2015)

- **Análisis complementario de las interacciones intermoleculares ARN-Solvente**

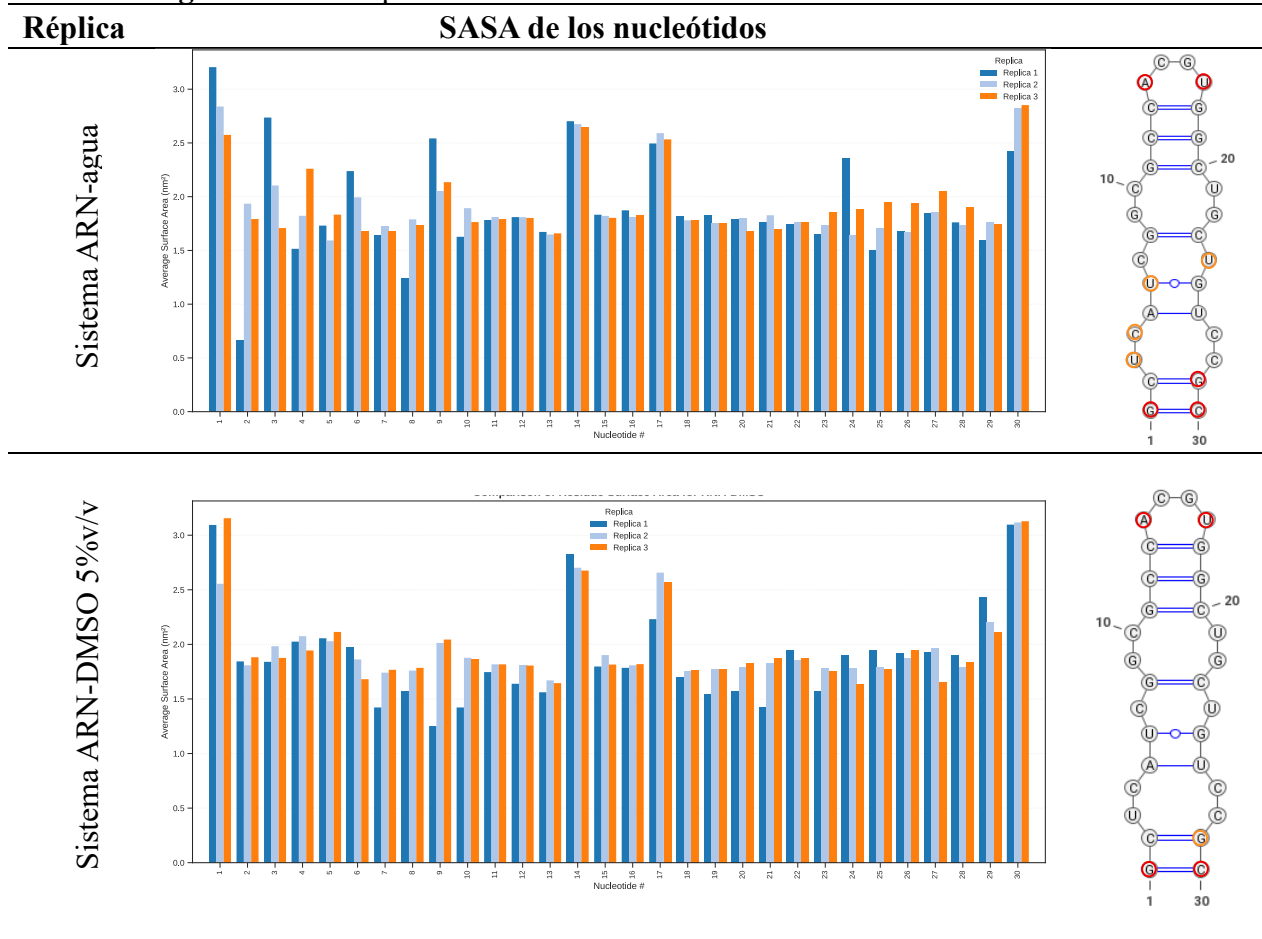
Los átomos seleccionados como donadores y aceptores de enlaces de hidrógeno de cada nucleótido para los análisis de las interacciones ARN-solvente, se ilustran en la *Figura A.3*.

**Figura A. 3.** Grupos donadores y aceptores de enlaces de hidrógeno de los nucleótidos del ARN



*Nota.* Los principales grupos donadores de enlaces de hidrógeno de los nucleótidos se señalan en color celeste (-CH, -NH, -NH<sub>2</sub> y -OH) y en rojo los principales átomos aceptores de enlaces de hidrógeno de los nucleótidos (-O y -N)

Con el objetivo de evaluar la accesibilidad y capacidad de interacción de cada nucleótido con el agua en las simulaciones en fase acuosa y en presencia de DMSO al 5% v/v, se identificaron los nucleótidos con mayor el área superficial accesible por el solvente (SASA) y se comparó el valor de SASA de cada nucleótido entre las réplicas de los dos sistemas.

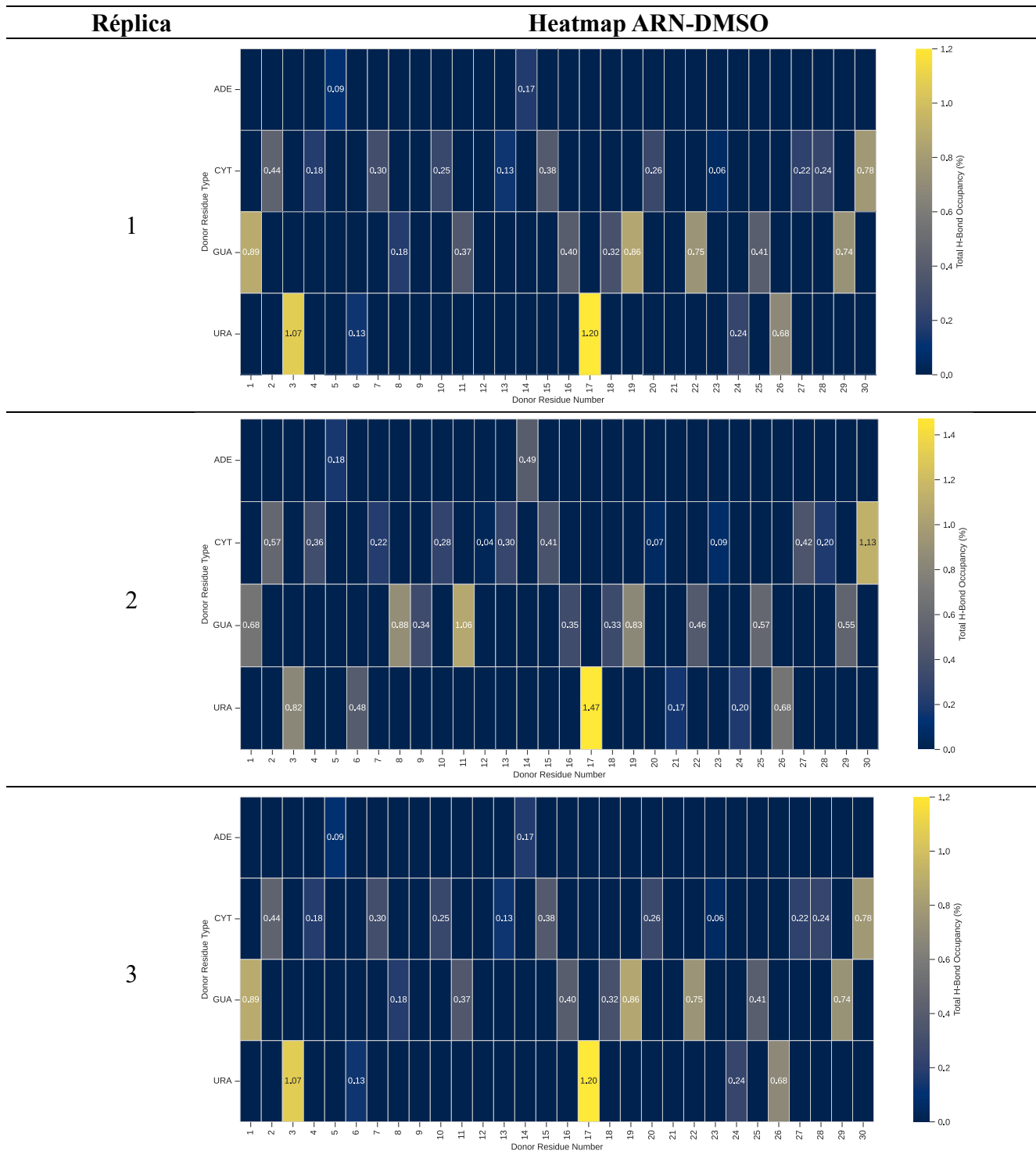
**Figura A. 4.** Comparación del SASA de cada nucleótido en las simulaciones

*Nota.* Diagramas de barras del valor de SASA de cada nucleótido en las tres réplicas de los sistemas de ARN en agua y ARN en DMSO al 5% v/v. A la derecha se representa la estructura secundaria del fragmento de ARN señalando los nucleótidos con mayor valor SASA.

De forma consistente, se observó que en todas las simulaciones de ARN en fase acuosa y en presencia de DMSO al 5%, los nucleótidos (G1, A14, U17 y C30) exhibieron los valores de SASA más altos. La alta exposición al solvente de estos nucleótidos se correlaciona directamente con su ubicación en la cadena, ya que estos nucleótidos forman parte de motivos estructurales más accesibles, como los bucles (loops) y en los extremos terminales 5' y 3'. No se observan cambios en los valores de SASA de los nucleótidos en presencia de DMSO. Se concluye que la accesibilidad

de las bases por parte del solvente depende directamente de su estructura secundaria y conformación 3D.

**Figura A. 5.** Mapa de calor de los tiempos de residencia de los enlaces de hidrógeno ARN-DMSO



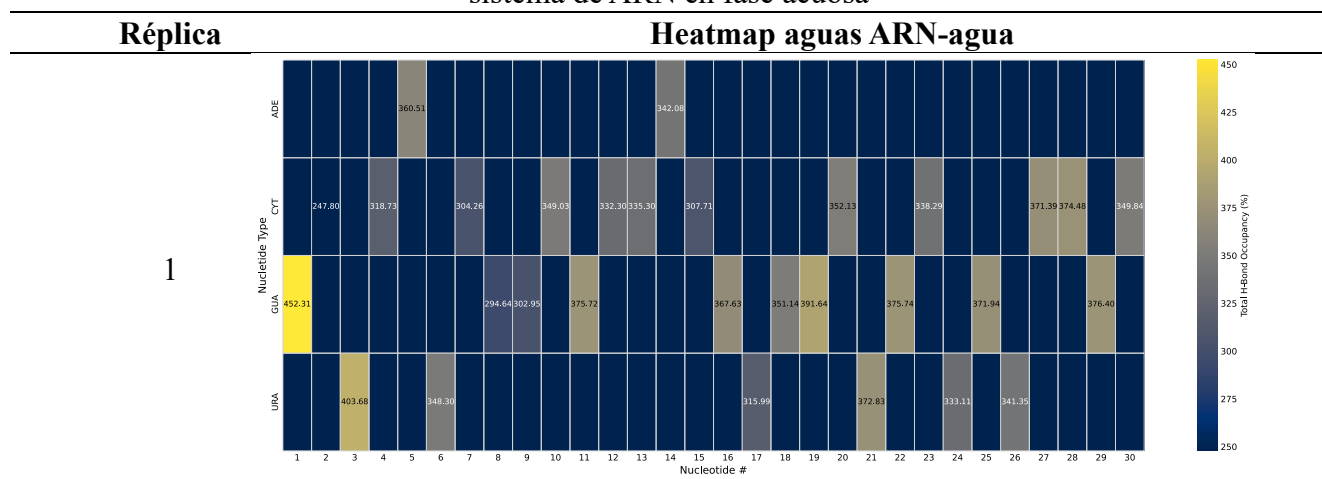
El porcentaje de la ocupancia del enlace de hidrógeno ARN-DMSO se determinó de acuerdo con el número total de frames en que hay presencia de enlace de hidrógeno nucleótido-DMSO con respecto a los frames de la trayectoria, de acuerdo con la siguiente relación:

$$Ocupancia = \frac{nf}{NF} * 100 \tag{ec. 19}$$

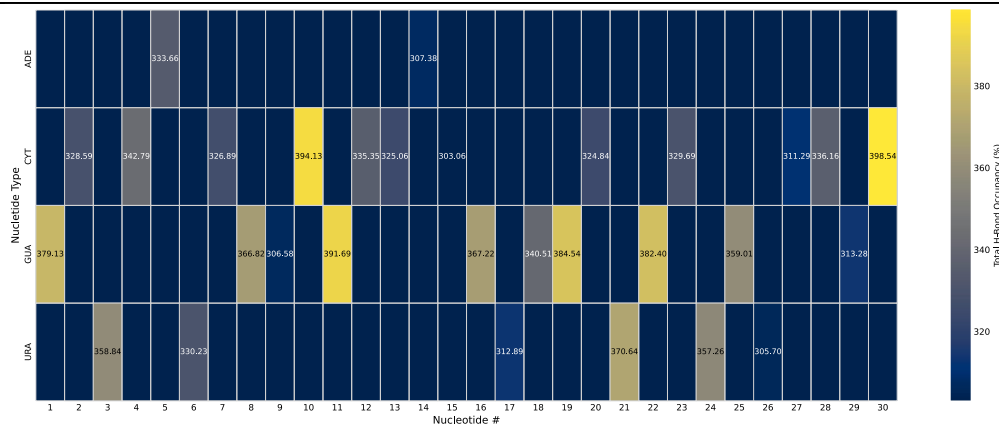
Donde nf = número de frames de la interacción y NF = número de frames totales de la simulación. Y se suman todas las interacciones con el mismo nucleótido.

Se obtienen bajas ocupancias de los enlaces de hidrógeno ARN-DMSO. Esto significa que los tiempos de residencia de estos enlaces de hidrógeno son cortos e indica que las interacciones ARN-DMSO son transitorias. Esto se explica ya que el DMSO es un aceptor de puentes de hidrógeno menos fuerte que el agua debida su estructura y densidad de carga. Con base en la *ec. 19* El porcentaje de tiempo de residencia de 1.2% significa que en los 200000 frames de la simulación hay 2400 frames con la formación de enlaces de hidrógeno ARN-DMSO.

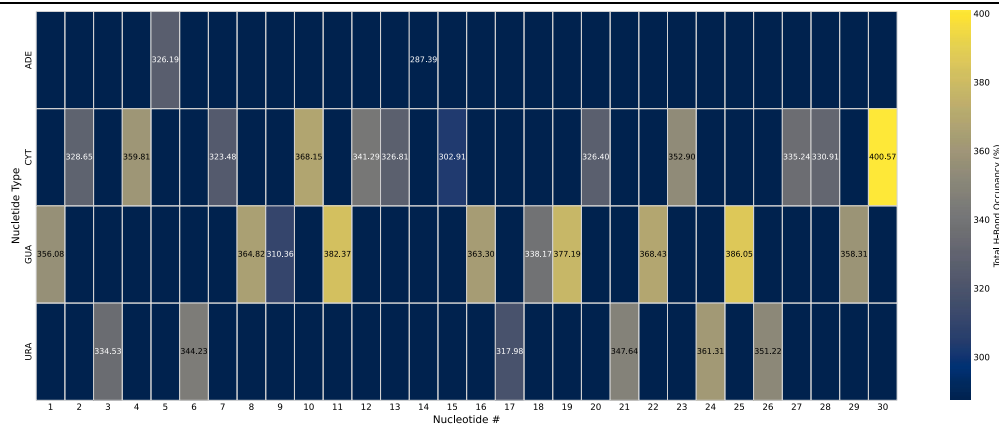
**Figura A. 6.** Mapa de calor de tiempos de residencia de los enlaces de hidrógeno ARN-agua del sistema de ARN en fase acuosa



2

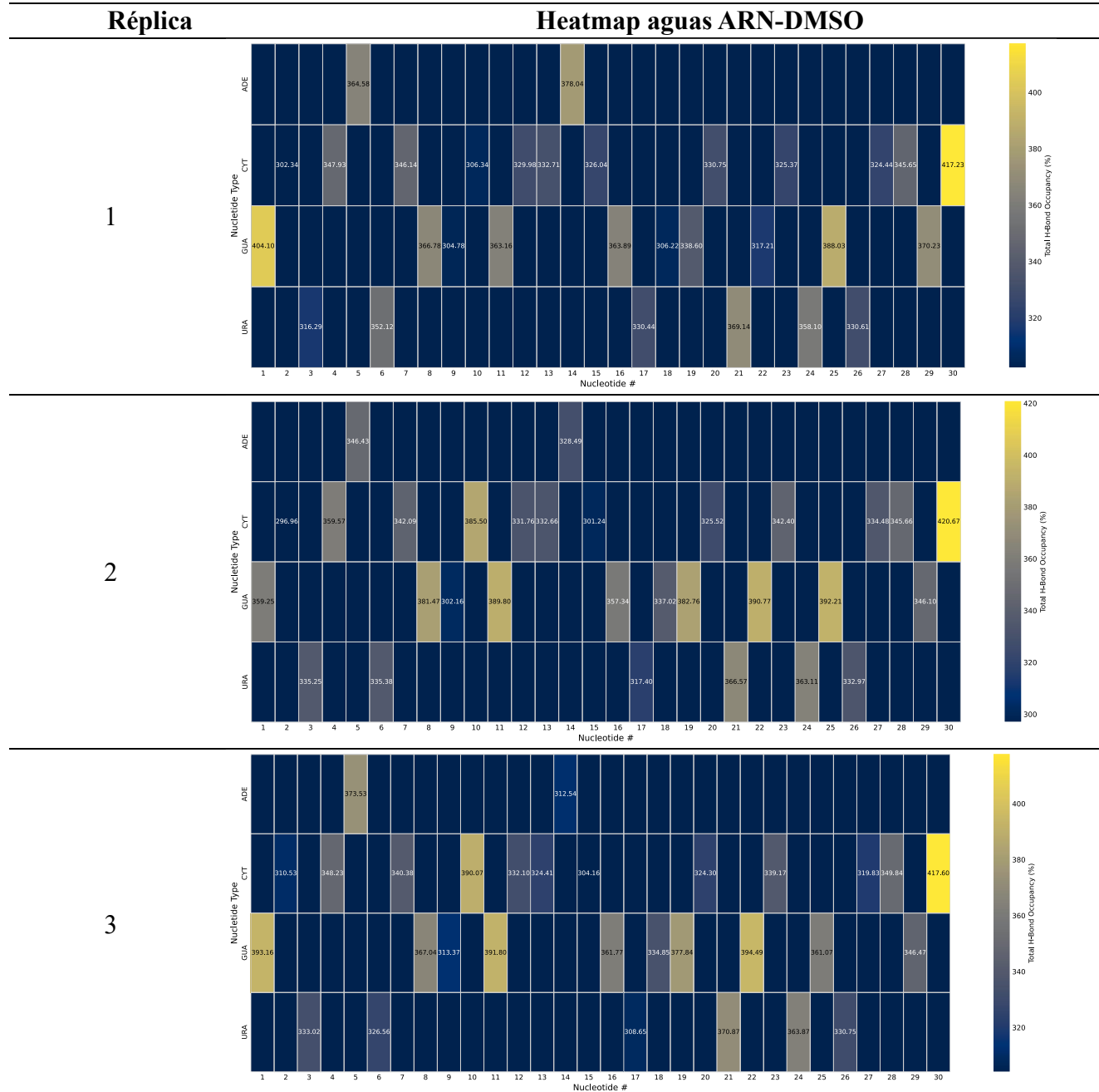


3



Los altos tiempos de residencia de los enlaces de hidrógeno ARN-agua se debe a que el agua tiene tiempos de residencia muy altos y a que más varias moléculas pueden formar enlaces de hidrógeno con cada nucleótido a la vez, por lo que al sumar los tiempos de residencia con cada nucleótido se obtienen valores mayores al 100%.

**Figura A. 7.** Mapa de calor de tiempos de residencia de los enlaces de hidrógeno ARN-agua del sistema de ARN con DMSO al 5% v/v



No se observó una diferencia significativa entre los tiempos de residencia de los enlaces de hidrógeno entre los nucleótidos y el agua entre los sistemas de ARN en fase acuosa y ARN con DMSO al 5%v/v. Esto indica que no es posible evidenciar un efecto del DMSO sobre la estructura de las aguas coordinadas a la superficie del fragmento de ARN mediante la metodología empleada.