



**ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**

**FACULTAD DE CIENCIAS**

**ESCUELA DE FÍSICA Y MATEMÁTICA**

**CARRERA DE BIOFÍSICA**

**“EVALUACIÓN DEL CAMPO DE FUERZA CHARMM27  
IMPLEMENTADO EN NAMD PARA LA SIMULACIÓN DE ADN G-  
CUÁDRUPLE TETRAMOLECULARES Y PARALELOS”**

TRABAJO DE TITULACIÓN

Previo a la obtención del título de:

**BIOFÍSICA**

**AUTORA:** Daniela Alejandra Barragán Rivera

**TUTOR:** Robert Cazar Ph. D.

Riobamba–Ecuador

2015

**ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**  
**FACULTAD DE CIENCIAS**  
**ESCUELA DE FÍSICA Y MATEMÁTICA**

El Tribunal del Trabajo de Titulación certifica que: El trabajo de investigación: **EVALUACIÓN DEL CAMPO DE FUERZA CHARMM27 IMPLEMENTADO EN NAMD PARA LA SIMULACIÓN DE ADN G-CUÁDRUPLE TETRAMOLECULARES Y PARALELOS**, de responsabilidad de la señorita Daniela Alejandra Barragán Rivera, ha sido prolijamente revisado por los Miembros del Tribunal del Trabajo de Titulación, quedando autorizada su presentación.

FIRMA

FECHA

Dr. Robert Cazar

\_\_\_\_\_

\_\_\_\_\_

**DIRECTOR DEL TRABAJO  
DE TITULACIÓN**

Dr. Dennis Cazar

\_\_\_\_\_

\_\_\_\_\_

**MIEMBRO DEL TRIBUNAL**

Yo, Daniela Alejandra Barragán Rivera, soy responsable de las ideas, doctrinas y resultados expuestos en esta Tesis y el patrimonio intelectual de la Tesis de Grado pertenece a la Escuela Superior Politécnica De Chimborazo.

---

DANIELA ALEJANDRA BARRAGÁN RIVERA

## **DEDICATORIA**

Este proyecto va dedicado a personas extraordinarias que aportaron en mi vida.

A mis queridos padres Gustavo Barragán y Norma Rivera, quienes han contribuido con mi crecimiento como persona líder, llena de valores, principios para conseguir mis objetivos. Para mí es un privilegio dedicarles a ustedes este trabajo y darles las gracias por darme tanto de su parte.

A mis hermanos María Isabel, Gustavo y Paulina quienes son mis amigos eternos y sinceros que me brindan el amor, compromiso y alegría día a día. Su ejemplo de superación, perseverancia, entrega y responsabilidad fueron mis principales claves para continuar con mis propósitos. Gracias por sus consejos.

A mis sobrinos Martín y Anahí quienes han sido mi motivación y felicidad. Ustedes me llenan de amor todos los días y me enseñan a vivir un día a la vez.

A mis cuñados Vanesa y Sebastián quienes siempre me motivaron a seguir adelante, gracias por sus consejos, reconozco su sencillez y cordialidad hacia mí.

Este logro es nuestro, mi esfuerzo es inspirado en ustedes. Los amo.

Daniela

## AGRADECIMIENTO

A Dios porque cada día llena de bendiciones mi vida y me da la oportunidad de crecer profesionalmente.

A mis padres, ellos son mis mayores promotores durante este proceso, por confiar en mí y por apoyarme en cada una de mis decisiones.

A mis hermanos, por estar siempre conmigo en todo momento y estar pendiente durante esta etapa.

A la Escuela Superior Politécnica de Chimborazo por haberme permitido formarme en ella y a cada docente que fueron parte de mi formación académica.

Al Grupo de Química Teórica Computacional, quienes contribuyeron para el desarrollo de este trabajo.

A Miguel Ángel Méndez, colaborador permanente de la Tesis, gracias por instruirme con excelencia y disposición.

A Robert Cazar por su valiosa colaboración y asesoramiento en la dirección de la presente tesis.

A Dennis Cazar por la confianza y el aporte brindado en la elaboración del trabajo de investigación.

A mis amigos quienes me alentaron y me entregaron su apoyo incondicional.

Daniela

## TABLA DE CONTENIDO

HOJA DE CERTIFICACION.....	ii
PÁGINA DE RESPONSABILIDAD.....	iii
DEDICATORIA .....	iv
AGRADECIMIENTO.....	v
TABLA DE CONTENIDO.....	vi
ÍNDICE DE TABLAS .....	ix
ÍNDICE DE ILUSTRACIONES.....	x
ÍNDICE DE ABREVIATURAS .....	xii
RESUMEN.....	xiii
SUMMARY .....	xiv
INTRODUCCIÓN .....	1
CAPÍTULO I.....	5
MARCO TEÓRICO REFERENCIAL.....	5
1. SIMULACIÓN MOLECULAR DE ÁCIDOS NUCLEICOS .....	5
1.1. Dinámica Molecular.....	5
1.1.1. Campo de Fuerzas .....	8
1.1.1.1 Términos Enlazantes .....	9
1.1.1.2 Términos no Enlazantes .....	10
1.1.2. Ensamblés.....	10
1.2. Estructura y Dinámica de los Ácidos Nucleicos .....	11
1.2.1. Estructura Primaria.....	11
1.2.2. Estructura Secundaria.....	12
1.2.2.1. Esqueleto Azúcar Fosfato y enlace Glucosídico .....	12
1.2.2.2. Empaquetamiento del Azúcar .....	13
1.2.2.3. Fuerzas estabilizantes: Puentes de hidrógeno y apilamiento.....	13

1.2.2.4.	Agua e Hidratación de ADN .....	14
1.2.3.	Estructuras canónicas de los Ácidos Nucleicos.....	14
1.2.4.	Estructuras no canónicas de los Ácidos Nucleicos.....	15
1.2.4.1.	G-Cuádruple .....	16
1.2.4.1.1.	Estructura y Topología .....	16
1.2.4.1.2.	Formación de ADN G-Cuádruples en sistemas biológicos.....	18
1.2.4.1.3.	Factores de Estabilidad.....	19
1.2.4.1.4.	Lugar y Función .....	20
1.2.4.1.5.	Aplicación .....	21
1.3.	Técnicas Experimentales.....	21
1.3.1.	Cristalografía de Rayos X .....	21
1.3.1.1.	Principio Físico .....	21
1.3.1.2.	Cristalografía ADN .....	22
1.3.2.	Resonancia Magnética Nuclear .....	23
1.3.2.1.	Principios Generales.....	23
CAPÍTULO II .....		26
MARCO METODOLÓGICO .....		26
2.1.	Material .....	26
2.1.1.	Recursos Computacionales .....	26
2.1.2.	Programas.....	27
2.2.	Métodos.....	28
2.2.1.	Modelado Molecular .....	28
2.2.2.	Análisis de trayectoria.....	30
2.2.3.	Análisis de las conformaciones estructurales de estructuras G-Cuádruples.....	32

CAPÍTULO III .....	41
MARCO DE RESULTADOS, DISCUSIÓN Y ANÁLISIS RESULTADOS .....	41
3.1. Preparación de Modelos a simular .....	41
3.2. Análisis de la Preparación del modelo .....	43
3.3. Conformaciones Estructurales.....	47
3.3.1. Conformaciones Estructurales de la estructura 1NP9 .....	47
3.3.2. Conformaciones Estructurales del Monómero 2GW0 .....	51
3.4. Análisis de las distancias de los puentes de Hidrógeno .....	54
3.4.1. Mapas de Calor de la estructura 1NP9 .....	55
3.4.2. Mapas de Calor de la estructura 2GW0.....	57
3.5. Comparación de los resultados de simulación versus los datos experimentales .....	58
CONCLUSIONES .....	61
RECOMENDACIONES .....	63
BIBLIOGRAFÍA	
ANEXOS	

## ÍNDICE DE TABLAS

<b>Tabla 1-1</b>	Resumen de parámetros estructurales de Estructuras Canónicas.....	14
<b>Tabla 2-1</b>	Radio Iónico de cationes estabilizantes de G-Cuádruples.....	20
<b>Tabla 1-2</b>	Características de las estructuras de ADN G-cuádruple estudiadas.....	29
<b>Tabla 2-2</b>	Histograma de Conformaciones seleccionadas por RMSD para solo los residuos de Guaninas del 1NP9 G-Cuádruple Tetramolecular Paralelo.....	34
<b>Tabla 3-2</b>	Histograma de Conformaciones seleccionadas por RMSD para el G-Cuádruple completo del 1NP9 G-Cuádruple Tetramolecular Paralelo.....	35
<b>Tabla 4-2</b>	Histograma de Conformaciones seleccionadas por Radio de Giro para solo los residuos de Guanina del Monómero 2GW0 G-Cuádruple Tetramolecular.....	37
<b>Tabla 5-2</b>	Histograma de Conformaciones seleccionadas por Radio de Giro para el G-Cuádruple completo del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo.....	38
<b>Tabla 1-3</b>	Conformaciones Estables y de Transición de la estructura 1NP9 G-Cuádruple Tetramolecular Paralelo.....	47
<b>Tabla 2-3</b>	Conformaciones Estables y de Transición de la estructura 2GW0 G-Cuádruple Tetramolecular Paralelo.....	51
<b>Tabla 3-3</b>	Cálculo de distancias de Estructuras Experimentales.....	59
<b>Tabla 4-3</b>	Comparación de distancias de los pisos de las G-Tétradas del 1NP9.....	60
<b>Tabla 5-3</b>	Comparación de distancias de los pisos de las G-Tétradas del Monómero 2GW0.....	60

## ÍNDICE DE ILUSTRACIONES

<b>Figura 1-1</b>	Esquema de una simulación de dinámica molecular.....	7
<b>Figura 2-1</b>	Condiciones Periódicas de Contorno.....	8
<b>Figura 3-1</b>	Purinas y Pirimidinas.....	12
<b>Figura 4-1</b>	Clasificación de las Estructuras Canónicas.....	15
<b>Figura 5-1</b>	Esquema de formas G-cuádruple.....	17
<b>Figura 6-1</b>	Tipo de elementos que unen las cadenas de G-Cuádruples.....	17
<b>Figura 7-1</b>	Diagrama Esquemático de Bucles para Cuádruples Bimoleculares.....	18
<b>Figura 8-1</b>	Ejemplo de apilamiento de pares de bases.....	19
<b>Figura 9-1</b>	Cuarteto de guaninas conectadas con presencia del ion K <sup>+</sup> en el canal central.....	20
<b>Figura 10-1</b>	Difracción de Rayos X.....	22
<b>Figura 1-2</b>	Proteína Tubulina en representación de New Cartoon.....	27
<b>Figura 2-2</b>	Protocolo de modelado molecular.....	28
<b>Figura 3-2</b>	G-cuádruples paralelos tetramoleculares.....	29
<b>Figura 4-2</b>	Preparación del sistema.....	29
<b>Figura 5-2</b>	Esquema del uso de las herramientas de GROMACS.....	31
<b>Figura 6-2</b>	Desviación estándar de la trayectoria de Dinámica Molecular durante 20 ns.....	33
<b>Figura 7-2</b>	Esquema de Clasificación Algoritmo k-NN.....	40
<b>Figura 8-2</b>	Pasos para realizar el Algoritmo de Clasificación k-NN.....	40
<b>Figura 1-3</b>	Resultado de solvatación.....	41
<b>Figura 2-3</b>	Resultado de la neutralización.....	42
<b>Figura 3-3</b>	Esquema para generar un archivo .psf.....	42
<b>Figura 4-3</b>	Disminución de la Energía del sistema ADN G-Cuádruple 1NP9.....	43
<b>Figura 5-3</b>	Disminución de la Energía del sistema ADN G-Cuádruple 2GW0.....	44
<b>Figura 6-3</b>	Temperatura de Termalización del sistema ADN G-Cuádruple 1NP9.....	44
<b>Figura 7-3</b>	Energía de Termalización del sistema ADN G-Cuádruple 1NP9.....	45
<b>Figura 8-3</b>	Temperatura de Termalización del sistema ADN G-Cuádruple del Monómero 2GW0.....	45
<b>Figura 9-3</b>	Energía de Termalización del sistema ADN G-Cuádruple del Monómero 2GW0.....	45

<b>Figura 10-3</b>	Presión de Termalización del sistema ADN G-Cuádruple 1NP9.....	46
	Presión de Termalización del sistema ADN G-Cuádruple del	
<b>Figura 11-3</b>	Monómero	46
	2GW0.....	
	Histograma de Clasificación de Conformaciones seleccionadas por	
<b>Figura 12-3</b>	RMSD para solo de los residuos de Guaninas del 1NP9 G-Cuádruple	48
	Tetramolecular Paralelo.....	
	Histograma de Clasificación de Conformaciones seleccionadas por	
<b>Figura 13-3</b>	RMSD para el G-Cuádruple completo de la estructura 1NP9 G-	
	Cuádruple Tetramolecular Paralelo.....	49
	Histograma de Clasificación de Conformaciones seleccionadas por	
<b>Figura 14-3</b>	Radio de Giro para solo los residuos de Guanina del Monómero 2GW0	52
	G-Cuádruple Tetramolecular Paralelo.....	
	Histograma de Clasificación de Conformaciones seleccionadas por	
<b>Figura 15-3</b>	Radio de Giro para el G-Cuádruple completo del Monómero 2GW0 G-	
	Cuádruple Tetramolecular Paralelo.....	53
<b>Figura 16-3</b>	Esquema de enlaces de Hidrógeno del G-Cuarteto de los sistemas.....	54
<b>Figura 17-3</b>	Mapas de Calor de la estructura 1NP9 G-Cuádruple.....	55
<b>Figura 18-3</b>	Mapas de Calor del Monómero 2GW0 G-Cuádruple.....	57
<b>Figura 19-3</b>	Esquema de medición distancia entre cuartetos de residuos de guanina..	59

## ÍNDICE DE ABREVIATURAS

<b>Å</b>	Angstrom
<b>A</b>	Adenina
<b>AND</b>	Ácido Desoxirribonucleico
<b>AMBER</b>	Del inglés Assisted Model Building with Energy Refinement
<b>C</b>	Citosina
<b>CHARMM</b>	Del inglés Chemistry at Harvard Macromolecular Mechanics
<b>Cl</b>	Cloro
<b>CSV</b>	Del inglés comma-separated values
<b>DM</b>	Dinámica Molecular
<b>fs</b>	Femtosegundos
<b>G</b>	Guanina
<b>GROMACS</b>	Del inglés GRONingen MACHine for Chemical Simulations
<b>K</b>	Potasio
<b>°K</b>	Kelvin
<b>Kcal/mol</b>	Kilo calorías/ mol
<b>k-NN</b>	Del inglés k- nearest neighbors
<b>Na</b>	Sodio
<b>NAMD</b>	Del inglés Nanoscale Molecular Dynamics
<b>OPLS</b>	Del inglés Optimized Potentials for Liquid Simulations
<b>PBC</b>	Del inglés Periodic Boundary Conditions
<b>PDB</b>	Del inglés Protein Data Bank
<b>ps</b>	Picosegundos
<b>PSF</b>	Del inglés Protein Structure File
<b>RMN</b>	Resonancia Magnética Nuclear
<b>RMSD</b>	Del inglés Root Media Standard Deviation
<b>T</b>	Temperatura
<b>TIP3</b>	Del inglés transferable intermolecular potential 3
<b>VMD</b>	Del inglés Visual Molecular Dynamics

## RESUMEN

El presente trabajo de investigación se realizó en el Laboratorio de Química Computacional y Teórica de la Universidad San Francisco de Quito en la ciudad de Quito con el propósito de estudiar el Ácido desoxirribonucleico (ADN) G-Cuádruple que es de gran interés debido a la participación en procesos celulares. La conformación molecular de los G-Cuádruples se ha obtenido experimentalmente vía Resonancia Magnética Nuclear (RMN) o Cristalografía de rayos X. Sin embargo, se desconoce los detalles del proceso de formación de estas estructuras a partir de segmentos de ADN de doble cadena. Con este fin, se da a conocer la evaluación del campo de fuerza CHARMM27 para estas estructuras mediante simulaciones utilizando el software Dinámica Molecular a Nanoescala (NAMD). Sé evaluó dos estructuras de ADN G-Cuádruple de similar secuencia, para la preparación del modelo se utilizó el paquete de Visualización de Dinámica Molecular (VMD) y se analizó las conformaciones estructurales a lo largo de una trayectoria de 200 nanosegundos con el uso de las herramientas de GROMACS Tools y Aprendizaje de Máquina. Los resultados obtenidos para las configuraciones de ambos sistemas fueron comparados con la información de datos experimentales y se encontró que las simulaciones de Dinámica Molecular son capaces de reproducir los valores experimentales con un margen de error menor al 10%. En conjunto, estos resultados proporcionan pruebas sustanciales de la formación de estructuras G-Cuádruple en condiciones fisiológicas (25° Centígrados en presencia de cloruro de sodio NaCl, 1 atmósfera). En este estudio para obtener resultados más robustos se recomienda a los investigadores tomar en cuenta análisis como Función de Distribución Radial, ángulos del esqueleto del ADN, Análisis de Componente Principal, entre otros.

### Palabras Clave:

<UNIVERSIDAD SAN FRANCISCO DE QUITO > <ÁCIDO DESOXIRRIBONUCLEICO  
[ADN]> <G-CUÁDRUPLE> <CAMPO DE FUERZA> < DINÁMICA MOLECULAR A  
NANOESCALA [NAMD]>

## SUMMARY

The current research was carried out at Computational and Theoretical Chemistry Laboratory of San Francisco de Quito University, located in Quito city in order to study the G-Quadruplex deoxyribonucleic acid (DNA) which is very important due to its participation in the cellular processes. The G-Quadruplex molecular conformation has been obtained experimentally through a Nuclear Magnetic Resonance (NMR) also called X-ray Crystallography. However, the details of the formation process in these structures is unknown from the double-stranded DNA segments. The evaluation of CHARMM27 force field for these structures by means of simulations using Nanoscale Molecular Dynamics (NAMD) is disclosed. Two same-sequence G-Quadruplex DNA structures were evaluated, to prepare the model, it was necessary to use Visual Molecular Dynamics program (VMD), the structural conformations in a trajectory of 200 nanoseconds were also analyzed with the use of GROMACS Tools and Machine Learning. The results obtained for the two system settings were compared with the experimental data information and it was shown that the molecular dynamics simulations are able to reproduce the experimental values with an error range less than 10%. All together, these results provide substantial evidence for G-Quadruplex structure formation in physiological conditions. It means: 25° grades centigrade with sodium Chloride NaCl, 1 atmosphere. According to this study, to obtain better results it is necessary that the researchers take into account the analyses such as: Radial Distribution Function, DNA backbone angles, Principal Component Analysis, among others.

### **Key words:**

<SAN FRANCISCO DE QUITO UNIVERSITY> <DEOXYRIBONUCLEIC ACID [DNA]> <G-QUADRUPLEX> <FORCE FIELD> <NANOSCALE MOLECULAR DYNAMICS [NAMD]>

## INTRODUCCIÓN

La Biología Molecular y la Genética han sido resultado del desarrollo de la Biología, ambas han adquirido gran importancia debido al desciframiento del genoma humano y las bases moleculares de las enfermedades cuyos mecanismos bioquímicos están siendo descubiertos día a día. Así, los nuevos conocimientos nos permiten comprender las estructuras complejas como los ácidos desoxirribonucleicos (ADN) que cumplen un papel en la duplicación celular.

Estos avances actuales sigue siendo una ciencia en elaboración como es el caso del ADN, aun cuando ya fue descubierto por Friedrich Miescher en 1869, quien realizó un experimento de la composición química en el esperma del salmón. (DAHM, 2005, pp. 274-288) En el transcurso de los años James Watson and Francis Crick en 1953 propusieron el modelo de ADN de doble hélice. Estos descubrimientos dieron el impulso a un creciente interés en otras formas de ADN. (VASUDEVAN, et al., 2011, pp. 2-65)

Recientemente, se encontraron otras estructuras más complejas y menos conocidas como de hélice cuádruple de ADN (TRISTÁN, 2013, <http://biotecmur.com>) que dan lugar a estructuras intermoleculares como intramoleculares según la asociación de sus cadenas. (GARGALLO, 2008, <http://www.ub.edu>) Estas estructuras, denominadas como G-cuarteto se observaron mediante biomarcadores fluorescentes en concentraciones milimolares de guanina. (HUPPERT, 2007, pp. 1471-2962) (NEIDLE & BALASUBRAMANIAN, 2007)

El estudio de las estructuras cuádruples del ADN tiene relevancia por encontrarse en los extremos teloméricos de los cromosomas y su papel fundamental en los procesos celulares como recombinación y transcripción. (STEGLE, et al., 2009, p. 374-382) (PARKINSON, et al., 2002, pp. 876-880) Por otra parte la diversidad estructural y función biológica del ADN G-Cuádruple ha sido estudiada mediante técnicas de Cristalografía de Rayos X y Resonancia Magnética Nuclear.

Actualmente los métodos teóricos computacionales facilitan el análisis de las secuencias biológicas, la búsqueda de genes y la predicción de estructuras de proteínas a partir de su secuencia de aminoácidos, entre otros. (FADRINA, et al., 2014, pp. 227-242)

Uno de los métodos teóricos, de la química computacional, para el estudio de biomoléculas es la Dinámica Molecular que consiste principalmente en describir el movimiento atómico a través del

tiempo de un sistema de partículas (en este caso átomos), por medio de campos de fuerza y así permiten una descripción confiable para estructuras como los ácidos nucleicos.

El campo de fuerza CHARMM27 presenta algunos parámetros diferentes con respecto al campo de fuerza CHARMM22; parámetros que describen de mejor manera el equilibrio de tipos de ADN A y B y la influencia del medio ambiente. (VANOMMESLAEGHE, et al., 2010, pp. 671-690)

El campo de fuerza permite describir matemáticamente la interacción entre los átomos del sistema y por tanto se pueden realizar predicciones cuantitativas, evaluar y validar mediante comparaciones entre datos obtenidos de la simulación versus los datos experimentales.

En este trabajo se evalúa el campo de fuerza CHARMM27 implementado en el software NAMD en su desempeño en la simulación de diferentes estructuras de ADN G-Cuádruple con la finalidad de analizar propiedades estructurales y dinámicas relacionadas con la función biológica.

En el modelamiento de estas estructuras se analizó las conformaciones estructurales estables y de transición presentes a lo largo de una trayectoria de 200ns con el uso de las herramientas de GROMACS Tools y Aprendizaje de Máquina. Además, se compararon los resultados obtenidos entre las configuraciones de ambos sistemas y con la información estructural de literatura disponible.

## **Antecedentes**

El campo de la biofísica computacional desarrolla métodos teóricos y algoritmos que son aplicados para el estudio de la estructura y dinámica de los ácidos nucleicos en su variedad de funciones biológicas. (BEVERIDGE & RAVISHANKER, 1994, pp. 246-255) Estos métodos computacionales han ido mejorando continuamente con el propósito de resolver las limitaciones experimentales y cumplen un rol importante en el estudio de las conformaciones e interacciones de las biomoléculas. (LAVERY, 1995, pp. 57-82)

Las simulaciones computacionales de ADN permiten la comprensión de la dinámica y estructura de las conformaciones, complementa los experimentos y hacen predicciones comprobables, por lo que su uso se ha incrementado dramáticamente esta última década. (MAFFEO, et al., 2014, p. 36)

Hoy en día, un número considerable de investigadores se encuentra realizando estudios con campos de fuerza en simulaciones moleculares de varios nanosegundos aplicando un conjunto de parámetros de campo de fuerza, como los paquetes disponibles AMBER, CHARMM27, y

GROMOS dando resultados razonables para la simulación de las conformaciones de ADN. (HAIDER & NEIDLE, 2010, pp. 17-37)

Por ejemplo, el estudio de estructuras como ADN G-Cuádruples mediante campos de fuerza han ayudado a explorar las principales características conformacionales (VIRGILIO, et al., 2012, pp. 461-475), su polimorfismo (LE, et al., 2014, pp. 5390-5405) y analizar los efectos de ciertas moléculas en su estructura, estabilidad y energía cinética, entre otros. (LE, et al., 2014, pp. 5390-5405) Sin embargo la implementación de CHARMM para la simulación de estas estructuras no convencionales de ácidos nucleicos aún no se ha realizado ampliamente.

Por otra parte, CHARMM27 es un modelo de campo de fuerza altamente optimizado y adaptable debido a procesos de optimización constante. Este campo de fuerza describe las formas A, B y Z del ADN en fase condensada y sistemas heterogéneos; también es compatible con proteínas y lípidos. (MACKERELL & NILSSON, 2006, pp. 73-94)

Algunos de los desafíos que enfrenta CHARMM27 son la problemática de reproducción precisa de los datos experimentales sobre bases y molécula de azúcar debido a la flexibilidad. (MACKERELL, et al., 2001, pp. 257-265); Además, es difícil llevar a cabo simulaciones precisas de algunos detalles estructurales; por ejemplo, grupos de hélice aparecen inestables, los bucles diagonales se pierden durante la simulación. (FADRNA, et al., 2009) Por lo tanto, es necesaria la optimización continua, pero sigue siendo una de las herramientas más utilizadas en simulaciones de ADN. (MACKERELL & NILSSON, 2006, pp. 73-94) (FADRNA, et al., 2009, pp. 2514-2530) (ZHU, et al., 2012, pp. 167-185)

## **Justificación**

Uno de los métodos para el estudio de biomoléculas, es la Dinámica Molecular que se caracteriza por describir los movimientos moleculares a través del tiempo. Esta técnica es una aplicación físico-química en estructuras como ácidos nucleicos que permite el análisis minucioso de la energía y dinámica, propiedades directamente relacionadas con la función biológica, así como también el desarrollo de compuestos que puedan modificar su estabilidad.

La Dinámica Molecular se basa en la determinación de parámetros para construir campos de fuerza de ácidos nucleicos y en el desarrollo de algoritmos, así como el uso de métodos teóricos más complejos para la evaluación de los efectos electrostáticos, que son importantes en los sistemas cargados. La implementación de nuevos métodos computacionales facilita la exploración de una variedad de sistemas biomoleculares que contienen disposiciones complejas como en el caso de ácidos nucleicos tales como ADN G-Cuádruples, dando lugar a resultados cada vez más confiables. (HAIDER & NEIDLE, 2010, pp. 17-37)

Las enfermedades genéticas en el ser humano y el cáncer están estrechamente relacionadas con el estudio del ADN G-cuádruple por su identificación en los telómeros, promotores oncogénicos y secuencias repetitivas. Al conocer la descripción completa de la estructura y dinámica del G-cuádruple adquirimos ideas fundamentales de la causa de estas patologías.

Por otra parte, algunas secuencias de oligómeros artificiales ricos en G, capaces de formar G-cuádruple, poseen actividades biológicas interesantes, tales como actividades anticoagulantes, anticancerígenos y anti-VIH. Las moléculas de G-cuádruple también tienen aplicaciones potenciales en diversos campos tales como la química y la nanotecnología. (MICHAEL, et al., 2012, pp. 11-24)

Teniendo en cuenta la importancia biológica de las estructuras del G-Cuádruple se han realizado estudios experimentales mediante Cristalografía y RMN pero no se ha logrado estudiar en detalle las diferentes conformaciones debido a la presencia de altas concentraciones de moléculas que conduce a una gran diversidad topológica; por ello la simulación molecular es esencial para estudiar estas estructuras.

## **Objetivos**

### **Objetivo General**

Evaluar el campo de fuerza CHARMM27 implementado en NAMD para la simulación de ADN G-cuádruple tetramoleculares y paralelos

### **Objetivos Específicos**

- Identificar conformaciones estables del ADN G-Cuádruple mediante la simulación de Dinámica Molecular.
- Clasificar las conformaciones estructurales presentes en estructuras de G-Cuádruple.
- Analizar la longitud de los puentes de hidrógeno Hoogsteen presentes en G-Cuádruple.
- Comparar los resultados de simulación de estructuras de ADN G-Cuádruple Tetramolecular Paralelo con los datos experimentales obtenidos por técnicas de RMN y Cristalografía de Rayos X.

# CAPÍTULO I

## MARCO TEÓRICO REFERENCIAL

### 1. SIMULACIÓN MOLECULAR DE ÁCIDOS NUCLEICOS

La simulación Molecular consiste en el uso de métodos computacionales que permiten modelar y predecir el comportamiento de las moléculas. Los métodos se basan en leyes físicas, químicas y bioquímicas de sistemas simples hasta sistemas complejos. La simulación molecular de sistemas se realiza mediante métodos clásicos o cuánticos, los cuales proporcionan la información espacial de la estructura, dinámica y termodinámica.

Los métodos computacionales tienen una gran ventaja sobre los métodos experimentales debido a la posibilidad de controlar rigurosamente parámetros y variables como el número de partículas, la temperatura, la presión, volumen y otros; produciendo experimentos que pueden ser visualizados y analizados a través de una interfaz gráfica.

El modelamiento molecular de ácidos nucleicos permite generar diferentes configuraciones estructurales de éstos, evaluar y representar interacciones atómicas entre pares de bases y simular posibles eventos a nivel cromosómico. En sistemas de interés biológico se utiliza ampliamente la dinámica molecular, ya que para sistemas biológicos de gran tamaño (centenas de miles a millones de átomos) ha entregado buenos resultados. (ZHAO, et al., 2013, pp. 643-646) (VANDIVORT, et al., 2011, <https://web.stanford.edu>)

#### 1.1. Dinámica Molecular

La dinámica molecular (DM) es un método que calcula el comportamiento de un sistema i.e. los promedios temporales de las propiedades del sistema. Este método fue utilizado por primera vez en líquidos por Alder y Wainwright en la década de los 50, siendo una herramienta teórica para el estudio de biomoléculas y sirve de ayuda para comprender las diferentes configuraciones adoptadas por el sistema. (AGUIRRE, 2009, <http://hera.ugr.es>)

Para estudiar las configuraciones del sistema molecular se debe resolver la ecuación de Movimiento de Newton. Las posiciones y velocidades de los sistemas se pueden describir a través de esta ecuación y si se conoce las posiciones y velocidades del sistema se puede predecir el estado del sistema en cualquier tiempo, por lo tanto el método es determinista.

### Principio Físico

En un sistema de  $N$  partículas, la energía potencial del sistema es una función de las coordenadas  $\vec{r}_i$ ,  $i= 1, \dots, N$ , de las partículas. La ecuación de Newton está descrita de la siguiente manera

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i \quad (1)$$

Donde  $m_i$  es la masa de la partícula  $i$  y  $t$  representa el tiempo. La aceleración  $a_i$  de la partícula  $i$  está definida como

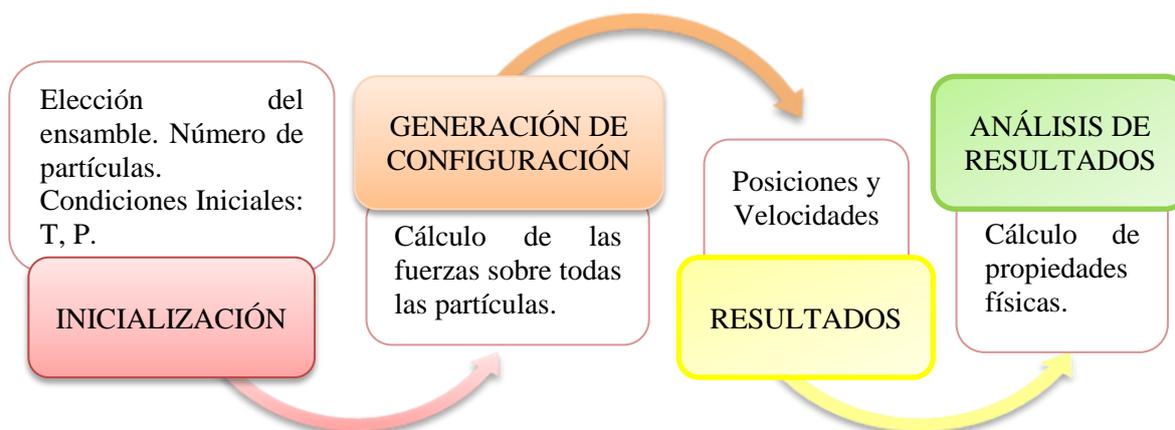
$$a_i = \frac{d^2 \vec{r}_i}{dt^2} \quad (2)$$

El resultado de la integración de la ecuación de Newton nos dará como resultado una trayectoria donde están calculadas las posiciones y velocidades de las partículas en función del tiempo. Existen varios métodos que simplifican la resolución de la ecuación de movimiento de Newton, debido a que su integración es compleja por su múltiple dependencia de las coordenadas de cada átomo. (AGUIRRE, 2009, <http://hera.ugr.es>)

El algoritmo de Verlet usa las posiciones y aceleraciones al tiempo  $t$ , y las posiciones del paso previo,  $r(t-\delta t)$  para calcular las nuevas posiciones en  $t+\delta t$ ,  $r(t+\delta t)$ . Las velocidades no aparecen explícitamente en el algoritmo de Verlet. Estas se pueden calcular de varias formas. Una aproximación muy simple es el dividir la diferencia en las posiciones en tiempos  $t+\delta t$  y  $t-\delta t$  by  $2\delta t$ ,

$$v(t) = [r(t + \delta t) - r(t - \delta t)] / 2\delta t \quad (3)$$

El esquema de una simulación de dinámica molecular es:



**Figura 1-1.** Esquema de una simulación de dinámica molecular  
Realizado por: Daniela Barragán

### Función de potencial

En la dinámica de un sistema químico para conocer todas las fuerzas del sistema se debe calcular la energía potencial. Cuando se tiene un sistema multidimensional, la energía potencial se conoce como la superficie de energía potencial. La superficie de energía potencial generalmente se calcula utilizando la aproximación Born Oppenheimer que considera que los electrones son más ligeros que los núcleos y se ajustan a cualquier cambio en la configuración nuclear. (JAQUET, 2002, <http://www.tc.chemie.uni-siegen.de>)

La superficie de Energía Potencial actúa como una función de energía potencial nuclear que depende de las posiciones nucleares. Este método permite describir los procesos moleculares, sus propiedades topológicas en las geometrías de equilibrio, niveles energéticos y propiedades dinámicas. (VALDÉS, 2006, <http://fama.iff.csic.es>)

La energía potencial de un sistema de N átomos se expresa de la siguiente manera:

$$V = \sum_{i=1}^N u_1 \vec{r}_i + \sum_{i=1}^N \sum_{j>1}^N u_2(\vec{r}_i, \vec{r}_j) + \sum_{i=1}^N \sum_{j>1}^N \sum_{k>j>1}^N u_3(\vec{r}_i, \vec{r}_j, \vec{r}_k) + \dots \quad (4)$$

Donde el primer término  $u_1$  representa las interacciones de un cuerpo,  $u_2$  las interacciones entre dos cuerpos, los términos mayores serán las interacciones de muchos cuerpos. El costo computacional que implica obtener las fuerzas que ejercen sobre los átomos es alto y no es viable para biomoléculas. Sin embargo actualmente se aplica la combinación de métodos como la Dinámica Molecular y la teoría del funcional de densidad. (AGUIRRE, 2009, <http://hera.ugr.es>)

## Restricción de los grados de libertad

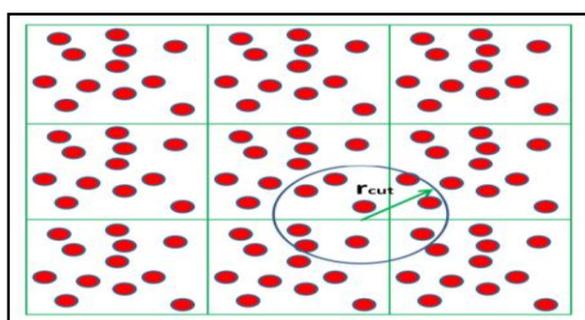
La dinámica molecular utiliza algoritmos de restricciones para enlaces, ángulos de torsión y ángulos para explorar un mayor número de conformaciones y evitar la distorsión aleatoria entre los enlaces y los ángulos diedros; por lo cual se desarrolló el método de SHAKE y el método RATTLE (FORESTER & SMITH, 1998, pp. 102-111)

## Condiciones Periódicas de Contorno

La simulación de un gran número de partículas implica el uso de condiciones periódicas de contorno (PBC) para disminuir el tamaño macroscópico del sistema dividiéndolo en múltiples copias pequeñas y minimizar los efectos superficiales de esta división.

La figura 2-1 muestra una replicación del sistema en 3 dimensiones cuando se aplica PBC. Cuando se aplica esta condición se introduce un nuevo concepto, radio de corte (AGUIRRE, 2009, <http://hera.ugr.es>), la propiedad de interés es calculada con precisión cuando las interacciones de las partículas se describen dentro del radio de corte y si las partículas se encuentran distantes al radio de corte se pierden las interacciones de van der Waals e interacciones electrostáticas.

Para la descripción de las interacciones a larga distancia en sistemas periódicos se usa el Método de Ewald. (LIANG, 2014, <http://www.diva-portal.org>)



**Figura 2-1.** Condiciones Periódicas de Contorno.

$r_{cut}$  es el radio de corte.

Fuente: (LIANG, 2014, <http://www.diva-portal.org>)

### 1.1.1. Campo de Fuerzas

Campo de fuerza es una función de energía potencial que describe las interacciones entre los átomos mediante un conjunto de términos que son: estiramiento, ángulo de enlace, flexión, rotaciones torsionales y diedros inadecuadas; los términos no enlazantes incluyen las interacciones

van der Waals y las interacciones electrostáticas. Por lo tanto la energía potencial de un sistema atómico se puede expresar como

$$U_{\text{potencial}} = U_{\text{enlazantes}} + U_{\text{no enlazantes}} \quad (5)$$

#### 1.1.1.1 Términos Enlazantes

**Estiramiento.-** este término describe el cambio de energía potencial causado por el cambio de enlaces covalentes entre dos átomos. El campo de fuerza de macromoléculas se representa con un término armónico.

$$\sum_{\text{str}} K_b (b - b_0)^2 \quad (6)$$

donde  $k_b$  es la constante de fuerza del enlace y  $b - b_0$  es la distancia desde equilibrio que el átomo se ha movido.

**Ángulo.-** este término describe la energía potencial causada por el cambio de ángulos formados entre enlaces covalentes vecinos

$$\sum_{\text{angle}} K_\theta (\theta - \theta_0)^2 \quad (7)$$

donde  $k_\theta$  que es la constante de fuerza del ángulo y  $\theta - \theta_0$  es el ángulo de equilibrio entre la 3 átomos enlazados.

**Ángulos de Torsión o Diedros.-** describe la rotación de los átomos alrededor de un enlace, es importante para describir la conformación de la molécula

$$\sum_{\text{dihedrals}} k_\phi [1 + \cos(n\phi - \delta)] \quad (8)$$

donde  $k_\chi$  es la constante de fuerza de diedro,  $n$  es la multiplicidad de la función,  $\chi$  es el ángulo diedro y  $\delta$  es el desplazamiento de fase.

**Diedros impropios.-** movimiento fuera del plano de átomos en una molécula planar

$$\sum_{\text{impropers}} K_w (w - w_0)^2 \quad (9)$$

donde  $k_w$  es la constante de fuerza y  $w - w_0$  es el de ángulo plano. Se definen diedros impropios cuando se quieren mantener ciertos grupos en un mismo plano o configuración determinada.

### 1.1.1.2 *Términos no Enlazantes*

La componente de Urey-Bradley (término cruzado para el ángulo de flexión usando 1,3 interacciones no enlazantes) donde  $k_U$  es la constante de fuerza y  $U$  es la distancia entre 1,3 átomos en el potencial armónico. Este término es característico del campo de fuerza CHARMM.

$$\sum_U K_{UB} (u - u_0)^2 \quad (10)$$

**Lennard-Jones.-** Las interacciones de van der Waals se describen por el potencial Lennard Jones donde  $\epsilon$  es la profundidad del pozo y  $r_{ij}$  es la distancia entre dos partículas

$$\sum_{\text{nonbond}} \epsilon \left[ \left( \frac{R_{\text{min}_{ij}}}{r_{ij}} \right)^{12} - \left( \frac{R_{\text{min}_{ij}}}{r_{ij}} \right)^6 \right] \quad (11)$$

En el potencial de Lennard-Jones, el término  $R_{\text{min}_{ij}}$  no es el mínimo del potencial, sino más bien donde el potencial de Lennard-Jones cruza el eje x (es decir donde el potencial de Lennard-Jones es cero).

**Coulomb.-** La interacción electrostática se produce por la distribución entre dos cargas puntuales  $Q_i$  y  $Q_j$  es calculada como

$$E_{\text{elec}} = \frac{Q_i Q_j}{4\pi\epsilon_0 r_{ij}} \quad (12)$$

donde  $\epsilon_0$  es la permitividad del vacío y  $r_{ij}$  es la distancia entre  $i$  y  $j$

### 1.1.2. *Ensamblés*

La teoría de ensamblés se puede conectar a la dinámica molecular con la termodinámica y así obtener propiedades de interés. Esta teoría se basa principalmente en la generación de los microestados accesibles por el sistema.

A través de la mecánica estadística se calculan los promedios de estos microestados y se relacionan con variables termodinámicas, dependiendo de las características del sistema. En el caso de sistemas biológicos, cada microestado representa una conformación  $i$  que tiene una energía  $E_i$ . (ROMERO, 2001, <http://bibing.us.es>)

**Ensamble Microcanónico.-** se utiliza en las simulaciones de dinámica molecular cuando el número total de átomos  $N$  y el volumen  $V$  se mantiene constante; por lo tanto la energía se conserva. (VESELY, 2005, <http://homepage.univie.ac.at>)

**Ensamble Canónico.-** consiste en mantener la temperatura y volumen constantes. Se emplea cuando se quiere explorar conformaciones de moléculas debido a que la energía es variable. Este ensamble entrega la energía de Helmholtz. (VESELY, 2005, <http://homepage.univie.ac.at>)

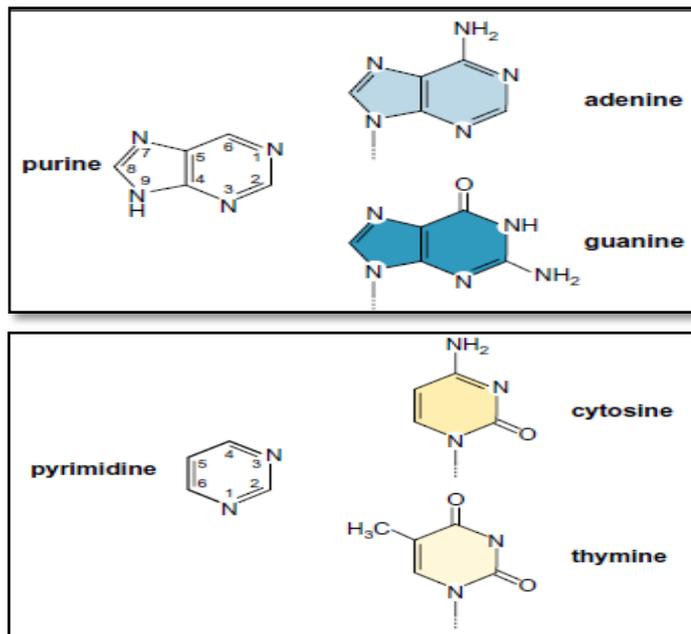
**Ensamble Isotérmico – Isobárico.-** se controla presión y temperatura constante, se utiliza para evaluar presión, densidad del sistema y optimizar volumen. Además, una característica importante de este ensamble, es que entrega la energía de Gibbs, la cual puede ser comparada con mayor facilidad con estudios energéticos experimentales. (CAI, 2011, <http://micro.stanford.edu>)

## 1.2. Estructura y Dinámica de los Ácidos Nucleicos

### 1.2.1. Estructura Primaria

La característica más importante del ADN es estar compuesta por dos cadenas de polinucleótidos, dando a lugar a una doble hélice. La columna vertebral de cada hebra de la hélice se compone de residuos de azúcar y fosfato. (HALL, 2002, <http://biology.kenyon.edu>) Cada nucleótido consta de una base nitrogenada ligada a un azúcar grupo fosfato. Las bases nitrogenadas que conforman los nucleótidos son moléculas aromáticas, planares, heterocíclicas, estructuralmente derivadas de una purina o pirimidina.

Las de purina son Adenina(A) y Guanina (G) y las bases de pirimidina Timina (T), citosina (C). Las bases de purina se enlazan con el Carbono (C1) de un azúcar furanósico a través de un Nitrógeno (N9). Pirimidinas se unen a través del átomo N1. Los nucleótidos están unidos a través de enlaces fosfodiéster 5'-3' que constituye el esqueleto de azúcar-fosfato. (HOFFORT, 2006, <http://biology.kenyon.edu>)



**Figura 3-1.** Purinas y Pirimidinas

Fuente: (HALL, 2002, <http://biology.kenyon.edu>)

## 1.2.2. Estructura Secundaria

### 1.2.2.1. Esqueleto Azúcar Fosfato y enlace Glucosídico

La flexibilidad conformacional de la unidad de nucleótido está limitada por una variedad de restricciones internas que limitan grandemente la libertad rotacional. La conformación de una molécula se describe generalmente por longitudes de enlace, ángulos de enlace y rotaciones de los grupos de átomos acerca de los enlaces por ángulos de torsión.

Hay siete ángulos de torsión que son importantes para describir la estructura secundaria de un nucleótido,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$  y  $\chi$ . Los primeros seis describen los ángulos de torsión del esqueleto azúcar-fosfato, el último ángulo de torsión,  $\chi$ , describe la rotación alrededor de los enlaces glucosídicos. El enlace glucosídico sólo se encuentra en una de las dos posiciones estéricamente permisibles: syn y anti.

En la conformación anti, la mayor parte de la base está apuntando lejos del azúcar, mientras que en la conformación syn, la mayor parte de la base está hacia el azúcar dando lugar a estrechos contactos interatómicos. La cantidad de impedimento estérico que sale en la conformación syn es afectada por la conformación del azúcar o el empaquetamiento del azúcar. (HOFFORT, 2006, <http://ecommons.usask.ca>)

### 1.2.2.2. *Empaquetamiento del Azúcar*

La disposición estructural del anillo glucosídico en los ácidos nucleicos se denomina empaquetamiento (*puckering*) que describe la flexibilidad del anillo ribosa, el cual es no plano. El azúcar puede ser empaquetado con 4 de los 5 átomos coplanares y el quinto átomo fuera del plano a una distancia de 0.5 Å.

Además el azúcar presenta conformaciones de tipo E (del inglés envelope) cuando uno de los carbonos queda fuera del plano a 0.5 Å y los demás átomos están dentro de un mismo plano; tipo T (del inglés twist) formado por los otros tres átomos donde dos átomos de carbono quedan desplazados fuera del plano y en lados opuestos. Si el carbono está dirigido en la misma cara del átomo C5' se denomina endo y si se ubican en la cara opuesta se denomina exo. (HOFFORT, 2006, <http://ecommons.usask.ca>)

### 1.2.2.3. *Fuerzas estabilizantes: Puentes de hidrógeno y apilamiento*

Las bases interactúan entre ellas cuando un átomo de hidrógeno forma un puente con un átomo electronegativo por efecto de las fuerzas electrostáticas. (HOFFORT, 2006, <http://ecommons.usask.ca>) Los puentes de hidrógeno se dan entre un átomo electronegativo denominado aceptor por poseer pares libres de electrones y un átomo dador el cual se encuentra unido a un hidrógeno ácido. En los nucleótidos, el grupo N-H es un buen dador y los nitrógenos del anillo son aceptores.

La interacción de apilamiento es un fenómeno complejo cuando las bases se apilan en forma vertical, se conoce también como interacción  $\pi$ - $\pi$  es perpendicular al plano de las bases y está estabilizado por fuerzas de dispersión London y efectos hidrofóbicos.

En solventes no polares, el enlace del hidrógeno es más prominente y se eliminan los efectos apilamiento base. En el agua, los efectos de base apilamiento son dominantes y las interacciones del enlace de hidrógeno son insignificantes debido a la competencia para atar los sitios con las moléculas de agua.

La descripción termodinámica de este tipo de interacciones ha sido analizada por diversas técnicas experimentales como osmometría, desnaturalización térmica, RMN, entre otras. Según los resultados obtenidos a medida que se reduce el tamaño del anillo existe una disminución progresiva de la estabilidad por lo tanto los apilamientos purina-purina son más estables mientras

que los apilamientos entre bases pirimidina-pirimidina son menos estables. (FRIEDMAN & HONIG, 1995, pp. 1528-1535)

En el caso de los G-Cuádruples tenemos secuencias ricas en purinas (guaninas) por tanto se espera una fuerte contribución por interacciones  $\pi$ - $\pi$ . Por otra parte, la energía de interacciones apilamiento base es afectada por interacciones de van der Waals, interacciones electrostáticas entre cargas atómicas parciales. (HOFFORT, 2006, <http://ecommons.usask.ca>)

#### 1.2.2.4. Agua e Hidratación de ADN

El agua es un activo estabilizador de estructuras secundarias y terciarias en macromoléculas. El agua ayuda a neutralizar repulsiones electrostáticas fosfato-fosfato. (HOFFORT, 2006, <http://ecommons.usask.ca>)

#### 1.2.3. Estructuras canónicas de los Ácidos Nucleicos

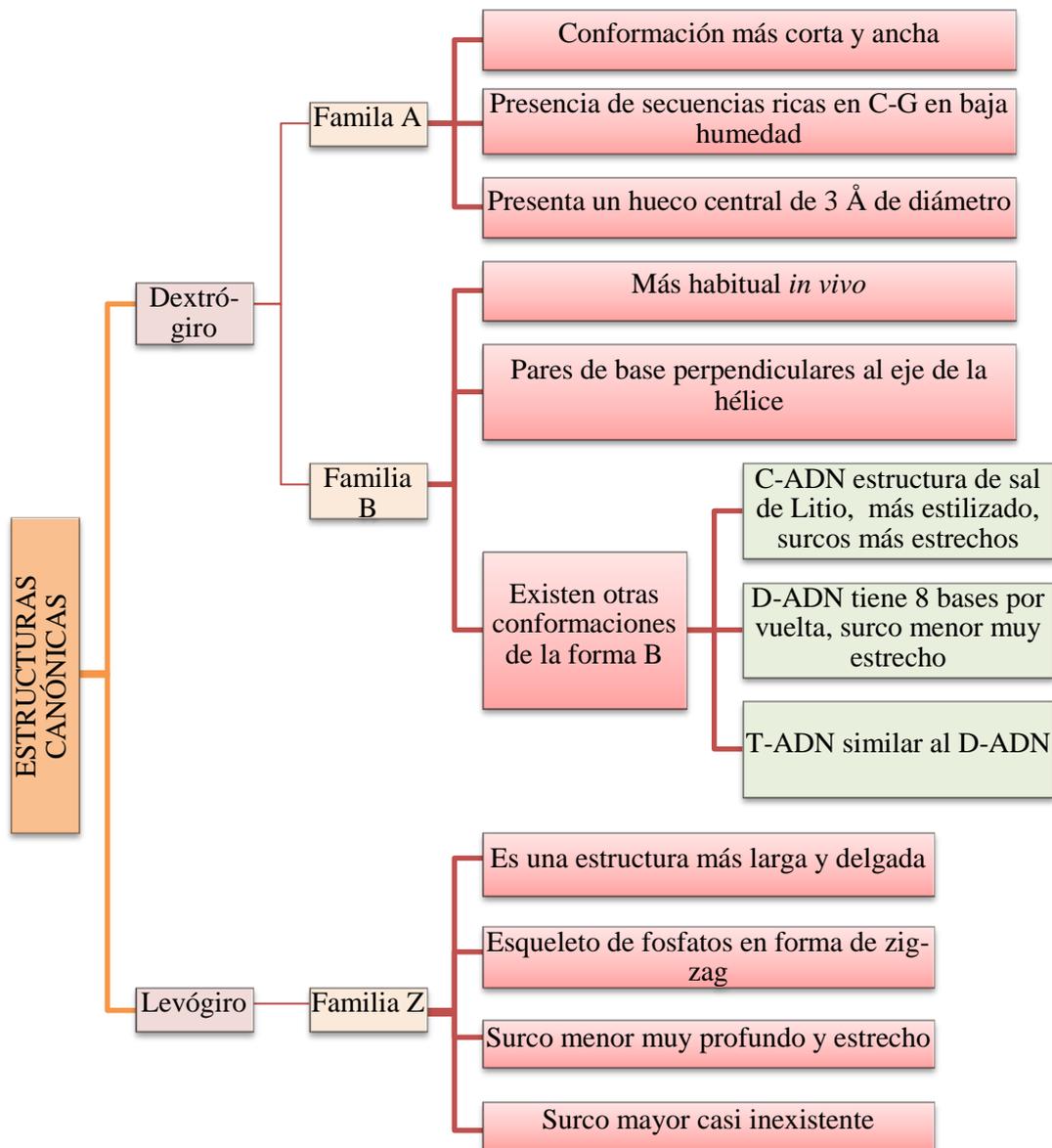
Las estructuras canónicas de los Ácidos Nucleicos son estructuras secundarias que se caracterizan por la asociación de dos cadenas antiparalelas. Este tipo de ADN puede ser dextrógiro o levógiro según el giro de las cadenas. En la tabla 1-1 se muestran los parámetros estructurales de las estructuras canónicas.

**Tabla 1.1:** Resumen de parámetros estructurales de Estructuras Canónicas

Tipo de estructura	ADN – B	ADN-A	ADN-Z
Altura por base(Å )	3.40	2.90	3.7
Torsión Helicoidal (°)	36.7	32.7	-10/-50
Ancho del surco (Å)	11.7/5.7	2.7/11	8.5
Polarización de la Cadena	+ -	+ -	+ -
Hélice	RH	RH	LH
Número de bases por vuelta	10.5	11.0	12.0
Ancho (Å)	18	26	
C1'-C1'	10		
Plegamiento de Azúcar	C2'	C3'	C2'/C3'

Fuente: (NEIDLE & BALASUBRAMANIAN, 2007)

Realizado por: Daniela Barragán



**Figura 4-1.** Clasificación de las Estructuras Canónicas  
 Realizado por: Daniela Barragán

#### 1.2.4. Estructuras no canónicas de los Ácidos Nucleicos

Los Ácidos nucleicos poseen una gran variedad conformacional formando diferentes estructuras según el emparejamiento de las bases. Por ejemplo se tiene ácidos nucleicos de tres cadenas donde la tercera hebra se une mediante puentes Hidrógeno de tipo Hoogsteen; también se puede dar ácidos nucleicos de cuatro cadenas o G-Cuádruples que toman gran interés investigativo por su intervención en procesos biológicos.

### 1.2.4.1. *G-Cuádruple*

#### 1.2.4.1.1. *Estructura y Topología*

G-cuádruples son secuencias de ADN caracterizados por tener bases ricas en guanina, cada base se encuentra conectadas por enlaces de hidrógeno Hoogsteen formando una estructura planar cuadrada denominada tétrada de guanina, (AL-NASIR & MADHANI, 2009, <http://www.al-nasir.com>) dependiendo la dirección de sus hebras pueden estar alineados en paralelo, antiparalelo o una mezcla de estas dos formas. (HUPPERT, 2007, pp. 1471-2962)

G-Cuádruples pueden estar formados por uno, dos, o cuatro hebras de ADN dando lugar a una gran variedad de topologías resultantes de una combinación de hebras, tamaños y secuencias. La topología más simple es aquella que está formada por la asociación de cuatro bases de guanina, se describe de la siguiente manera:



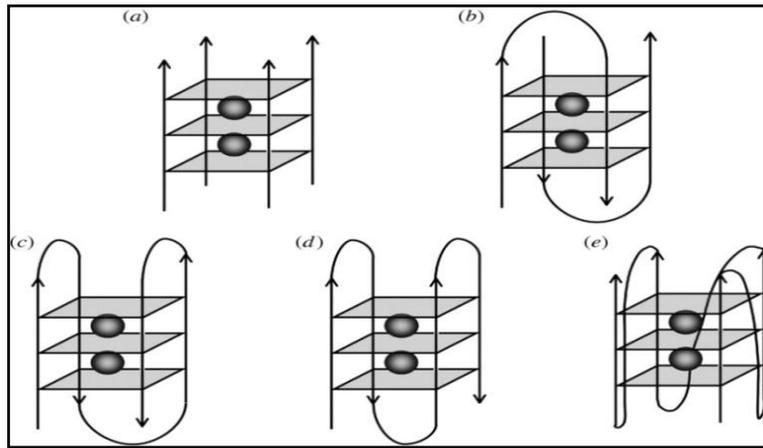
Donde  $X_n$  es alguna base de longitud  $n$  y  $G_p$  representa cualquier número de guaninas involucradas en la formación de la tétrada de longitud  $p$ . Estas Estructuras se denominan intermoleculares o tetramoleculares; puede darse el caso de G-Cuádruples formados a partir de dos hebras separadas ricas en bases de guanina, estas estructuras se conocen como bimoleculares.

También se puede formar secuencias que contengan bases distintas de guanina formando G-Cuádruples de una sola hebra, estas estructura se denominan intramoleculares y se expresa de la siguiente manera: (NEIDLE & BALASUBRAMANIAN, 2007) (BURGE, et al., 2006, pp. 5402-5415)



donde  $X_n$  es cualquier base distinta a la guanina de longitud  $n$ ,  $G_o$  representa el número de guaninas de una tétrada de longitud  $o$  y  $X_p$  representa algún nucleótido de longitud  $p$  involucrado en la formación del bucle. (NEIDLE & BALASUBRAMANIAN, 2007)

La topología de los G-Cuádruples depende de cómo se conecten las hebras entre sí. La figura 5-1 muestra un esquema de las posibles maneras en las que las hebras se pueden asociar.



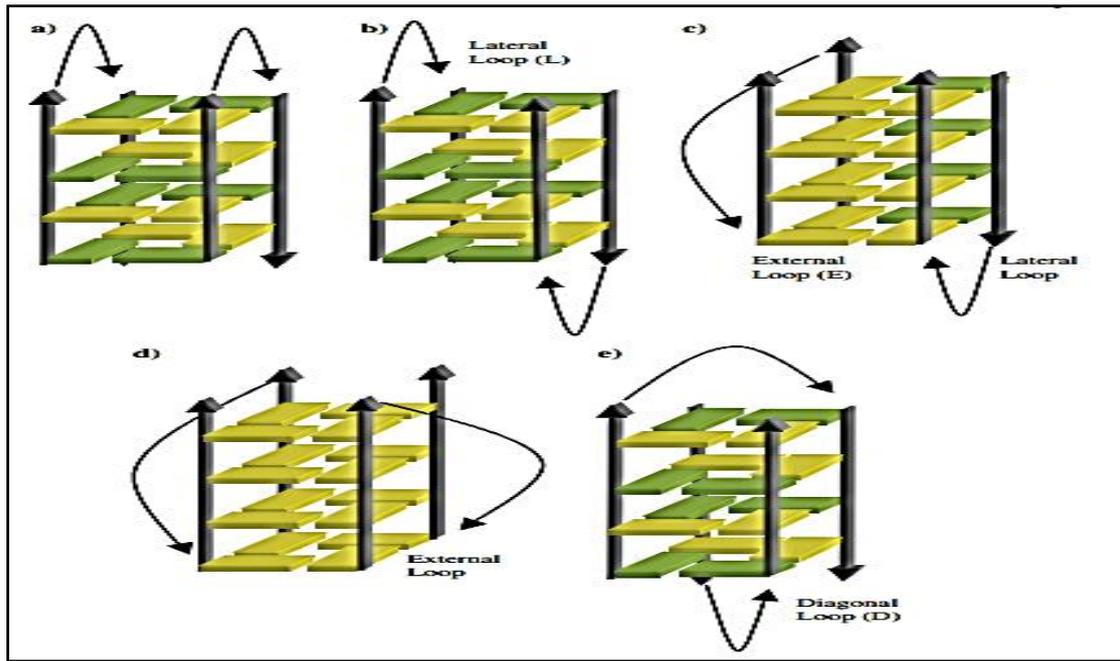
**Figura 4-1.** Esquemas de formas G-cuádruple.  
 (a)Tetramolecular Paralelo (b)Bimolecular Antiparalelo (c) y  
 (d)Unimolecular Antiparalelo (e)Unimolecular Paralelo  
**Fuente:** (HUPPERT, 2007)

Las estructuras de G-Cuádruple se clasifican según las polaridades de sus hebras y la localización de los bucles que unen las hebras de guanina en cuádruples dando lugar a hebras simples o dobles. (BURGE, et al., 2006, pp. 5402-5415) El elemento que une las hebras al cuádruple pueden ser diagonal, lateral o externo. (Figura 6-1) Estos nucleótidos enlazantes son importantes en la determinación de la estabilidad en longitud y secuencia. (NEIDLE & BALASUBRAMANIAN, 2007)

HÉLICE	LATERAL	DIAGONAL
<ul style="list-style-type: none"> <li>•Hebras paralelas adyacentes se conectan al bucle para enlazarse a G-tétrada inferior con G-tétrada superior.</li> <li>•Esta característica se ha encontrado en estructuras cristalinas y en secuencias teloméricas humanas y también en cuádruples no teloméricos.</li> </ul>	<ul style="list-style-type: none"> <li>•Bucles unidos por G-hebras adyacentes.</li> </ul>	<ul style="list-style-type: none"> <li>•Es un bucle antiparalelo que une G-hebras opuestas y se han observado en estructuras de secuencia telomérica de <i>Oxytricha nova</i>.</li> </ul>

**Figura 5-1.** Tipo de elementos que unen las cadenas de G-Cuádruples  
**Fuente:** (BURGE, et al., 2006, pp. 5402-5415)

La figura 7-1 muestra un esquema de los bucles en estructuras cuádruples bimoleculares.



**Figura 6-1.** Diagrama Esquemático de Bucles para Cuádruples Bimoleculares.

a) Bucles Laterales conectados a hebras antiparalelas en la misma cara b) Bucles laterales conectados a hebras antiparalelas en cualquiera de los extremos c) Bucles diagonales conectados a hebras antiparalelas opuestas d) Bucles externos conectados a hebras paralelas juntas e) Bucles diagonales conectados a hebras antiparalelas.

Fuente: (NEIDLE & BALASUBRAMANIAN, 2007)

#### 1.2.4.1.2. Formación de ADN G-Cuádruples en sistemas biológicos

La primera formación G- Cuádruple se observó en el ADN telomérico de varios organismos. La existencia *in vivo* de los G-Cuádruples de estas estructuras se comprobó mediante anticuerpos específicos para G-Cuádruples paralelos y antiparalelos en ADN telomérico de *ciliate Stylonychia lemnae*.

La detección de G-Cuádruples en los extremos cromosómicos humanos se realizó utilizando ligandos específicos. (YANG & OKAMOTO, 2010, pp. 619-646) Además, G-Cuádruples se han encontrado en otras regiones del genoma como sitios de switch de inmunoglobina, ADN ribosomal, ciertas regiones de ARN. (YANG & OKAMOTO, 2010, pp. 619-646)

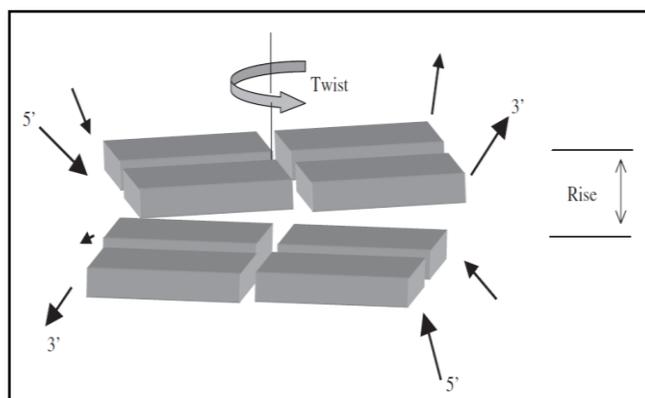
Las estructuras más estudiadas de G-Cuádruples provienen de las secuencias repetitivas de telómeros humanos. En 1993 se obtuvo esta estructura mediante RMN y resultó en una estructura de forma antiparalela, con lazos (bucles) en la parte superior e inferior.

En 2004, sin embargo, una estructura cristalina de rayos X se resolvió para la misma secuencia dando como resultado una forma paralela completamente diferente, con bucles alrededor de los lados de la estructura. Estos trabajos y otros sugieren que en solución en realidad existe una mezcla de estas formas y otros, coexistiendo en equilibrio. (HUPPERT, 2007, pp. 1471-2962) Este fenómeno se podría esperar en otras secuencias, dando lugar a una diversidad de conformaciones que podrían ser de suma importancia para la relevancia biológica.

#### 1.2.4.1.3. Factores de Estabilidad

Los factores que estabilizan en estructuras del tipo G-Cuádruple son: Fuerzas de Base - apilamiento, enlace de hidrógeno, estructura de hidratación.

**Base – Apilamiento.-** La estabilidad es provista a las hélices de ADN por el apilamiento de las bases a través de una combinación de fuerzas hidrofóbicas, electrostáticas y las fuerzas de Van der Waals.



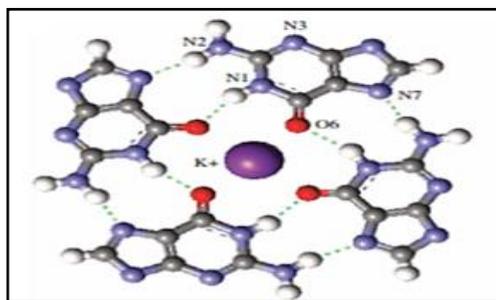
**Figura 7-1.** Ejemplo de apilamiento de pares de bases.

**Fuente:** (NEIDLE & BALASUBRAMANIAN, 2007)

**Enlace de hidrógeno.-** La estructura del G-Cuádruple presenta 8 enlaces de hidrógeno entre las cuatro bases, siendo dos enlaces por base y confieren estabilidad al sistema. (HUPPERT, 2007, pp. 1471-2962)

**Estructura de Hidratación.-** Los G-Cuádruples contienen surcos de moléculas de agua alrededor de los enlaces de hidrógeno aceptores y donadores, como el grupo amino exocíclico Nitrógeno (N2) y el átomo Nitrógeno (N3) para cuádruples. La simetría en las hebras paralelas de cuádruples está reflejada en la estructura de hidratación simétrica en los cuatro surcos formados por las hebras de las tétradas. (NEIDLE & BALASUBRAMANIAN, 2007)

**Iones Metales.-** la formación y estabilización de las estructuras G-Cuádruples requiere de la presencia de cationes monovalentes como  $K^+$  y  $Na^+$  que se encuentran en el centro de cada par de tétradas. (YANG & OKAMOTO, 2010, pp. 619-646)



**Figura 8-1.** Cuarteto de guaninas con presencia del ion  $K^+$  en el canal central  
Realizado por: Daniela Barragán

La ubicación precisa de los cationes entre las tétradas es dependiente de la naturaleza de los iones; con iones de  $Na^+$  dentro del canal se observa una amplia gama de geometrías. Por ejemplo, en algunas estructuras, se observa que un ion  $Na^+$  que está en plano con una G- tétrada; mientras que en otras está entre las dos G-tétradas.

Los iones de  $K^+$  están siempre equidistantes entre cada plano de la tétrada. También las estructuras G-Cuádruples pueden ser sustituidas por otros iones. Estos iones se enumeran en la tabla 2-1 con sus diversos radios iónicos.

**Tabla 2-1:** Radio Iónico de cationes estabilizantes de G-Cuádruples

Elemento	$K^+$	$Na^+$	$NH_4^+$	$Rb^+$	$Cs^+$	$Li^+$	$Ca^{2+}$
Radio Iónico	1.52 1.33 <sup>100</sup>	1.16 (0.97)	1.43 <sup>71</sup>	1.66	1.81	0.9	0.99

Fuente: (NEIDLE & BALASUBRAMANIAN, 2007)

Realizado por: Daniela Barragán, 2015

#### 1.2.4.1.4. Lugar y Función

En los extremos de los cromosomas existen regiones especializadas de ADN denominados telómeros. Estos extremos cromosómicos se resintetizan utilizando la enzima telomerasa, cuya función es activar los sistemas de reparación y proteger los extremos del ADN. Además, los telómeros juegan un papel importante en el mantenimiento de la inmortalización celular en varias líneas celulares. (CHEMIE.DE, 1997, <http://www.quimica.e>)

#### 1.2.4.1.5. *Aplicación*

G-Cuádruples Tetramoleculares debido a sus propiedades de simetría, limitado polimorfismo e inercia cinética son de gran interés en aplicaciones nanotecnológicas y biotecnológicas. (THAO, et al., 2012, pp. 2012-2334) (ZHOU, et al., 2014, pp. 159-168)

Por ejemplo estas estructuras sirven para construir nano-hilos de guanina utilizados en nanoelectrónica, nanotecnología y biosensores. (JISSY, et al., 2011, pp. 12530-12546) Además el G-Cuádruple es un principal candidato para combatir el cáncer. Estructuras a base de ADN denominadas aptámeros tienen un gran potencial terapéutico y diagnóstico. Además otros aptámeros pueden ser empleados en aplicaciones químicas. (JISSY, et al., 2011, pp. 12530-12546)

### 1.3. Técnicas Experimentales

#### 1.3.1. *Cristalografía de Rayos X*

Un cristal es un sólido que consiste en repeticiones periódicas. (WILEY & SONS, 1922). En solvente acuoso los cristales de proteínas y ácidos nucleicos contienen cantidades significativas, son frágiles y sensibles a una variedad de factores ambientales. Las moléculas en los cristales se mantienen unidos por fuerzas débiles como enlaces de hidrógeno e interacciones hidrofóbicas.

Para cristalizar la macromolécula se tiene que separar de la solución acuosa y formar un ordenado en fase sólida cristalina. Se debe controlar el nivel de saturación y sobresaturación en la solución. Hay muchos factores que afectan el crecimiento de cristales como la pureza de proteína, el pH de la solución, la fuerza iónica, la temperatura, la concentración de la proteína, entre otros. Estos factores deben ser analizados hasta obtener un cristal difractado de calidad.

##### 1.3.1.1. *Principio Físico*

Bragg demostró que un conjunto de planos paralelos con índice  $hkl$  y espaciado interplanar  $d_{hkl}$  produce una difracción de haz cuando se incide rayos x de longitud de onda  $\lambda$  sobre los planos en un ángulo  $\theta$  y se refleja en el mismo ángulo donde  $n$  es un entero.

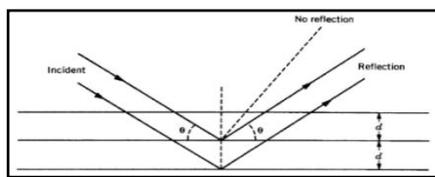
Esta observación se describe en la ecuación de Bragg, en la forma

$$n\lambda = 2d \sin \theta \quad (13)$$

donde  $\lambda$  es la longitud de onda de los rayos X incidente (en el rango entre 0,001 y 50 Å),  $d$  es la distancia interplanar del cristal, y  $n$  ( $= 1 ; 2 , \dots$ ) es un entero y se asocia con una franja de interferencia dada de un patrón de difracción. (WILEY & SONS, 1922)

$$n\lambda = 2d_{hkl} \sin \theta \quad (14)$$

Cuando se cumplen estas condiciones las ondas emergen desde el cristal en fase una con la otra, sufren interferencia constructiva produciendo una difracción fuerte. Cuando no se cumplen estas condiciones, las ondas emergen desde el cristal fuera de fase una de la otra, y sufren interferencia destructiva produciendo una reflexión débil o ausente. El conocer la estructura cristalina de las macromoléculas nos proporciona una comprensión de las características estructurales de un compuesto.



**Figura 9-1.** Difracción de Rayos X  
Fuente: (WILEY & SONS, 1922)

### 1.3.1.2. *Cristalografía ADN*

Oligonucleótidos de ADN son más difíciles de cristalizar que las proteínas. Los cristales de ácidos nucleicos son heterogéneos, esto indica que la calidad promedio de la difracción es inferior que en cristales proteicos. La plasticidad del ADN contribuye a la obtención de buena difracción en los cristales. Varias formas de ADN pueden coexistir en el mismo cristal.

La molécula de ADN está hidratada y los cristales tienen un contenido de solvente alto dando como resultado un alto grado de libertad para el esqueleto fosfodiéster, iones y solvente. Es necesario examinar las tendencias específicas para la cristalización de oligonucleótidos y cómo estos pueden afectar la calidad de difracción.

### 1.3.2. *Resonancia Magnética Nuclear*

La Resonancia magnética nuclear (RMN) es un fenómeno de absorción, similar a ultravioleta (UV) e infrarrojo (IR), pero la energía de RMN es la radiación de radio-frecuencia emitida por los núcleos expuestos a un campo magnético. La RMN es una herramienta importante para el estudio de la estructura molecular y el comportamiento de estructuras en la química. (WILEY & SONS, 1922).

Esta técnica experimental está asociada con el momento angular intrínseco de spin y las propiedades magnéticas de los núcleos atómicos. El efecto Zeeman o desdoblamiento energético se da cuando ocurre la interacción entre el momento magnético del núcleo y el campo, además mediante la absorción y emisión de fotones electromagnéticos con la frecuencia adecuada pueden producirse transiciones entre esos nuevos estados energéticos. (LARRAONA & AREVALO, 2008, <http://nuclear.fis.ucm.es>)

#### 1.3.2.1. *Principios Generales*

##### **El campo magnético y momento magnético**

Un campo magnético se produce por el movimiento de la carga eléctrica. Se caracteriza por dos vectores: B, la inducción magnética o densidad de flujo magnético, y H, la intensidad de campo magnético (también llamada intensidad). (WILEY & SONS, 1922)

La unidad Sistema Internacional de B es el tesla (T), también llamado Wb/m<sup>2</sup>. La unidad de H es en amperios por metro, donde un amperio es igual a un culombio por segundo.

Los vectores B y H están relacionados por la ecuación

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} \quad (15)$$

El momento magnético  $\mu_0$  es producido por el spin del electrón (momento magnético del spin) o el movimiento orbital (momento magnético orbital). Su unidad es magnetón Bohr

$$1 \mu_s = \frac{eh}{4\pi mc'} \quad (16)$$

donde  $e$  es la carga eléctrica,  $h$  es la constante de Planck,  $m$  es la masa del electrón, y  $c'$  es la velocidad de la luz.

## Propiedades magnéticas de los núcleos

El spin de un núcleo se puede describir en términos de dos números cuánticos: el número del spin cuántico, o spin  $I$ , y el número cuántico spin magnético, o de proyección  $m_I$ . Los dos números están relacionados como se muestra en la ecuación: (WILEY & SONS, 1922)

$$m_I = 0, \pm 1, \pm 2, \dots, \pm I \quad (17)$$

Dado que, para un núcleo, el número  $I$  que sólo puede tomar dos valores,  $1$  ó  $1/2$ , el número  $m_I$  sólo puede tomar valores de  $0, \pm 1/2$ , and  $\pm 1$ . Desde  $I$  y  $m_I$ , dos cantidades,  $I$  y  $I_z$  están definidas:

$$I = \sqrt{I(I+1)}\hbar \quad (18)$$

$$I_z = m_I \hbar \quad (19)$$

donde  $\hbar = h/2\pi$ . La cantidad  $I$  se llama spin de momento nuclear y la cantidad  $I_z$  se llama la componente  $z$  del spin del momento nuclear.

El momento nuclear magnético  $\mu$  se expresa en términos del factor  $g$  nuclear,  $g_I$  y el magnetón nuclear,  $\mu_0$ , mediante el uso de la ecuación  $\mu = g_I \mu_0$

$$\mu_0 = e\hbar/2m \quad (20)$$

donde  $m$  es la masa de los núcleos. El factor  $g$  varía con el núcleo. (WILEY & SONS, 1922)

## Precesión y frecuencia de Larmor

Si un campo magnético externo de intensidad  $H$  actúa sobre el núcleo, se experimenta una fuerza que se expresa en forma de un producto vectorial:

$$\Gamma = \mu \times \beta \quad (21)$$

Esta fuerza da lugar a un movimiento conocido como precesión de Larmor: (VALENCIA, 2012, <http://www.uv.es>)

$$\mathbf{v} = \frac{1}{2\pi} \gamma \beta \quad (22)$$

## **Espectroscopia de RMN**

La técnica espectroscópica permite obtener información física, química y estructural de las moléculas y se basa en la incidencia de radiación electromagnética sobre un núcleo que este ocupando un nivel inferior logre pasar al nivel superior. Estas transiciones origina la Espectroscopia de Resonancia Magnética Nuclear. (VALENCIA, 2012, <http://www.uv.es>)

## CAPÍTULO II

### MARCO METODOLÓGICO

#### 2.1. Material

##### 2.1.1. Recursos Computacionales

El proyecto se realizó en la Universidad San Francisco de Quito, en el laboratorio del grupo de Química Computacional y Teórica. Para él envío de cálculos se utilizó el Sistema de Cómputo de Alto Rendimiento de la USFQ (HPC-USFQ) que está compuesto por 8 nodos

Procesador	⇒	Intel(R) Xeon(R) CPU
Frecuencia	⇒	2.4 GHz
RAM	⇒	500 GB
Sistema operativo	⇒	GNU/Linux

El HPC-USFQ es un sistema compuesto por un conjunto de servidores multi-procesador destinado para proyectos de investigación sobre simulación. En el HPC están instalados los siguientes programas: Gaussian 09, CRYSTAL 09, VASP 5.3.3., LAMMPS, Gromacs y NAMD.

Para el análisis e interpretación de resultados se utilizó un ordenador con las siguientes características:

RAM	⇒	500 GB
Sistema operativo	⇒	GNU/Linux

Para las pruebas y análisis redacción del trabajo se utilizó una laptop:

Procesador	⇒	iCore 3
Frecuencia	⇒	2.4 GHz
RAM	⇒	4 GB
Sistema operativo	⇒	64bits

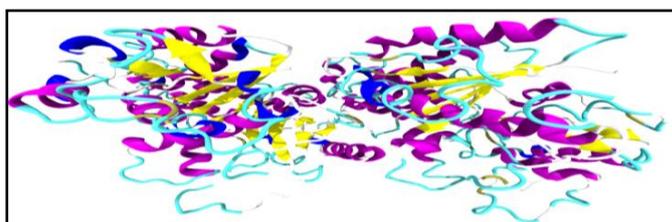
### 2.1.2. Programas

**NAMD.-** es un código en paralelo diseñado para la simulación de dinámica molecular de alto rendimiento de grandes sistemas. NAMD escala a cientos de procesadores de alta gama de plataformas paralelas, así como decenas de procesadores en grupos de productos básicos de bajo costo y también funciona en computadoras de escritorio y portátiles individuales. NAMD trabaja con campos de fuerza AMBER y CHARMM, parámetros y formatos de archivo. (PHILLIPS, et al., 2005, pp. 1781-1802)

**VMD.-** es un programa computacional, diseñado para modelación, visualización, análisis de representación tridimensionales de los sistemas biológicos como proteínas, ácidos nucleicos, lípidos. Esta herramienta lee archivos descargados desde una Base de Datos (Protein Data Bank) y permite analizar los resultados de una trayectoria de dinámica molecular.

VMD fue desarrollado por el grupo de Biofísica Teórica y Computacional de la Universidad de Illinois y el instituto Beckman. La versión inicial de VMD fue escrita por William Humphrey, Andrew Dalke, Ken Hamer, Jon Leech, y James Phillips y fue lanzada en 1995.

Esta versión es compatible con las simulaciones de NAMD. (HUMPHREY, et al., 1996, pp. 27-38) Proporciona una amplia variedad de representación de los átomos por ejemplo: líneas, puntos, esferas, cilindros, tubos, enlaces, dibujados animados, estructuras secundarias de proteína, entre otros.



**Figura 10-2.** Proteína Tubulina en representación de New Cartoon

Realizado por: Daniela Barragán

**GROMACS.-** se desarrolló en el departamento de Química Biofísica de la Universidad de Groningen, es un paquete versátil para realizar dinámica molecular, es un software libre diseñado para moléculas bioquímicas y polímeros. Además incluye un paquete denominado GROMACS TOOLS utilizado para análisis de trayectoria. (MABRAHAM, 2009, <http://www.gromacs.org/>)

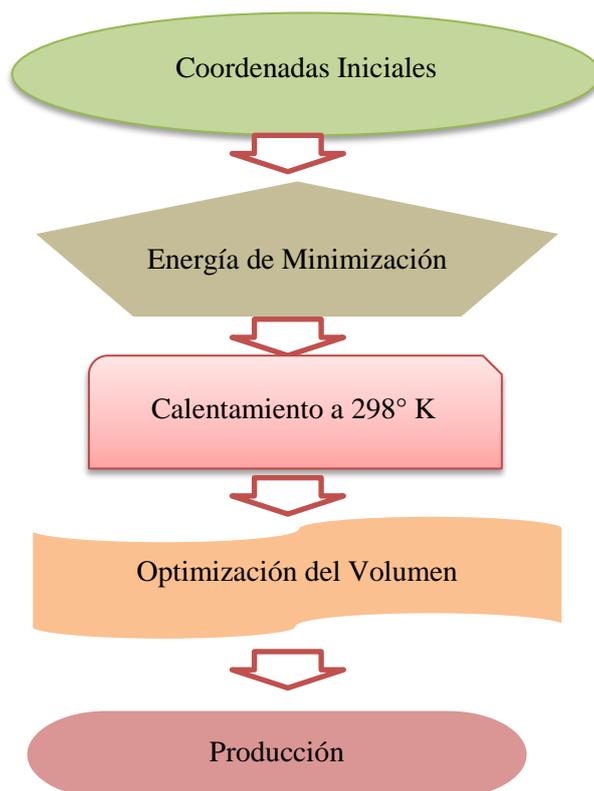
**Wolfram Mathematica.-** es una herramienta utilizada en el entorno de investigación, análisis, creación de interfaces. En este proyecto se utilizó para realizar las gráficas de diferentes parámetros evaluados en las diferentes etapas de modelación. (Este software está disponible para

investigadores y estudiantes mediante una licencia adquirida por CEDIA-Ecuador) (WOLFRAM, 1987, <http://www.wolfram.com>)

## 2.2. Métodos

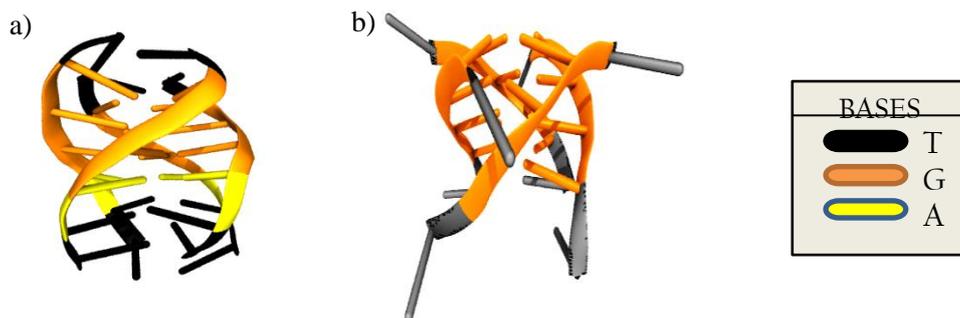
### 2.2.1. Modelado Molecular

En Dinámica Molecular los pasos para realizar una simulación son mostrados en la Figura 2-2.



**Figura 11-2.** Protocolo de modelado molecular  
Realizado por: Daniela Barragán

Las simulaciones fueron realizadas con el campo de fuerza CHARMM27 implementado en NAMD. Los modelos y trayectorias fueron visualizados en VMD para Windows versión 1.9. Las coordenadas iniciales de las estructuras fueron obtenidas de Protein Data Bank (Berman, et al., 2000, pp. 235-242). Todas las estructuras de estudio son G-cuádruples tetramoleculares paralelos como se muestra en la figura 3-2.



**Figura 12-2.** G-cuádruplex paralelos tetramoleculares. a) 1NP9 b) Monómero 2GW0  
Realizado por: Daniela Barragán

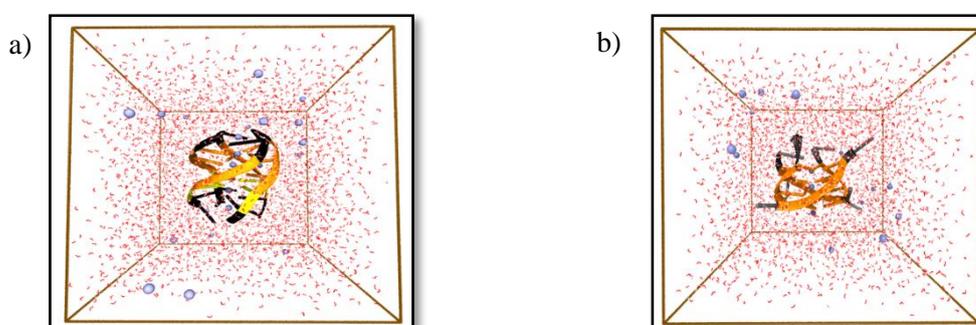
**Tabla 1-2:** Características de las estructuras de ADN G-cuádruple estudiadas

Estructura	Secuencia	Origen	Coordenadas Experimentales	Iones del canal	N. de moléculas de Agua
1NP9	5'-d(TTAGGGT)-3'	Telómero Humano	Resonancia Magnética Nuclear	K <sup>+</sup>	12405
Monómero 2GW0	5'-d(TGGGGT)-3'	Telómero Eucariótico	Difracción de Rayos X	Ca <sup>2+</sup> , Na <sup>+</sup>	11265

Realizado por: Daniela Barragán

Los sistemas fueron solvatados con moléculas de agua del modelo TIP3 contenidas en una caja de dimensiones que aseguran una mínima distancia de 12 Å de la biomolécula al borde de la caja. Además los sistemas fueron neutralizados con iones Na<sup>+</sup> y Cl<sup>-</sup>, colocados aleatoriamente a una distancia del soluto de 5 Å y una distancia entre iones 5 Å y una distancia entre iones de 5 Å.

Se empleó condiciones periódicas de borde y se utilizó el método de Ewald para evaluar las interacciones de largo alcance con un radio de corte de 12 Å.



**Figura 13-2.** Preparación del sistema.  
a) Estructura G-Cuádruple 1NP9 b) Estructura G-Cuádruple del Monómero 2GW0  
Realizado por: Daniela Barragán

La primera etapa de minimización fue de 300 pasos con restricción de átomos del ADN G-Cuádruple y los iones del canal para asegurar que solo las moléculas de agua y los iones neutralizantes se optimicen durante la primera etapa.

Luego se procedió a realizar la segunda etapa de minimización del sistema a 5000 pasos sin restricción de movimientos de los átomos, con el fin de minimizar la estructura de ADN G-Cuádruple. Después de la minimización, los sistemas están energéticamente estables y se elimina la distorsión estérica entre los átomos.

Los sistemas fueron sometidos a calentamiento a 298 K, con el uso del ensamble NVT donde el volumen y la temperatura son constantes; la simulación corresponde a la etapa de dinámica molecular 600 ps (300 000 pasos de simulación).

Para optimizar el volumen de los sistemas se aplicó el ensamble isotérmico-isobárico con 298 K y a 1 atmósfera. La producción tuvo una trayectoria de 200 ns con el ensamble NVT (corresponde a 100 millones de pasos de simulación).

Las coordenadas del sistema fueron guardadas cada 500 fs durante la simulación. El tamaño del “timestep” usado para integrar cada paso de la simulación fue de 2 fs por paso.

### 2.2.2. *Análisis de trayectoria*

La dinámica molecular proporciona diferentes tipos de información según los resultados obtenidos, en este caso de estudio se analizó la información estructural de la dinámica y la información estructural promediada.

#### **Información estructural de la dinámica**

Las fluctuaciones temporales de una molécula se encuentran en detalle durante la evaluación de la trayectoria mediante el cálculo del RMSD (desviación cuadrática media) respecto a una estructura de referencia o inicial, la cual se compara con cada una de las estructuras hasta el final de la trayectoria

$$RMSD_k = \left[ \frac{1}{N} \sum_{l=1}^{3N} (x_{kl} - x_l)^2 \right]^{1/2} \quad (23)$$

Donde  $x_{kl}$  representa la l-coordenada en la estructura k,  $x_l$  es el valor de la l coordenada en la estructura de referencia y N es el número de átomos.

En este estudio otro parámetro que se empleó fue el radio de giro que proporciona la información sobre la flexibilidad de la estructura y se calcula de la siguiente manera

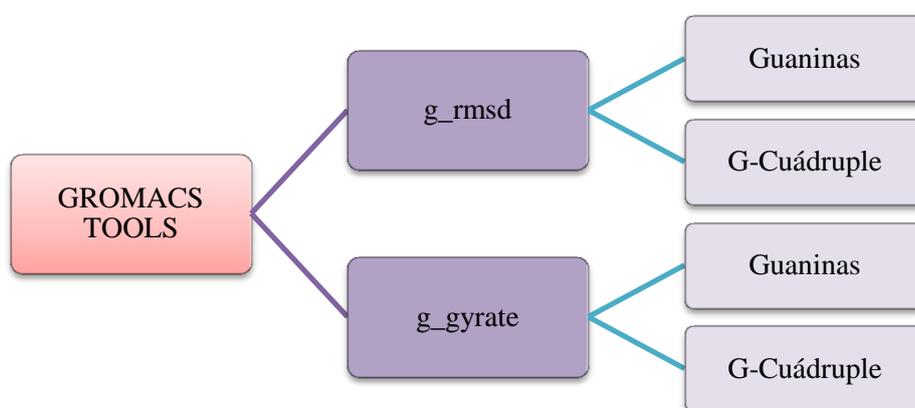
$$R_{gyr} = \left( \frac{\sum r_i^2 m_i}{\sum m_i} \right)^{1/2} \quad (24)$$

Donde  $m_i$  es la masa del átomo  $i$  y  $r_i$  es la posición del átomo respecto al centro de masas de la molécula.

### Información Estructural Promediada

La información estructural promediada calcula diferentes parámetros estructurales, por ejemplo las distancia de los puentes de hidrógeno. En este trabajo se presenta Mapas de Calor de los valores de distancia de los puentes de hidrógeno de las conformaciones estables y transición encontradas en mayor proporción durante la simulación de 200 ns.

Las simulaciones de Dinámica Molecular implican trayectorias grandes, su respectivo estudio se puede realizar mediante visualización o análisis cuantitativo. En este proyecto para visualizar trayectorias de dinámica molecular, se utilizó VMD (Visual Molecular Dynamics) y para el análisis se utilizó las herramientas de GROMACS `g_rmsd` y `g_gyrate` donde se analizó los residuos de guanina y la estructura del G-Cuádruple en su totalidad por la tanto en los dos sistemas se analizaron los cambios y detalles estructurales.



**Figura 14-2.** Esquema del uso de las herramientas de GROMACS. Uso para análisis estructurales de los residuos de Guanina y de la estructura G-Cuádruple  
Realizado por: Daniela Barragán

### 2.2.3. *Análisis de las conformaciones estructurales de estructuras G-Cuádruples*

Las trayectorias de dinámica molecular generan información de las posiciones y velocidades del sistema analizado. En este caso la trayectoria fue de 200 ns donde se obtuvo un total de 400 000 conformaciones estructurales.

Estas conformaciones fueron seleccionadas según criterios que serán descritos a continuación y se usaron como referencia inicial para el algoritmo de Clasificación Vecino Cercano k (k-NN) con el objetivo de comparar con las conformaciones de técnicas experimentales como RMN y Cristalografía para evaluar el campo de fuerza de CHARMM27.

#### **Criterios de selección**

En la etapa final de una simulación de dinámica molecular se debe analizar los resultados de parámetros termodinámicos y estructurales del sistema.

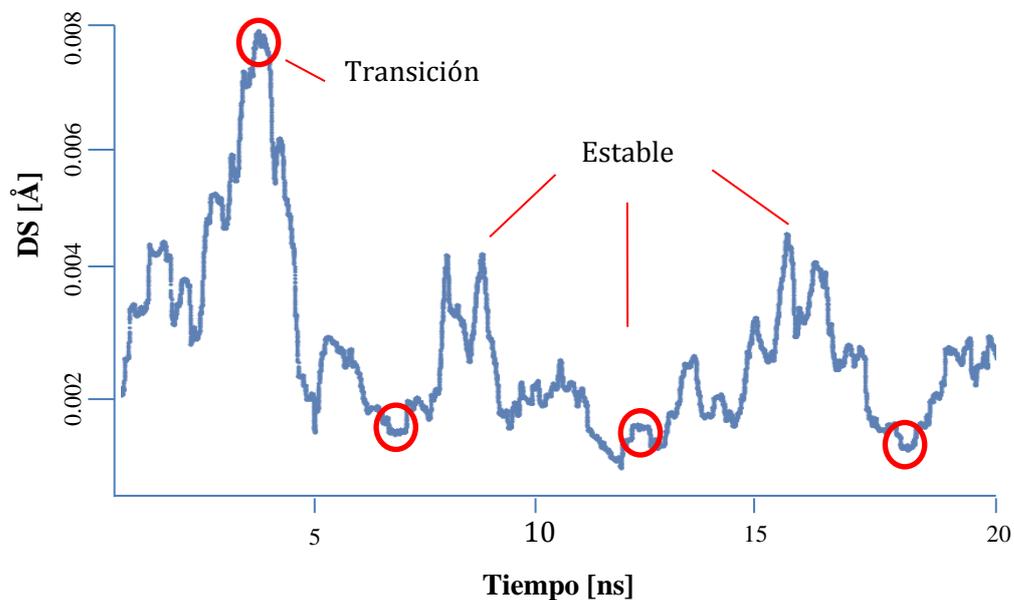
El criterio empleado es seleccionar zonas con valores inferiores donde el parámetro se mantenga constante, a la anterior zona se denomina conformación estable. Por otra parte, se identificaron zonas con valores superiores denominándose conformaciones estructurales de transición.

En estudio de este tipo, los parámetros más empleados para comparación de conformaciones estructurales son los parámetros de RMSD y Radio de Giro. Valores altos para ambos casos indican que la estructura ha cambiado en mayor medida con respecto a una estructura inicial.

Estos parámetros fueron calculados en base al movimiento o cambio de las guaninas y de la estructura completa, por lo cual son indicadores aceptables para identificar estructuras estables y de transición.

Por ejemplo para la estructura 1NP9 Calculados los valores de RMSD y Radio de Giro sobre el G-Cuádruple se grafican por medio del software Wolfram Mathematica y de los mismos parámetros se realizó una desviación estándar (DS) cada 1000 puntos debido a la dificultad presentada al incluir toda la trayectoria.

En la figura 6-2 se muestra una trayectoria de 20 ns con un total de 40 000 conformaciones estructurales. Se destacan cuatro zonas comprendidas por 500 conformaciones de las cuales, tres zonas con valores inferiores denominadas “estables” y por otro lado, una zona con valores superiores denominada “transición”.



**Figura 15-2.** Desviación estándar de la trayectoria de Dinámica Molecular durante 20 ns  
Realizado por: Daniela Barragán

### Validación

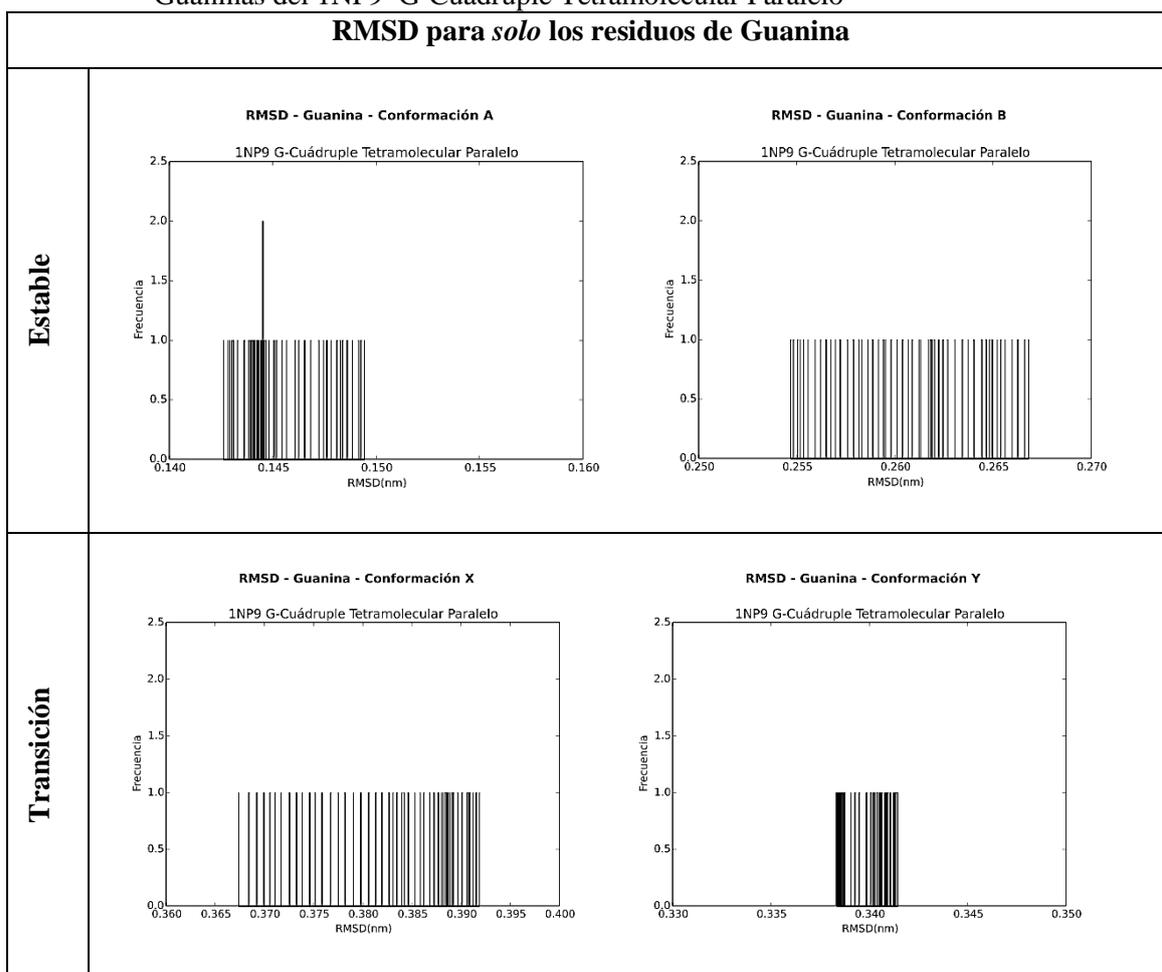
Luego se realizó la validación de las estructuras seleccionadas para ello se utilizó el software VMD, que permite confirmar que las conformaciones estructurales seleccionadas sean diferentes entre los parámetros de RMSD y Radio de Giro.

Cada estructura fue alineada con referencia a una estructura inicial debido a los movimientos de translación y rotación. Además se realizaron histogramas de cada conformación para corroborar que las conformaciones estables tengan valores inferiores y las conformaciones de transición presenten valores superiores.

Los histogramas se explican en las siguientes tablas.

En la Tabla 2-2 se observan histogramas de los valores de RMSD de *solo* los residuos de Guanina para cada conformación estructural. Las conformaciones estables presentan valores menores a 0.27 nm en cambio las conformaciones de transición tienen valores a partir de 0.33 hasta 0.4 nm.

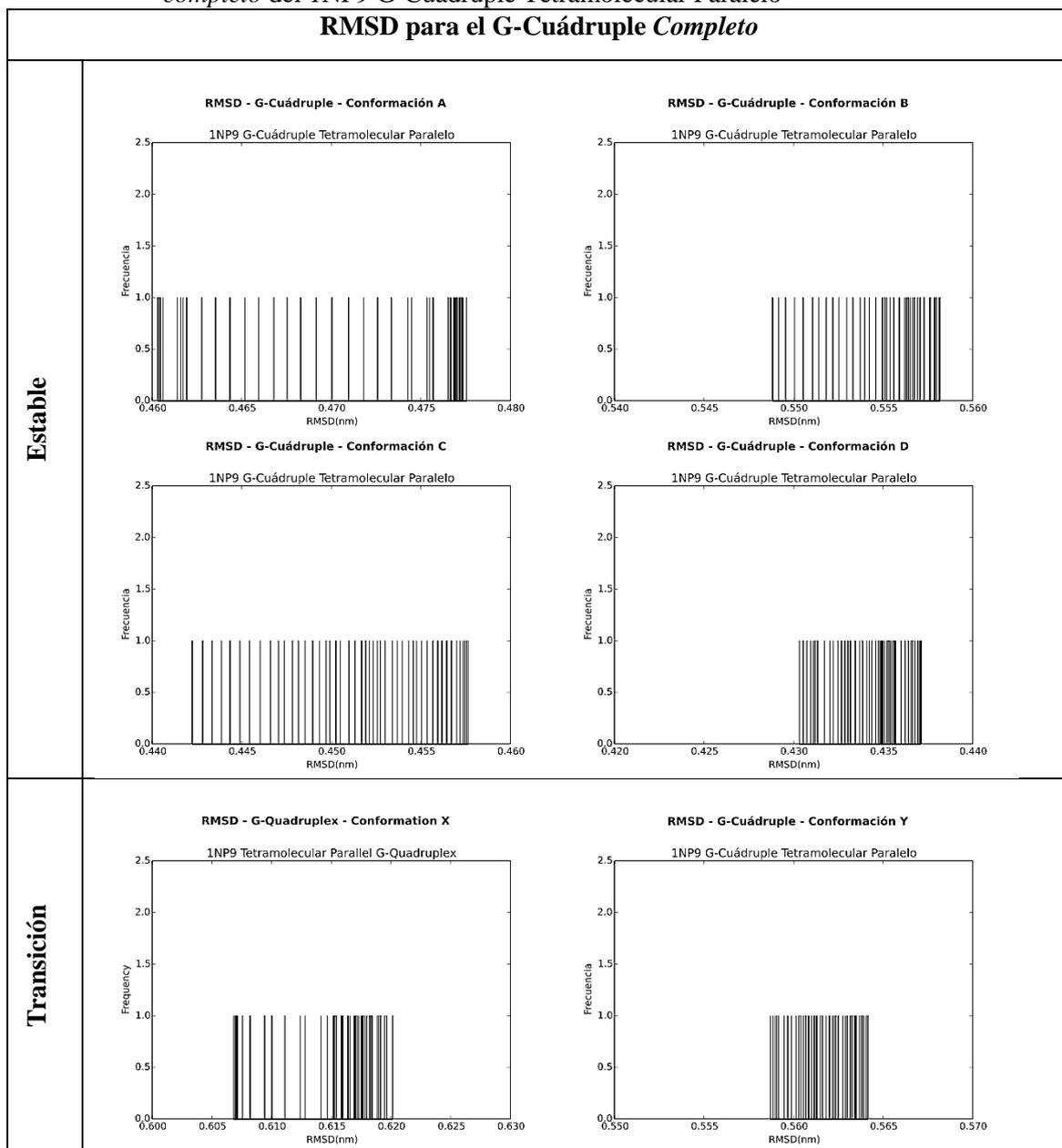
**Tabla 2-2:** Histograma de Conformaciones seleccionadas por RMSD para *solo* los residuos de Guaninas del 1NP9 G-Cuádruple Tetramolecular Paralelo



Realizado por: Daniela Barragán

En la Tabla 3-2 se muestran histogramas de los valores de RMSD para el G-Cuádruple *completo* para cada conformación estructural. Las conformaciones estables presentan valores de RMSD de 0.420 a 0.55 nm y las conformaciones de transición valores de 0.56 hasta 0.62 nm. Por tanto Se observa un aumento en el RMSD, al compararlo con el obtenido en el cálculo sobre *solo* los residuos de guanina. Este aumento es esperado porque se toma en cuenta los movimientos de las timinas y además las diferentes conformaciones del esqueleto.

**Tabla 3-2:** Histograma de Conformaciones seleccionadas por RMSD para el G-Cuádruple completo del 1NP9 G-Cuádruple Tetramolecular Paralelo



Realizado por: Daniela Barragán

**Anexo A** Histograma de Conformaciones seleccionadas por Radio de Giro para *solo* los residuos de Guaninas del 1NP9 G-Cuádruple Tetramolecular Paralelo

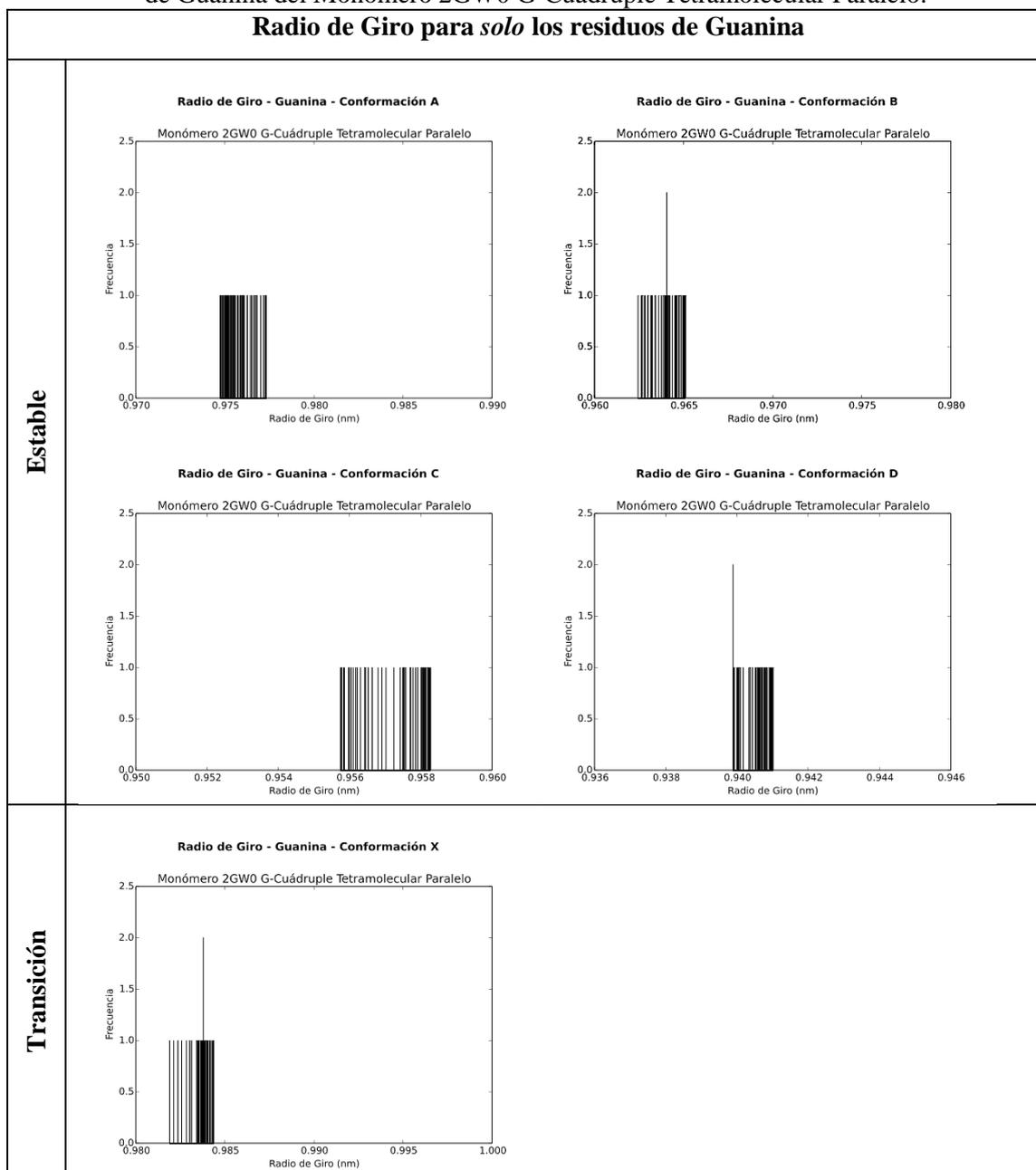
Estos histogramas corresponden a los valores de Radio de Giro para *solo* los residuos de Guanina para cada conformación estructural. Las conformaciones estables presentan valores desde 0.85-0.90 nm y la conformación de transición valores superiores a 0.91nm.

**Anexo B** Histograma de Conformaciones seleccionadas por Radio de Giro para el G-Cuádruple *completo* del 1NP9 G-Cuádruple Tetramolecular Paralelo

En el anexo se presenta histogramas de los valores de Radio de Giro para los residuos del G-Cuádruple para cada conformación estructural. Las conformaciones estables presentan valores desde 0.85-0.90 nm y la conformación de transición cuyo valor es superior a 0.92 nm.

En la Tabla 4-2 se muestran histogramas de los valores de Radio de Giro para *solo* los residuos de Guanina para cada conformación estructural. Las conformaciones estables presentan valores menores a 0.98 nm en cambio la conformación de transición tiene valores superiores a 0.98 nm.

**Tabla 4-2:** Histograma de Conformaciones seleccionadas por Radio de Giro para *solo* los residuos de Guanina del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo.

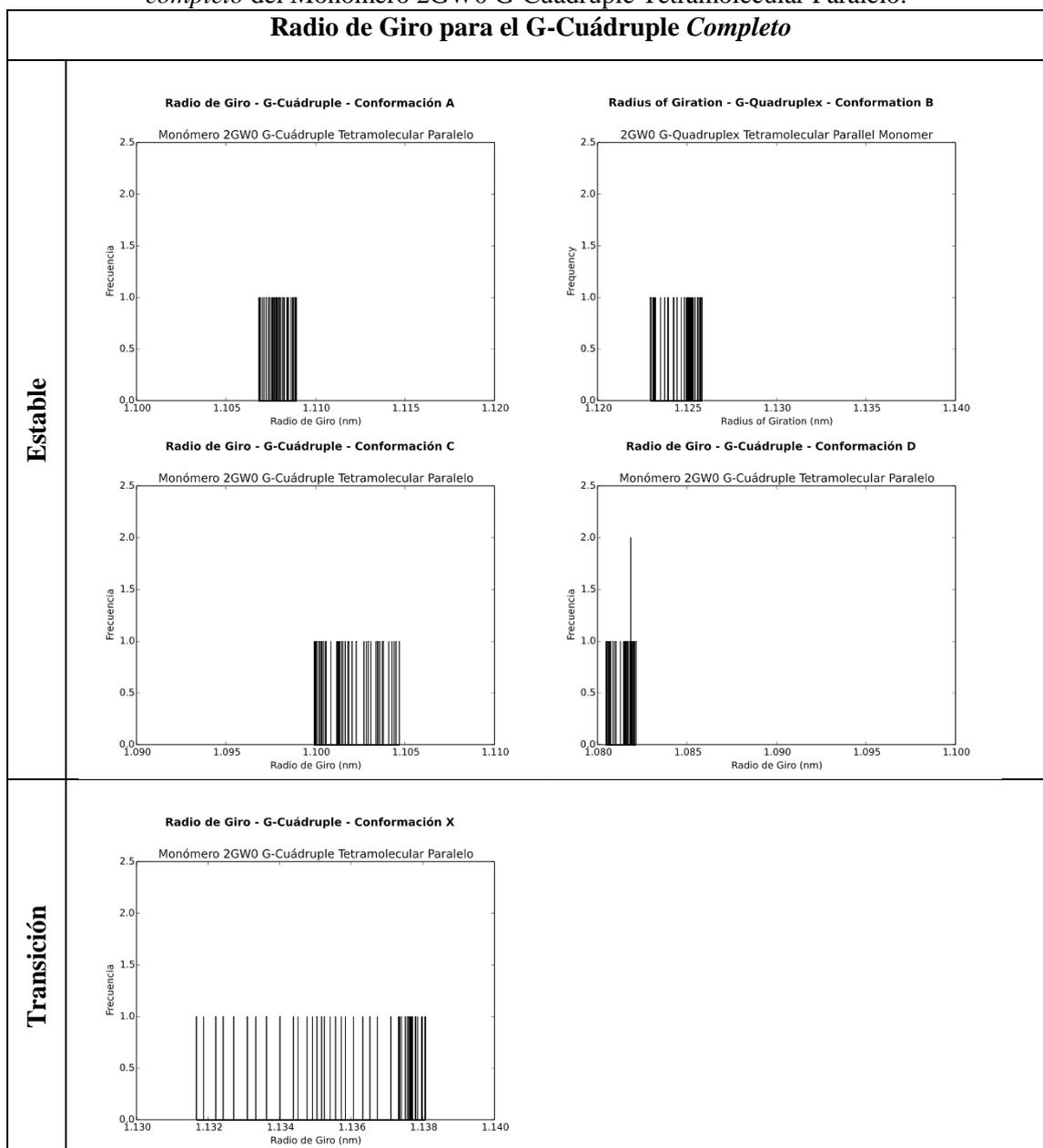


Realizado por: Daniela Barragán

En la Tabla 5-2 se muestran histogramas de los valores de Radio de Giro para el G-Cuádruple *completo* para cada conformación estructural. Las conformaciones estables presentan valores desde 1.08-1.13 nm y la conformación de transición valores superiores a 1.13 nm.

En este análisis los valores de Radio de Giro aumentaron porque se analiza la estructura completa del G-Cuádruple, por tanto es más larga en una dimensión debido a la presencia de timinas en los extremos de la estructura.

**Tabla 5-2:** Histograma de Conformaciones seleccionadas por Radio de Giro para el G-Cuádruple *completo* del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo.



Realizado por: Daniela Barragán

### **Anexo C** Histograma de Conformaciones seleccionadas por RMSD para *solo* los residuos de Guanina del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo

En este anexo se presentan histogramas de los valores de RMSD de los residuos del G-Cuádruple para cada conformación estructural. Las conformaciones estables presentan valores desde 0.23-0.29 nm y la conformación de transición tiene valores superiores a 0.32 nm.

### **Anexo D** Histograma de Conformaciones seleccionadas por RMSD para el G-Cuádruple *completo* del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo

En este anexo se presenta histogramas de los valores de RMSD de los residuos del G-Cuádruple para cada conformación estructural. Las conformaciones estables presentan valores desde 0.41-0.50 nm y la conformación de transición tiene un valor superior a 0.52 nm.

### **Clasificación de las conformaciones estructurales**

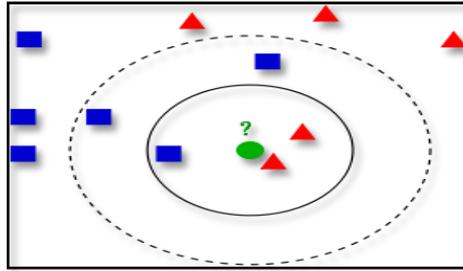
Las conformaciones estructurales seleccionadas de los sistemas 1NP9 y del monómero 2GW0 fueron clasificadas mediante el algoritmo k-NN que permitió evaluar las frecuencias de las conformaciones y determinar las conformaciones más frecuentes y las que estuvieron en menor presencia durante la simulación.

Vecino cercano k (kNN) es un algoritmo de Aprendizaje de Máquina, consiste en clasificar la clase más frecuente a la que pertenecen sus vecinos cercanos K entre los datos del experimento.

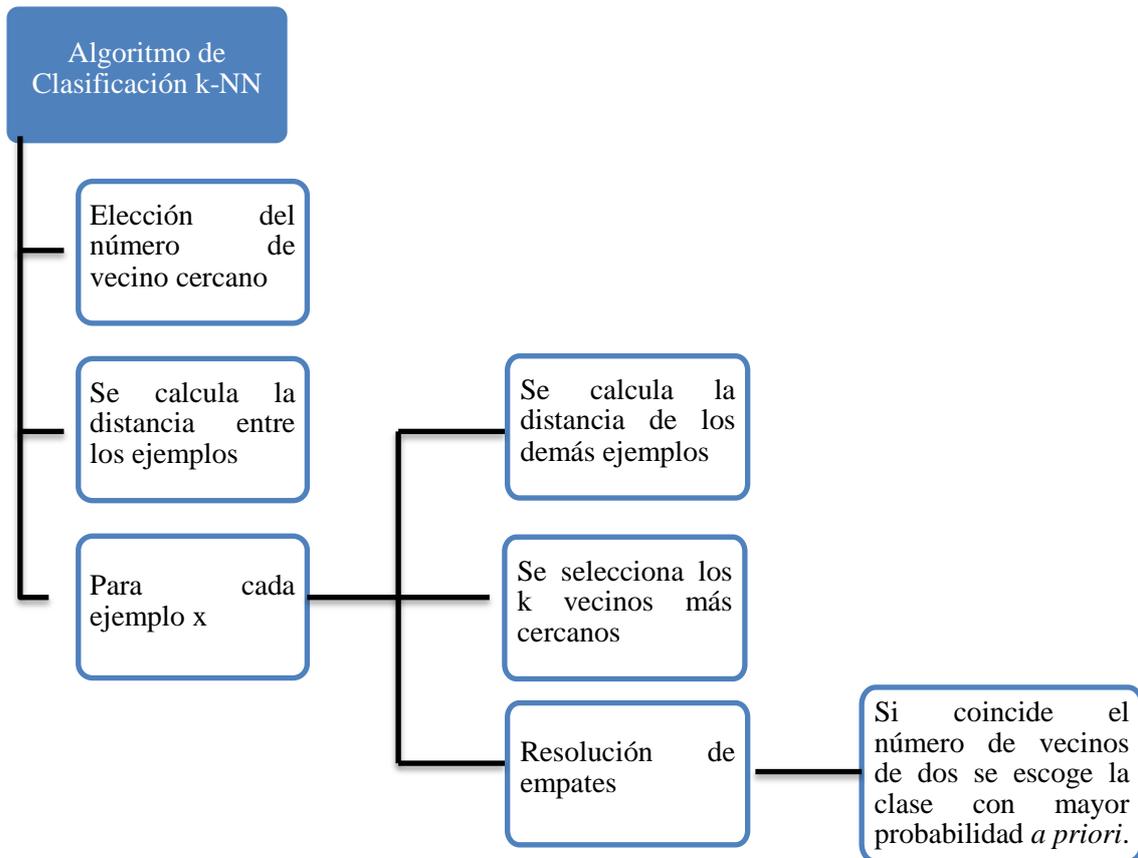
Para la elección del k se debe considerar los siguientes casos:

- Si k es muy pequeño el modelo clasificador toma importancia a los datos atípicos, esto se debe a la presencia de ruido.
- Si k es muy grande, el modelo tiende a clasificar a la clase más grande, se consideran más vecinos y el ruido desaparece.

El valor de K fue de 10 después de realizar una serie de pruebas de validación donde la clasificación sea adecuada. En la Figura 7-2, se muestra la clasificación cuando  $K=1$ , la muestra a clasificar pertenece a la clase roja debido a mayor número de vecinos cercanos k, mientras si  $k=3$ , hay mayor cantidad de vecinos cercanos azules por lo tanto la muestra pertenece a la clase azul.



**Figura 16-2.** Esquema de Clasificación Algoritmo k-NN  
Fuente: (SANCHO, 2014)



**Figura 17-2.** Pasos para realizar el Algoritmo de Clasificación k-NN  
Realizado por: Daniela Barragán

## CAPÍTULO III

### MARCO DE RESULTADOS, DISCUSIÓN Y ANÁLISIS RESULTADOS

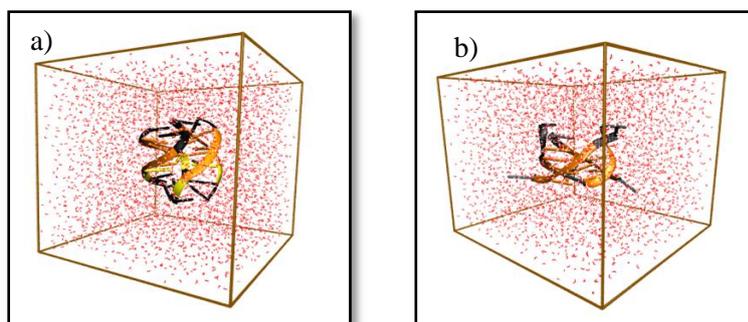
En este capítulo se muestran los resultados obtenidos, se ha distribuido en 5 secciones donde la primera sección y segunda sección son los resultados de la preparación del modelo donde el campo de fuerza CHARMM27 simula los sistemas de manera adecuada. Además se presenta gráficas de comportamiento de energía, temperatura y presión.

En la Tercera sección se realiza un análisis mediante histogramas en el cual se identifica conformaciones estructurales presentes a lo largo de la trayectoria. Posteriormente, en la cuarta sección se evalúa la información adquirida a partir de las distancias de los puentes de hidrógeno en cada G-tétrada empleándose mapas de calor para su visualización.

Finalmente, en la quinta sección se presenta la evaluación de los resultados obtenidos vía simulación con respecto a valores experimentales de técnicas RMN y Cristalografía.

#### 3.1. Preparación de Modelos a simular

La construcción de la caja de agua en los modelos permitió simular el comportamiento del ADN frente a un solvente. Las dimensiones de la caja para el modelo 1NP9 son 50.89x51.19x56.35 Å y del monómero 2GW0 son 54.30x54.29x45 Å

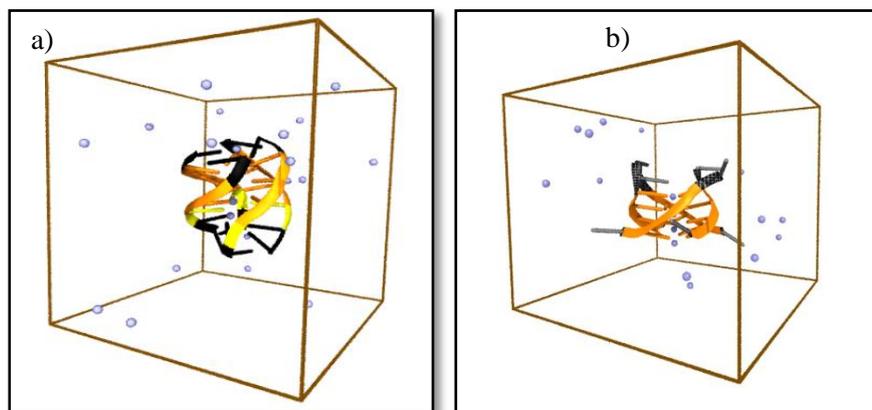


**Figura 1-3.** Resultado de solvatación.

a) G-Cuádruple 1NP9 b) G-Cuádruple del Monómero 2GW0

Realizado por: Daniela Barragán

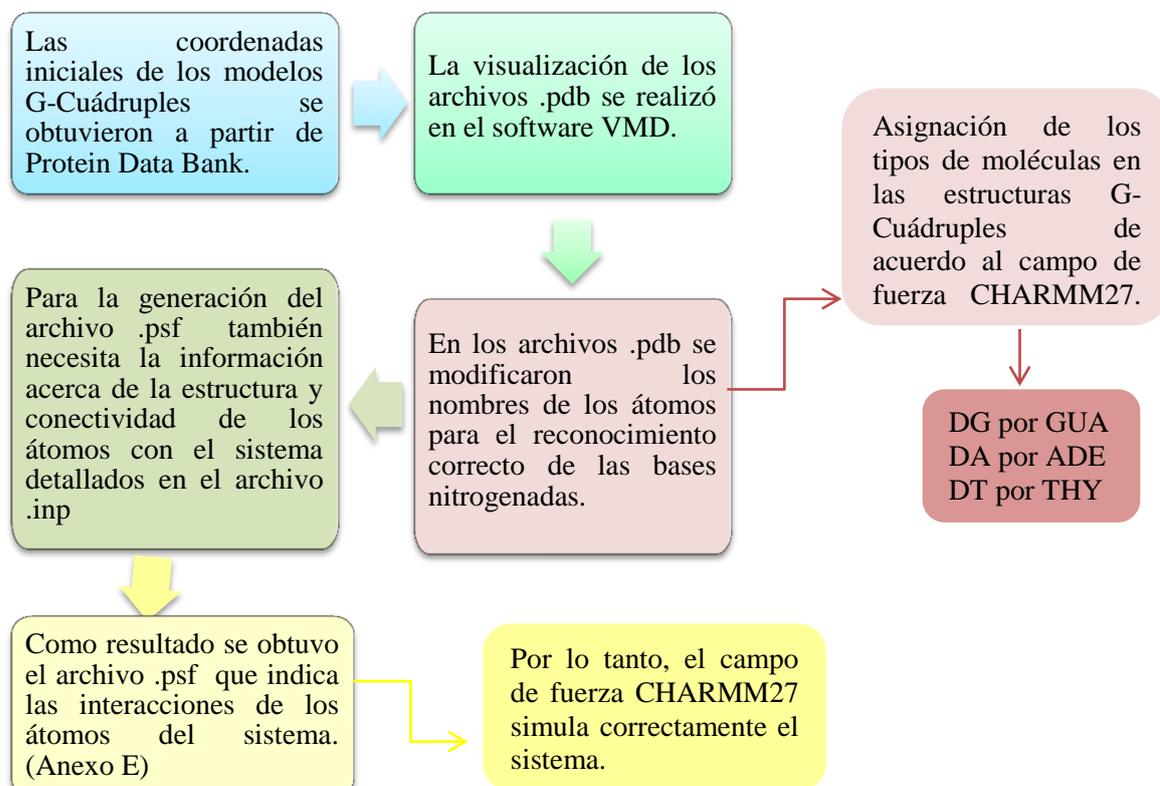
Debido a la presencia de cargas negativas presentes en los fosfatos de la molécula de ADN, el uso de la neutralización da como resultado la adición de iones a la caja.



**Figura 2-3.** Resultado de la neutralización. Los iones  $\text{Na}^+$  y  $\text{Cl}^-$  están representados de color morado a) G-Cuádruple 1NP9 b) G-Cuádruple del Monómero 2GW0

Realizado por: Daniela Barragán

### Generación del archivo .psf para los modelos de G-Cuádruples



**Figura 3-3.** Esquema para generar un archivo .psf

Realizado por: Daniela Barragán

### 3.2. Análisis de la Preparación del modelo

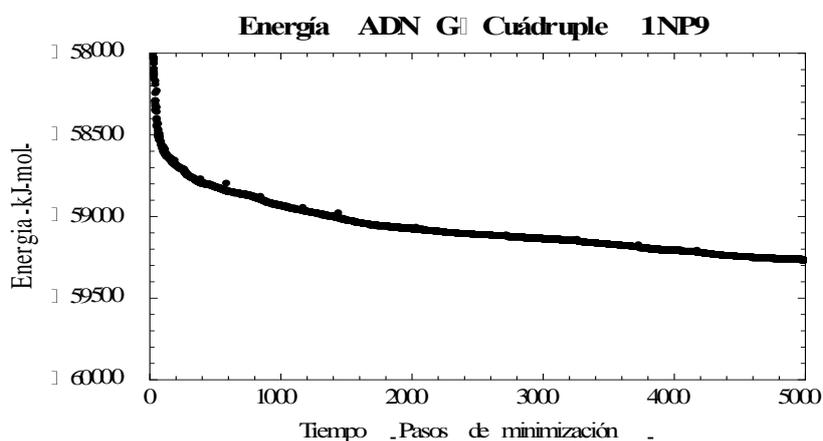
#### Minimización

En la etapa de minimización, las moléculas del sistema junto con el solvente se relajaron eliminándose impedimentos estéricos creados por la posición aleatoria de los átomos de hidrógeno durante la generación del psf.

El sistema fue sometido a un algoritmo de minimización adecuado para explorar la superficie de energía potencial y así encontrar, en lo posible, un mínimo energético de la estructura. Durante la minimización los iones se movieron ligeramente de su posición pero en general permanecieron en el centro del canal. Estos iones son indispensables para la estabilidad de la molécula, debido a que su ausencia provoca que la estructura se modifique y desestabilice. En el anexo F se muestra el archivo de configuración NAMD usado para minimizar la energía potencial de los sistemas.

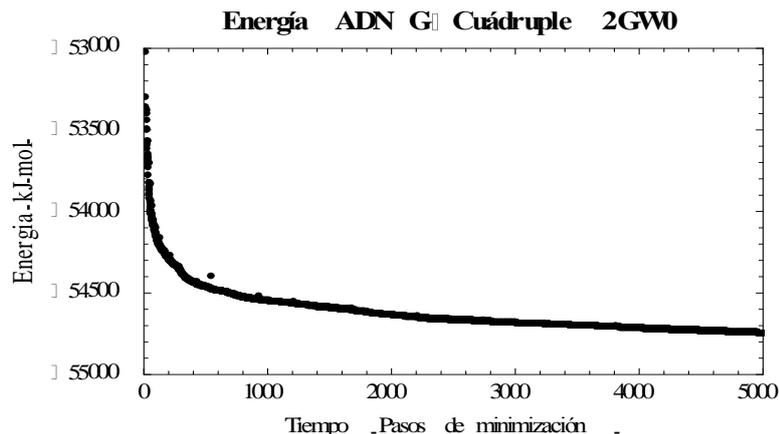
En la segunda etapa de minimización, no se emplean restricciones en el sistema por lo que el ADN puede explorar sus conformaciones en conjunto con el solvente.

Las figuras 4-3 y 5-3 muestran la disminución de energía potencial del sistema durante los pasos de minimización en la segunda etapa de minimización hasta llegar a un mínimo energía. Los valores de energía promedio son: -59263.73 Kcal/mol y -54742.45 Kcal/mol respectivamente.



**Figura 4-3.** Disminución de la Energía del sistema ADN G-Cuádruple 1NP9

Realizado por: Daniela Barragán



**Figura 5-3.** Disminución de la Energía del sistema ADN G-Cuádruple del monómero 2GW0

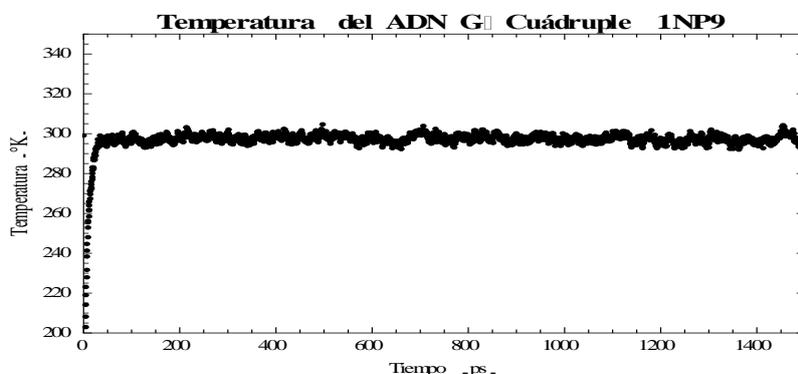
Realizado por: Daniela Barragán

### Termalización

Luego de la construcción de la caja y de las etapas de minimización se procede a calentar al sistema sometido al ensamble canónico (NVT), el cual controla la temperatura. Se presentan resultados del incremento gradual de la temperatura y la energía potencial en función del tiempo hasta alcanzar el valor constante de la simulación.

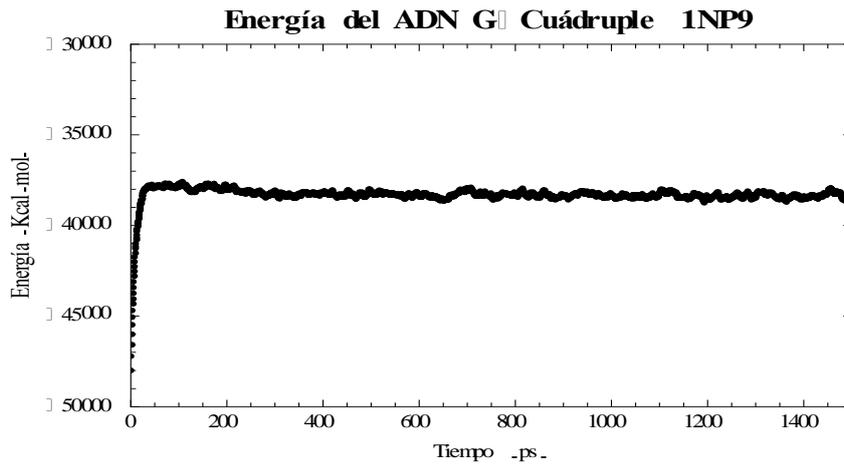
En la termalización se llega a una temperatura de 298 K para los sistemas con una desviación estándar de 7.34 K para la estructura 1NP9 y para el monómero 2GW0 7.52 K. Estos resultados se muestran en las figuras 6-3 y 8-3. El termostato Langevin mantuvo la temperatura constante.

La energía potencial de la estructura 1NP9 G-Cuádruple tiene una media de  $-38312.1 \pm 537.629$  kcal/mol. (Gráfica 7-3) La energía potencial del monómero 2GW0 llega a ser constante en una media de  $-35928.05 \pm 476.50$  kcal/mol. (Gráfica 9-3)



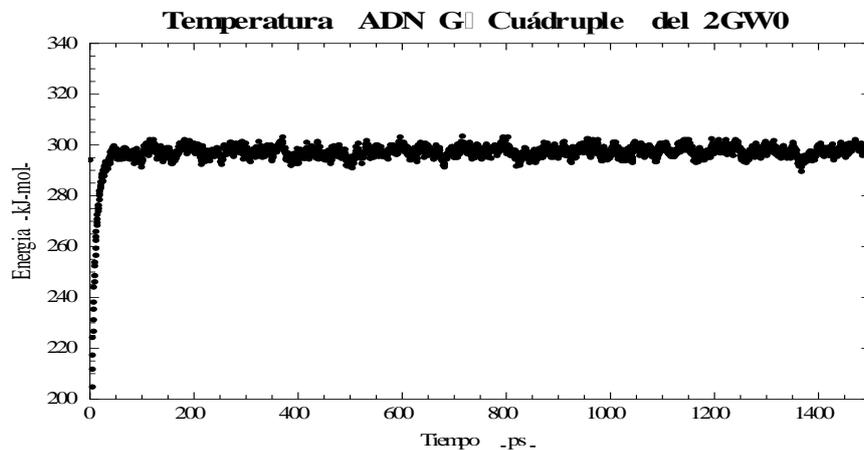
**Figura 6-3.** Temperatura de Termalización del sistema ADN G-Cuádruple 1NP9

Realizado por: Daniela Barragán



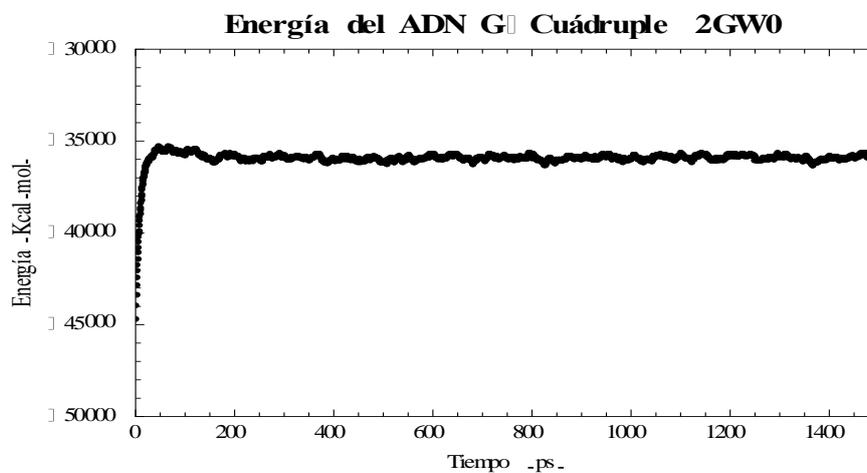
**Figura 7-3.** Energía de Termalización del sistema ADN G-Cuádruple 1NP9

Realizado por: Daniela Barragán



**Figura 8-3.** Temperatura de Termalización del sistema ADN G-Cuádruple del Monómero 2GW0

Realizado por: Daniela Barragán

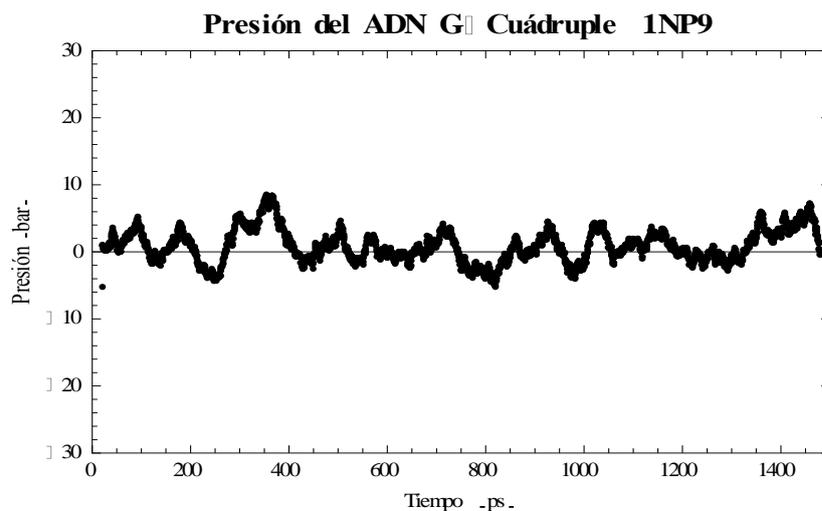


**Figura 9-3.** Energía de Termalización del sistema ADN G-Cuádruple 2GW0

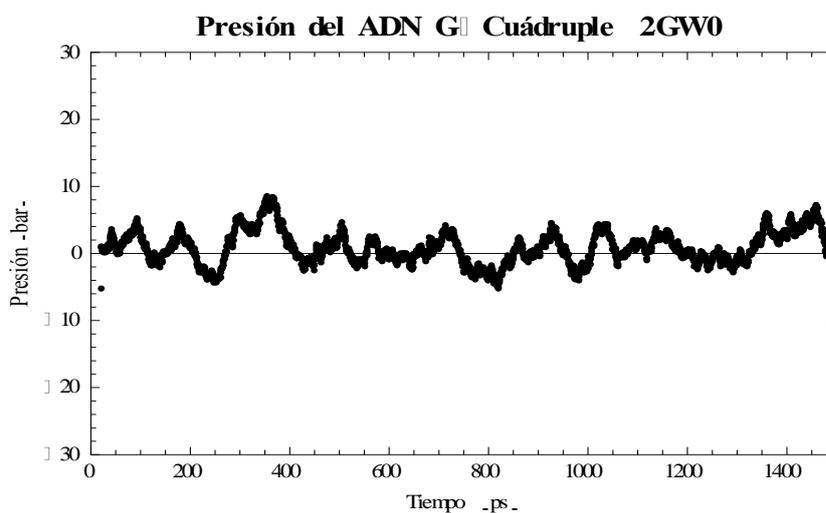
Realizado por: Daniela Barragán

## Optimización del Volumen

El ensamble isotérmico-isobárico (NPT) se empleó para mantener constante la presión y además optimizar el volumen del sistema (al ser variable en este ensamble). Se obtuvieron los siguientes resultados en el modelo 1NP9, el valor de la presión promedio es  $0,68 \pm 0,21$  bar y del monómero es  $0,88 \pm 0,22$  bar. Estos valores se muestran en las gráficas 10-3 y 11-3.



**Figura 10-3.** Presión de Termalización del sistema ADN G-Cuádruple 1NP9  
Realizado por: Daniela Barragán



**Figura 11-3.** Presión de Termalización del sistema ADN G-Cuádruple del Monómero 2GW0  
Realizado por: Daniela Barragán

### 3.3. Conformaciones Estructurales

#### 3.3.1. Conformaciones Estructurales de la estructura 1NP9

En el sistema tetramolecular paralelo con secuencia 5'-d (TTAGGGT)-3' se observó:

- Cambios importantes en el esqueleto sin necesariamente cambiar la estabilidad de los cuartetos de Guanina en la estructura.
- Una gran variabilidad conformacional en la posición de las bases nitrogenadas que se encuentran a los extremos de la molécula que no intervienen en la formación de los G-cuartetos.
- Las estructuras de transición presentan cambios importantes incluso en las bases nitrogenadas que interactúan por medio de puentes de hidrógeno.

En la tabla 1-3 se muestran las diferentes conformaciones estructurales estables y de transición encontradas mediante el análisis de los valores obtenidos por RMSD y Radio de Giro, evaluadas en los residuos de Guanina y la estructura completa del G-Cuádruple.

**Tabla 1-3:** Conformaciones Estables y de Transición de la estructura 1NP9 G-Cuádruple Tetramolecular Paralelo

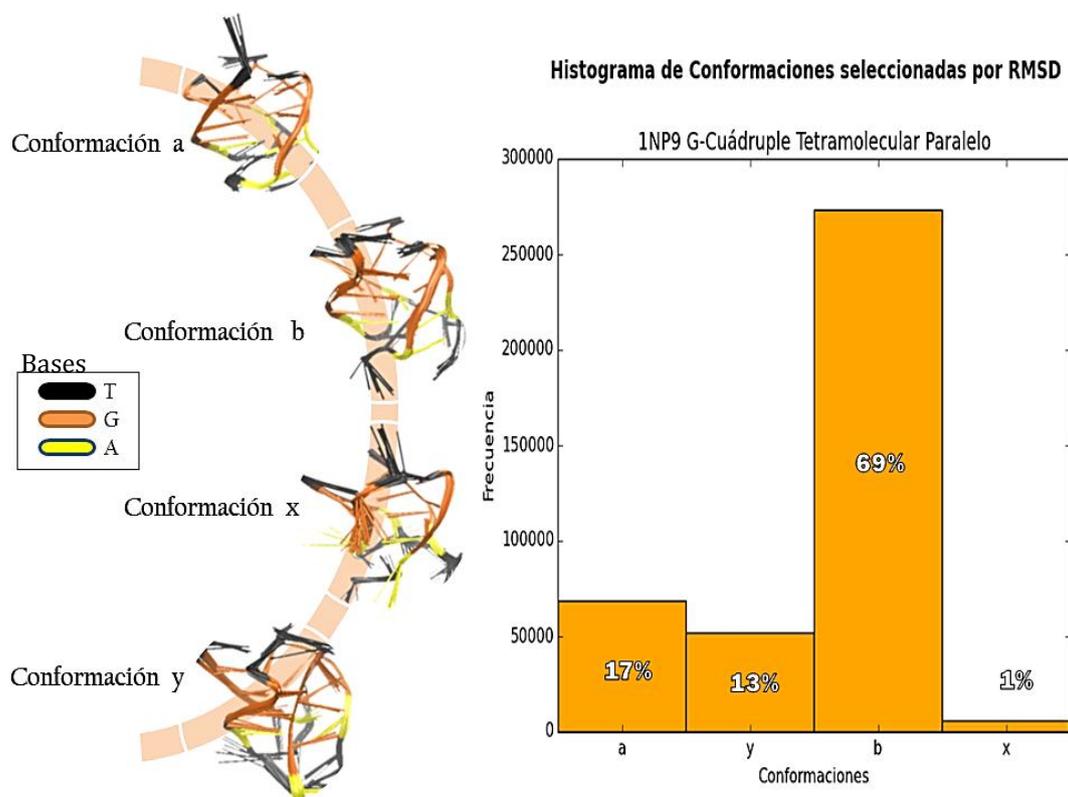
G-Cuádruple Tetramolecular Paralelo		1NP9	
		Estables	Transición
RMSD	Guaninas	2	2
	G-Cuádruple	4	2
Radio de Giro	Guaninas	4	1
	G-Cuádruple	4	1

Realizado por: Daniela Barragán

Las conformaciones seleccionadas se muestran en histogramas obtenidos a partir del algoritmo k-NN, además cada conformación tiene una representación gráfica de los movimientos conformacionales observadas. El análisis en las simulaciones vía dinámica molecular muestra en el histograma una mayor presencia de conformaciones estables que de las conformaciones de transición.

### Histograma de las conformaciones seleccionadas por RMSD para *solo* en los residuos de guaninas de la estructura 1NP9 G-Cuádruple Tetramolecular Paralelo

En la figura 12-3 se detalla las conformaciones estructurales seleccionadas por RMSD. Las conformaciones estables se identifican como “a” y “b” y se distingue desplazamientos en los residuos de la Timina. Estas bases nitrogenadas de Timinas se caracterizan por tener más libertad de movimiento debido que no forman parten de la estructura de los G-cuartetos. Además, las bases de adenina y guanina de estas conformaciones no presentan grandes desplazamientos de sus bases por la presencia de puentes de hidrógeno. La conformación de transición denotada con “x” tiene mayor variabilidad conformacional en las bases nitrogenadas. Estas bases se ubican en posiciones externas al esqueleto del ADN y tienen mayores desplazamientos. En la conformación de transición, denotada con “y”, los residuos de adenina y timina poseen mayor movimiento, cambiando la ubicación de las bases con respecto el esqueleto del ADN. El algoritmo de clasificación k-NN aplicado durante todas la trayectoria obtuvo que la conformación estable “a” está presente en un 17%, la conformación estable “b” está en un 64% y en menor proporción las conformaciones de transición “x” presente en 1% y la conformación “y” en un 13%.



**Figura 12-3.** Histograma de Clasificación de Conformaciones seleccionadas por RMSD para *solo* de los residuos de Guaninas del 1NP9 G-Cuádruple Tetramolecular Paralelo

Realizado por: Daniela Barragán

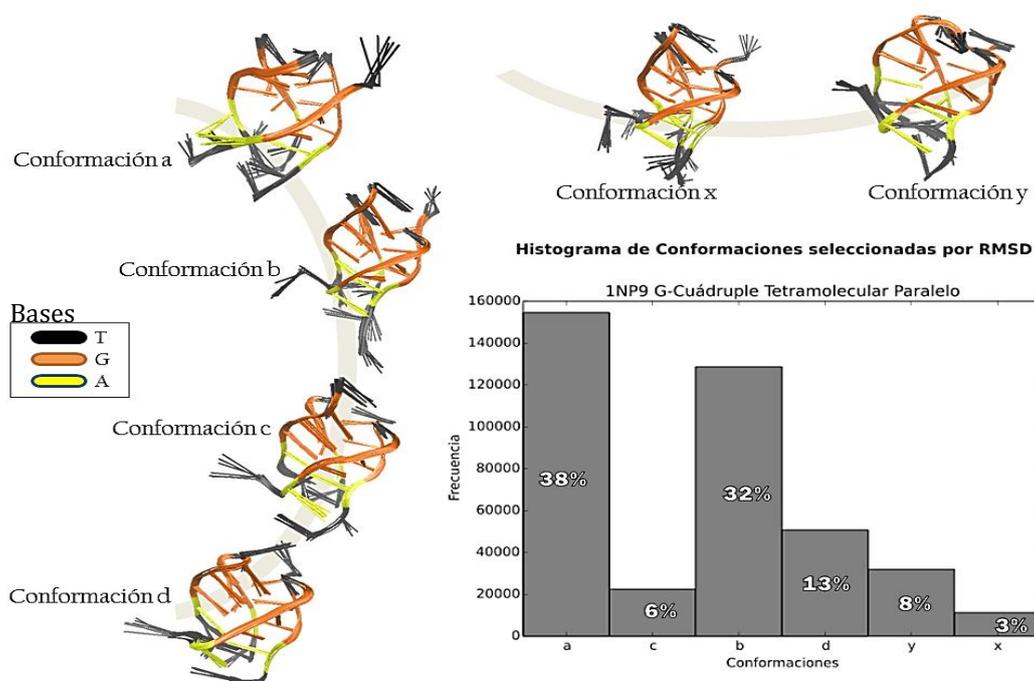
### Histograma de las conformaciones seleccionadas por RMSD para el G-Cuádruple *completo* del 1NP9 G-Cuádruple Tetramolecular Paralelo

En la figura 13-3 se describe las conformaciones estructurales de la siguiente manera:

Las conformaciones estables se diferencian por los movimientos de los residuos de timinas y el esqueleto que adquiere diversas formas según la posición de las bases. En la conformación estable “a” se observa que los residuos de adenina no tienen una posición paralela con respecto a las demás bases de la cadena sin embargo en las demás conformaciones no se observa este cambio. En general, la conformación “a” y la conformación estable “b” se caracterizan porque no se presentan cambios en los G-cuartetos que confieren estabilidad a la estructura global.

La conformación de transición “x” se identifica por un importante desplazamiento de una de las cadenas causando movimientos en las bases de timinas y adeninas. La conformación de transición “y”, a diferencia de la otra estructura de transición, no presenta desplazamientos en los G-cuartetos y las bases de adeninas. Esta estructura podría considerarse que no fue afectada por los cambios estructurales; sin embargo, las bases de los extremos de la estructura si tienen un desplazamiento significativo.

En el histograma se presenta 4 conformaciones estables y 2 de transición; encontrándose en un 38% de la conformación estable “a” y en menor cantidad, la conformación de transición x presente en un 3%.



**Figura 13-3.** Histograma de Clasificación de Conformaciones seleccionadas por RMSD para el G-Cuádruple *completo* de la estructura 1NP9 G-Cuádruple Tetramolecular Paralelo

Realizado por: Daniela Barragán

En el **Anexo G** Histograma de Clasificación de Conformaciones seleccionadas por Radio de Giro para *solo* los residuos de Guaninas del 1NP9 G-Cuádruple Tetramolecular Paralelo

En las conformaciones seleccionadas se observa que las tétradas de guanina son invariantes, ya que las bases de adenina tienen ligeros cambios. En verdad estas conformaciones se diferencian entre sí por el desplazamiento de las bases de timina. La estructura de transición “x” muestra inestabilidad en toda la estructura observándose diversos movimientos fuera del esqueleto de la cadena. Siguiendo los criterios de clasificación se encontraron 4 conformaciones estables y 1 de transición. El algoritmo de clasificación k-NN aplicado sobre toda la trayectoria encontró que donde la conformación “b” está en un 45% con respecto a las demás y la de transición está presente en 1% en la simulación.

En el **Anexo H** Histograma de Clasificación de Conformaciones seleccionadas por Radio de Giro para el G-Cuádruple *completo* del 1NP9 G-Cuádruple Tetramolecular Paralelo

En este caso las conformaciones estables se diferencian por la posición de las bases ubicadas en los extremos de la cadena y por desplazamientos en las bases de adenina. En la conformación estable “a” la base de adenina adopta una posición fuera del esqueleto mientras que en la conformación “b” existe un desplazamiento en la base de adenina de una de las cadenas.

Las conformaciones estables “c” y “d” se caracterizan porque las bases de la estructura G-Cuádruple son constantes con respecto a una posición inicial pero también se observa ligeros desplazamientos en las bases de timina. Las conformaciones estables mencionadas presentan estabilidad en los G-cuartetos.

En la estructura de transición “x” se observa que 3 de las cadenas sufren cambios estructurales afectando en su totalidad a todas las bases nitrogenadas.

El histograma de clasificación de las conformaciones de esta estructura indica que en mayor proporción está presente la conformación “b” en 56% y la conformación más baja fue la “x” en 1%.

### 3.3.2. Conformaciones Estructurales del Monómero 2GW0

En el sistema tetramolecular paralelo con secuencia 5'- d(TGGGGT)-3' se muestra que:

- Existe una gran estabilidad de los cuartetos de Guanina en la estructura, por otra parte, las bases nitrogenadas que se encuentran en los extremos presentan variabilidad conformacional.
- Las estructuras de transición presentan un movimiento más libre en las bases timinas y también se observa variabilidad en algunas bases de las tétradas de guaninas.

En la tabla 2-3 se muestran las diferentes conformaciones estructurales estables y de transición encontradas mediante el análisis de los valores obtenidos por RMSD y Radio de Giro; evaluadas *solo* en los residuos de Guanina y en la estructura del G-Cuádruple *completo*.

**Tabla 2-3:** Conformaciones Estables y de Transición de la estructura 2GW0 G-Cuádruple Tetramolecular Paralelo

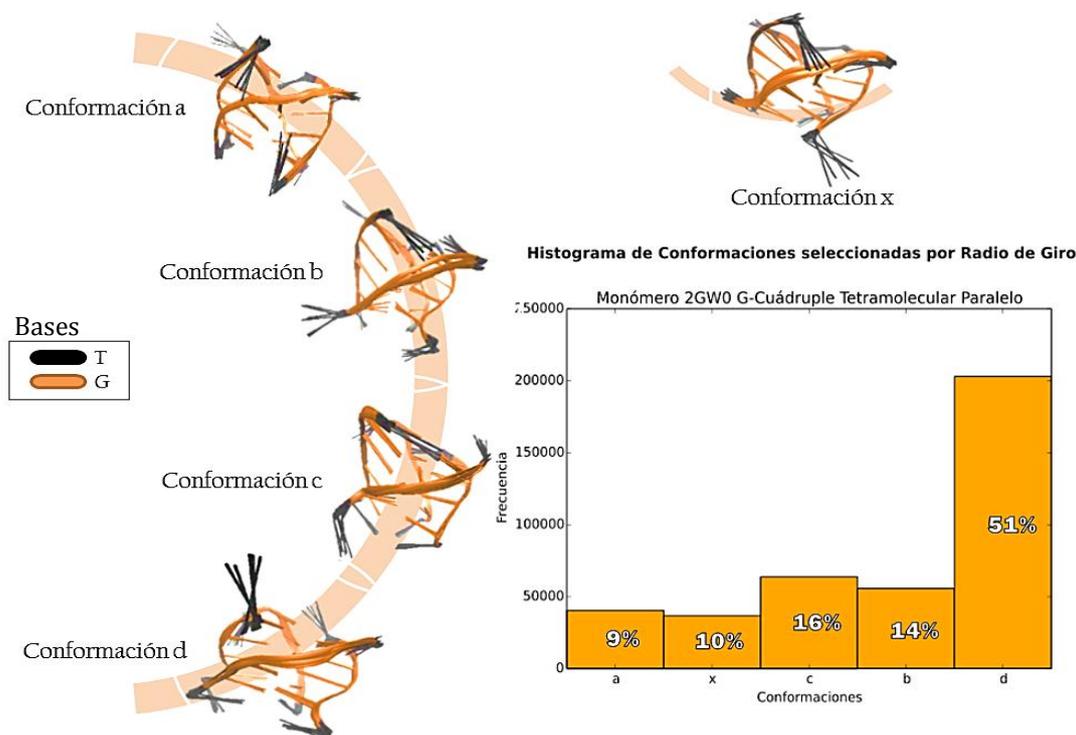
G-Cuádruple Paralelo	Tetramolecular	Monómero 2GW0	
		Estables	Transición
RMSD	Guaninas	4	2
	G-Cuádruple	3	1
Radio de Giro	Guaninas	4	1
	G-Cuádruple	4	1

Realizado por: Daniela Barragán

#### Histograma de conformaciones seleccionadas por Radio de Giro para *solo* los residuos de guaninas del monómero 2GW0 G-Cuádruple Tetramolecular Paralelo

En la figura 14-3 se presenta las conformaciones estructurales clasificadas por el algoritmo k-NN. Las conformaciones estructurales son distintas por la posición de las bases nitrogenadas y la disposición del esqueleto de la estructura. Las bases de los G-cuartetos en las conformaciones estables mantienen una disposición invariable sin embargo en la conformación “b” se observa un desplazamiento de una de las bases de la cadena. Además, los residuos de las timinas presentan variación conformacional en todas las estructuras. En la estructura de transición “x” se encontró que varias de las bases que conforman tétradas de guaninas son afectadas por variación conformacional.

En el histograma se muestra que la conformación estable en mayor proporción es la “d” con 51% mientras que la conformación de transición representa el 10% de las conformaciones en la simulación.

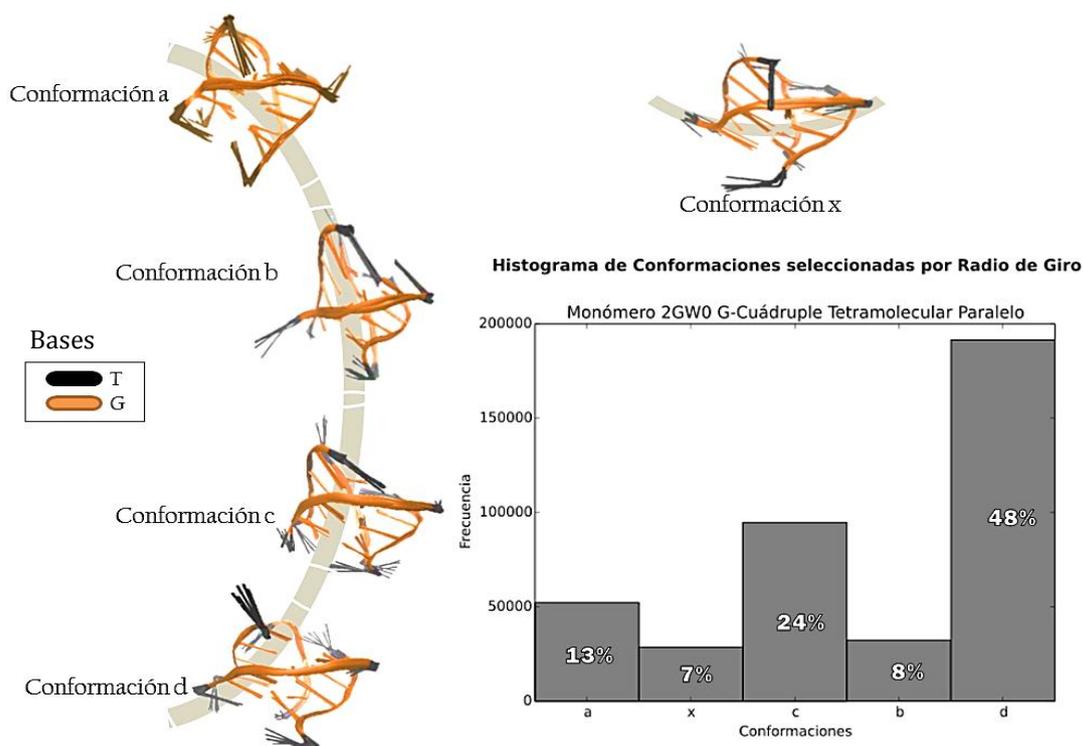


**Figura 14-3.** Histograma de Clasificación de Conformaciones seleccionadas por Radio de Giro para *solo* los residuos de Guanina del monómero 2GW0 G-Cuádruple Tetramolecular Paralelo

Realizado por: Daniela Barragán

### Histograma de conformaciones seleccionadas por Radio de Giro para el G-Cuádruple completo del monómero 2GW0 G-Cuádruple Tetramolecular Paralelo

Las conformaciones estables presentan movimientos en las bases de timina por encontrarse en los extremos de la estructura sin embargo los G-cuartetos mantienen su conformación estructural. Las conformaciones de transición se caracterizan por desplazamientos en algunas bases nitrogenadas de la estructura. Se seleccionaron 4 conformaciones estables y una de transición; encontrándose en un 48% la conformación estable “d” y en menor proporción la conformación de transición “x” presente en 7% (Gráfica 15-3) En general, las conformaciones de esta estructura presentan bastante estabilidad en las G-tétradas incluso en la conformación de transición demostrando que la presencia de las bases de guanina cumplen el rol de estabilización de la estructura.



**Figura 15-3.** Histograma de Clasificación de Conformaciones seleccionadas por Radio de Giro para el G-Cuádruple *completo* del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo

Realizado por: Daniela Barragán

**Anexo I** Histograma de Clasificación seleccionadas por RMSD de *solo* los residuos de Guanina del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo

Las posiciones de las bases que se encuentran en los extremos de las cadenas son distintas en cada conformación. La conformación “a” y “b” presentan un desplazamiento en una base de guanina de la misma cadena. En la conformación “c” se observa que las G-tétradas son estables. La conformación “d” es la estructura más estable en comparación a las anteriores. La conformación “x” muestra desplazamiento de la timina en una cadena mientras que las demás bases nitrogenadas son estables. La conformación “y” muestra una de las cadenas donde se observan cambios indicativos de flexibilidad en una base de las guaninas.

Este histograma muestra 4 conformaciones estables y 2 conformaciones transición. La conformación “d” está presente en un 46% y la de transición “x” en un 5% .

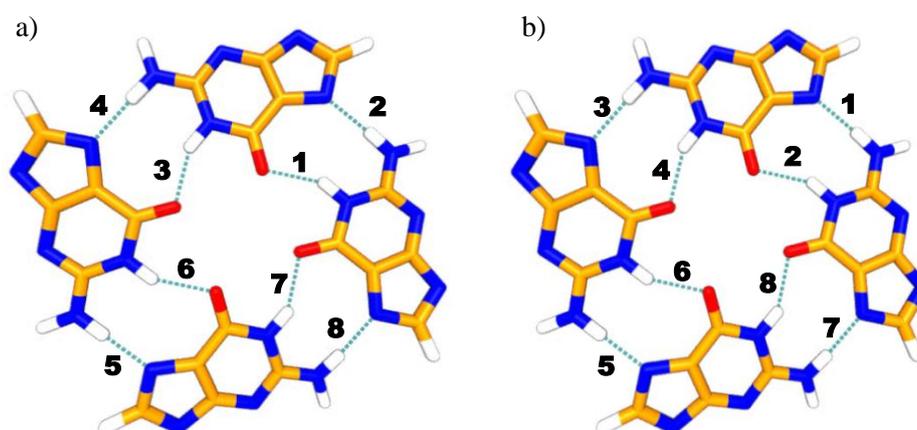
## Anexo J Histograma de Clasificación seleccionadas por RMSD para el G-Cuádruple *completo* del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo

Las conformaciones estables y de transición se diferencian por las bases nitrogenadas que no forman parte de los G-cuartetos. Las conformaciones “a” y “b” de algunas guaninas tienen variación en sus posiciones con respecto a su estructura de referencia. La conformación “c” presenta estabilidad en los G-cuartetos. La conformación “x” tiene desplazamientos en las bases de los G-cuartetos notándose mayor libertad de movimiento

El histograma muestra la clasificación 3 conformaciones estables y 1 de transición. Los resultados más representativos fueron de las conformaciones más estable fue “b” presente en un 37% que la conformación mostró menor presencia es “x” en 22%.

### 3.4. Análisis de las distancias de los puentes de Hidrógeno

Los residuos de guanina se unen con sí mismas mediante puentes de hidrógeno del tipo Hoogsteen que pueden ser de dos tipos N-H··N y O··H-N. El rango de longitud de los puentes de Hidrógeno es de  $2.4 < d < 4.5$  Å. En los mapas de calor se muestra los valores de las distancia de la conformación estable y de transición respectivamente presentes en mayor proporción como se determina por k-NN en base a los parámetros de RMSD y Radio de Giro. En la figura 16-3 la componente del eje “x” representa 25 valores de longitud de puentes de hidrógeno de la conformación en estudio. La componente “y” indica la etiqueta de los puentes de hidrógeno de las estructuras.

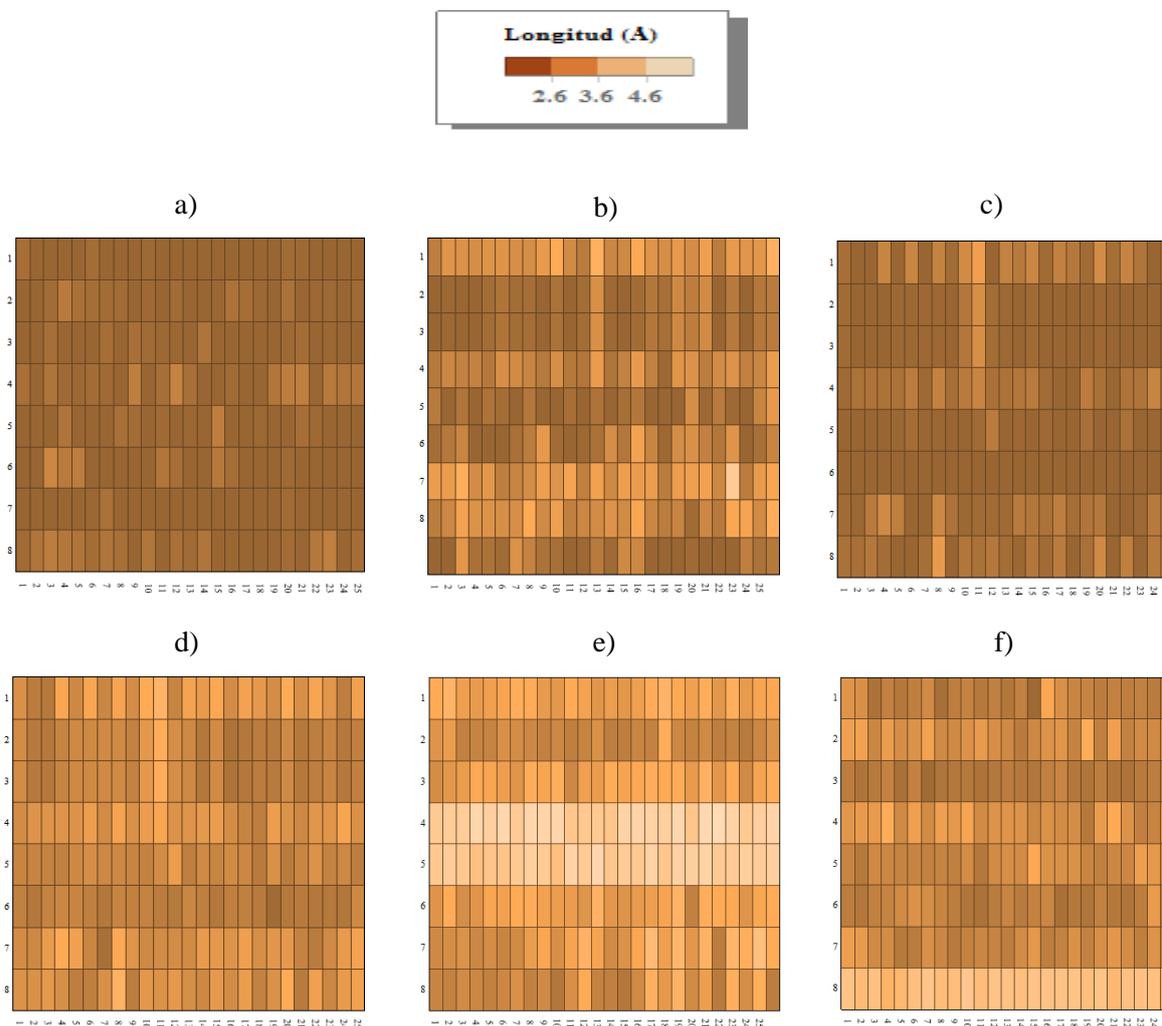


**Figura 16-3.** Esquema de enlaces de Hidrógeno del G-cuarteto de los sistemas.

a) Sistema 1NP9 G-Cuádruple b) Sistema Monómero 2GW0 G-Cuádruple

Realizado por: Daniela Barragán

### 3.4.1. Mapas de Calor de la estructura INP9



**Figura 17-3.** Mapas de Calor de la estructura INP9 G-Cuádruple.

a) Cuarteto 1 de la Conformación estable “b” b) Cuarteto 2 de la Conformación estable “b” c) Cuarteto 3 de la Conformación estable “b” d) Cuarteto 1 de la Conformación transición “x” e) Cuarteto 2 de la Conformación transición “x” f) Cuarteto 3 de la Conformación transición “x”. Conformaciones pertenecen al parámetro RMSD para *solo* los Residuos de Guanina

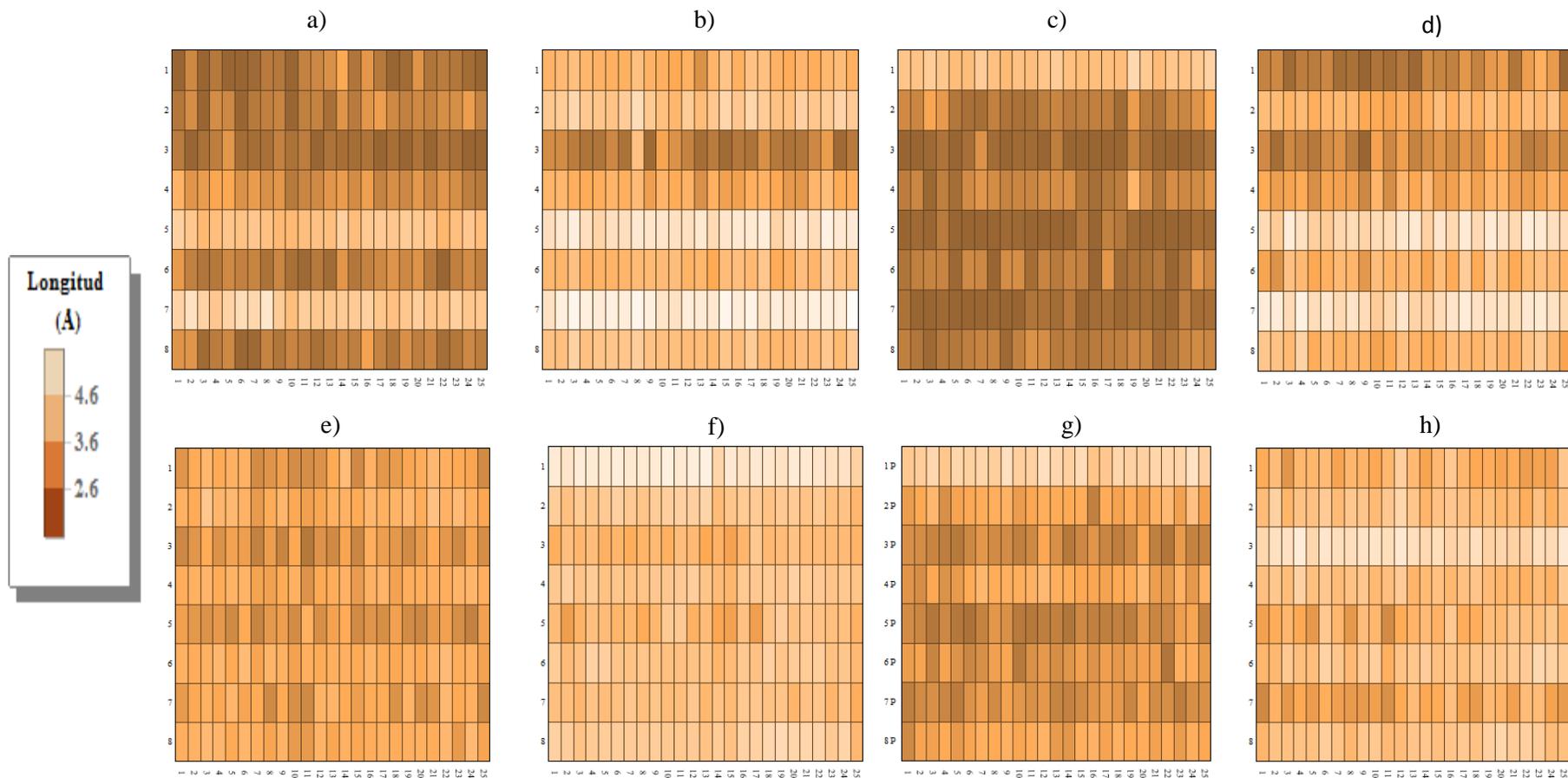
**Realizado por:** Daniela Barragán

En el mapa de calor de la estructura INP9 G-Cuádruple Tetramolecular Paralelo, el color café indica menor longitud de puente de hidrógeno y el color blanco indica mayor longitud de los puentes de hidrógeno. Una menor longitud del puente de hidrógeno, cercana a 2.4 Å representa un enlace más fuerte, esta longitud se puede asociar con la estabilización de la estructura por sus enlaces de hidrógeno.

En la figura 17-3, los mapas de calor de las estructuras estables presentan distancias desde 1.6 hasta 3.6 Å mientras que la estructuras de transición presentan distancias desde 2.2 hasta 5.6 Å. Además al comparar la estructura de transición con la estructura estable presentan un color blanco en tres enlaces de hidrógeno que se encuentran en los extremos de la G-tétrada debido a mayor libertad de movimiento.

En los **Anexos K, L, M** se muestran los mapas de calor calculados para conformaciones estables y de transición seleccionadas a partir del RMSD del G-Cuádruple *Completo*, Radio de Giro de solo los residuos de Guanina y Radio de Giro de G-Cuádruple Completo respectivamente, los cuales presentan un comportamiento similar a lo anteriormente señalado.

3.4.2. Mapas de Calor de la estructura 2GW0



**Figura 18-3.** Mapas de Calor del Monómero 2GW0 G-Cuádruple.

a) Cuarteto 1 de la Conformación estable “a” b) Cuarteto 2 de la Conformación estable “a” c) Cuarteto 3 de la Conformación estable “a” d) Cuarteto 4 de la Conformación estable “a” e) Cuarteto 1 de la Conformación transición “x” f) Cuarteto 2 de la Conformación transición “x” g) Cuarteto 3 de la Conformación transición “x” h) Cuarteto 4 de la Conformación transición “x”. Conformaciones pertenecen al parámetro Radio de Giro del G-Cuádruple *Completo*.

Realizado por: Daniela Barragán

Los mapas de calor del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo muestran, un aumento en la longitud de los enlaces de hidrógeno en comparación con la estructura 1NP9; esto se distingue por el color tomate claro. Por otro parte, al comparar las conformaciones estables con las conformaciones de transición, se observa que mantienen el color café y tomate en la mayoría de los patrones, lo que permite concluir que las estructuras no presentan una marcada variación conformacional durante la simulación.

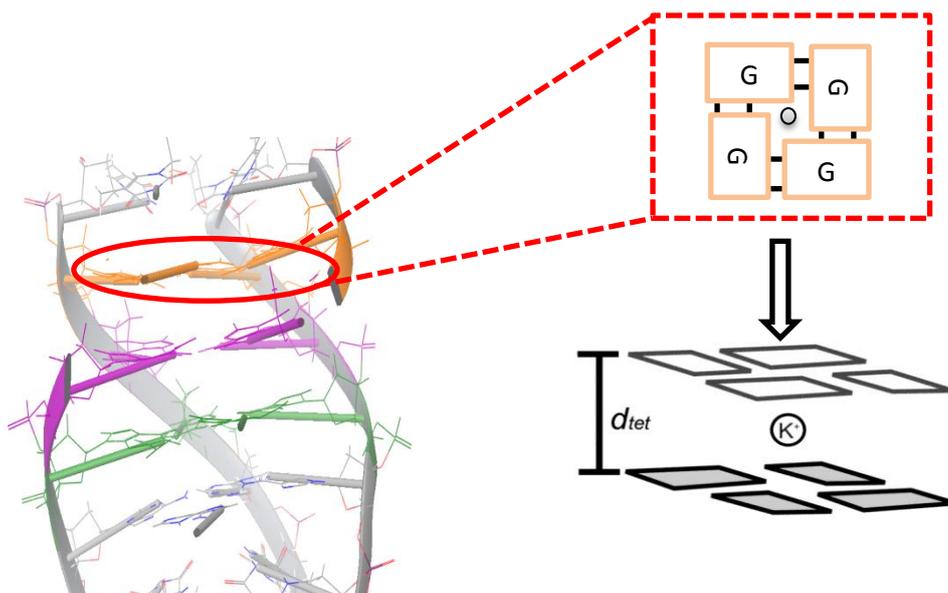
Otro aspecto a considerar en las conformaciones estables, es que existen longitudes de enlace mayores a 4.6 Å en puentes de hidrógeno ubicados en la parte externa (5 y 7), lo que se espera debido al mayor grado de libertad que se observa en esas regiones.

En los **Anexos N, O** se muestran los mapas de calor calculados para conformaciones estables y de transición seleccionadas a partir del RMSD de *solo* los residuos Guaninas, RMSD del G-Cuádruple *Completo* y Radio de Giro de *solo* los residuos de Guaninas respectivamente, los cuales presentan un comportamiento similar al análisis anteriormente mencionado.

### **3.5. Comparación de los resultados de simulación versus los datos experimentales**

En este análisis se realiza la comparación de los resultados vía simulación con el campo de fuerza CHARMM27 versus los datos experimentales obtenidos por técnicas de RMN y Cristalografía. Las estructuras G-Cuádruples se caracterizan por ser ricas en secuencias de guanina que son capaces de formar tétradas.

Las tétradas (cuartetos) de guanina se apilan una sobre otra para formar estas estructuras. (Figura 19-3) El modelo 1NP9 tiene tres G-tétradas y el monómero 2Gw0 tiene cuatro G-tétradas. Se calcularon las distancias entre los cuartetos de cada sistema simulado, como también las distancias entre cuartetos de estructuras experimentales con el objetivo de evaluar el campo de fuerza CHARMM27.



**Figura 19-3** Esquema de medición distancia entre cuartetos de residuos de guanina  
**Realizado por:** Daniela Barragán

En la Tabla 3-3 se muestra el cálculo de distancias promedio  $d_{tet}$  entre pisos de técnicas experimentales.

**Tabla 3-3:** Cálculo de distancias de Estructuras Experimentales

Promedio de Distancia Experimental (Å)	
RMN (1NP9)	$3,59 \pm 0,08$
Cristalografía Dímero (2GW0)	$3,36 \pm 0,34$
Cristalografía Monómero (2GW0)	$3,27 \pm 0,11$

**Realizado por:** Daniela Barragán

En la tabla 4-3, se compara valores de distancias  $d_{tet}$  entre los cuartetos de G-tétradas encontrados en las simulaciones realizadas versus los valores experimentales. Se observa un buen arreglo entre las distancias promediadas durante la simulación de dinámica molecular para el modelo 1NP9 y el promedio de distancia en la estructura de RMN. Sin embargo, sólo una conformación sobrepasa el 10% de error y corresponde al promedio de  $d_{tet}$  para las estructuras de transición. Además al comparar esta estructura con los promedios experimentales de los cristales 2GW0, tanto monómero como dímero, se tiene un porcentaje de error mayor, llegando a un máximo de 23,2%; esto se podría relacionar por un lado porque a las secuencias distintas (1NP9 5'-TAGGGT-3' versus 2GW0 5'-TGGGGT-3') y a la metodología de obtención de los modelos tridimensionales, que en este caso corresponde a RMN y Cristalografía de Rayos X, los cuales se realizan en fase líquida y sólida respectivamente, afectando directamente en la conformación final de la estructura.

Estos resultados sugieren que la distancia promedio  $d_{tet}$  del sistema 1NP9 es sobreestimada durante la simulación de dinámica molecular, con respecto a los valores experimentales.

**Tabla 4-3:** Comparación de distancias de los pisos de las G-Tétradas del 1NP9

<b>1NP9 G-Cuádruple Tetramolecular Paralelo</b>				
<b>CONFORMACIONES</b>	<b>PROMEDIO (Å)</b>	<b>PORCENTAJE DE ERROR</b>		
		<b>RMN (1NP9)</b>	<b>Cristalografía del Dímero (2GW0)</b>	<b>Cristalografía del Monómero (2GW0)</b>
Conformaciones Estables <sup>a</sup>	3,8±0,1	6,0	13,2	16,4
Conformaciones Estables <sup>b</sup>	3,9±0,2	9,6	17,2	20,4
Conformaciones Transición <sup>a</sup>	4,0±0,2	12,3	19,9	23,2
Conformaciones Transición <sup>b</sup>	3,9±0,2	8,9	16,4	19,6

<sup>a</sup>RMSD, <sup>b</sup>Radio de Giro

Realizado por: Daniela Barragán

En la tabla 5-3 se detalla la comparación entre el promedio de distancia  $d_{tet}$  calculado para la simulación de dinámica molecular del monómero 2GW0 donde se observa un buen ajuste con los valores experimentales (monómero y dímero 2GW0 y la estructura 1NP9); la mayoría de los valores tiene un error inferior al 10%. En este caso, los promedios de distancias  $d_{tet}$  son subestimados por el campo de fuerza debido a su bajo error refleja un buen comportamiento al reproducir dichas distancias.

**Tabla 5-3:** Comparación de distancias de los pisos de las G-Tétradas del Monómero 2GW0

<b>Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo</b>				
<b>CONFORMACIONES</b>	<b>PROMEDIO (Å)</b>	<b>PORCENTAJE DE ERROR</b>		
		<b>RMN (1NP9)</b>	<b>Cristalografía del Dímero (2GW0)</b>	<b>Cristalografía del Monómero (2GW0)</b>
Conformaciones Estables <sup>a</sup>	3,4±0,2	4,7	1,8	4,6
Conformaciones Estables <sup>b</sup>	3,2±0,2	9,2	3	0,3
Conformaciones Transición <sup>a</sup>	3,3±0,1	8	1,7	1
Conformaciones Transición <sup>b</sup>	3,6±0,1	1,5	8,5	11,5

<sup>a</sup>RMSD, <sup>b</sup>Radio de Giro

Realizado por: Daniela Barragán

El campo de fuerza reproduce de manera adecuada el comportamiento de los dos sistemas estudiados, sin embargo cuantitativamente el porcentaje de error menor se encontró en el monómero 2GW0. Esta estructura se caracteriza por tener mayor presencia de G-tétradas que son capaces de formar puentes de Hidrógeno Hoogsteen, siendo más estable que la estructura 1NP9.

## CONCLUSIONES

- En este proyecto se evaluó la modelización de estructuras de ADN G-Cuádruples aplicando el campo de fuerza CHARMM27 donde se analizó su comportamiento del punto de vista estructural mediante diversos parámetros, tales como estabilidad conformacional (RMSD y Radio de Giro), longitud de puentes de hidrógeno de Hoogsteen y distancia entre tétradas de guanina( $d_{tet}$ ).
- En la simulación de Dinámica molecular se obtuvo un total de 400 000 conformaciones, las cuales se clasificaron en conformaciones estables y de transición mediante el uso de parámetros como RMSD y Radio de Giro. Las conformaciones de transición se identificaron en las dinámicas iniciales de 20 a 60 ns siendo indicativo que las estructuras estables están presentes a partir de las siguientes dinámicas, esto sugiere que la exploración de conformaciones estables fue mayor.
- El algoritmo k-NN permitió clasificar conformaciones estructurales mediante histogramas reveló mayor porcentaje en las estructuras estables que las estructuras de transición durante toda la trayectoria. Este análisis demostró que las estructuras de G-cuádruple una vez formadas muestran variabilidad conformacional en las bases que rodean la estructura y en el esqueleto de las bases involucradas en la región de G-cuádruple creando posiblemente sitios de reconocimiento específico *in vivo*. Además se comprobó la estabilidad que confiere a la estructura global la presencia de los residuos de Guanina en las estructura G-Cuádruples.
- Para profundizar más sobre el comportamiento del campo de fuerza CHARMM27 para los sistemas en estudio se realizó un análisis de los puentes de hidrógeno de Hoogsteen, evaluando su longitud. La longitud de los puentes de hidrógeno de las tétradas de guanina, presentadas mediante mapas de calor, en su mayoría se encontraron dentro del rango  $2.4 < d < 4.5 \text{ \AA}$ , lo que permite concluir que dichos enlaces se mantienen a lo largo de las simulaciones. Se observaron fluctuaciones más altas en los puentes de hidrógeno ubicados en la parte externa de las tétradas, lo que se espera debido a un mayor grado de libertad.
- Se calcularon las distancias entre las tétradas de guanina a partir de las conformaciones identificadas mediante parámetros como RMSD y Radio de Giro. Para la interpretación de esta información de sistemas de ácidos nucleicos no convencionales, se recurrió a los datos experimentales disponibles que proporcionan la evidencia más directa sobre la estructura dominante del G-Cuádruple analizada. Las simulaciones de Dinámica Molecular de ADN G-Cuádruples tetramoleculares paralelos se reprodujeron con un margen de error menor al 10% con respecto a las estructuras estables seleccionadas. Además, en los sistemas no se evidenció

una variación importante de las distancias entre las G-tétradas, cumpliendo así un rol importante de estabilización de la estructura. El porcentaje de error de las distancias entre las tétradas de la estructura 2GW0 (secuencia 5'-TGGGGT-3') fue inferior con respecto al resultado del sistema 1NP9 (secuencia 5'-TAGGGT-3'), esto demuestra que la presencia más tétradas de guanina en la secuencia 2GW0 (4 G-tétradas versus 3 G-tétradas) confiere mayor estabilidad a la estructura.

En resumen, los resultados obtenidos muestran un buen modelamiento molecular del sistema bajo condiciones fisiológicas (presencia de solvente y iones); lo que permitiría realizar estudios futuros, más profundos utilizando el campo de fuerza estudiado por ejemplo, se podría realizar simulaciones para desarrollar fármacos para tratamiento del cáncer, el cual está estrechamente relacionada con el reconocimiento de secuencias específicas del ADN que podrían formar conformaciones de G-Cuádruple.

## RECOMENDACIONES

- La validación de un campo de fuerza en estructuras no convencionales de ADN implica un análisis más amplio es así que se recomienda evaluar:  
Parámetros helicoidales como por ejemplo la rotación por residuo y la inclinación por par de bases. Distancias atómicas como el tiempo de residencia de agua e interacciones del Solvente junto con el sistema.
- Efectuar un análisis de los iones monovalentes en el interior de la estructura G-Cuádruple con el fin de estudiar su rol en la conformación estructural
- Analizar parámetros energéticos como la energía libre del sistema, sus contribuciones intramolecular y de solvatación, energía de interacción entre bases o el potencial electrostático molecular.
- Realizar nuevas simulaciones de dinámica molecular donde se compare este proyecto con otros campos de fuerza como por ejemplo AMBER, OPLS, GROMOS, COMPASS, etc.
- Evaluar el campo de fuerza con otras topologías de G-Cuádruples.

## BIBLIOGRAFÍA

- [1] **AGUIRRE, A.** *Estudio Mecanocuántico, Docking y Dinámica Molecular de Tioazucareos como inhibidores de la proteína fucosidasa: Algoritmos para el análisis conformacional y programa para el cálculo de constantes de acoplamiento vecinales* [En línea] (Tesis de pregrado). Universidad de Granada, Madrid, España. 2009. pp. [Consulta: 2014-09-08]  
Disponible en: <http://hera.ugr.es/tesisugr/18196305.pdf>.
- [2] **AL-NASIR, J. & MADHANI.** *Jaime Alnasir's Homepage* [en línea], 2009. [Consulta: 18 de agosto del 2014].  
Disponible en: <http://www.al-nasir.com/www/Jamie/Articles/Pharmacy/G-Quadruplex.shtml>.
- [3] **BERMAN, H. M. y otros.** "The Protein Data Bank". *Nucleic Acids Research* [en línea], 2000, 28, pp. 235-242. [Consulta: 20 de septiembre 2014].  
Disponible en: <http://doi.org/10.1093/nar/28.1.235>
- [4] **BEVERIDGE, D. & RAVISHANKER, G.** "Molecular dynamics studies of DNA". *Current Opinion in Structural Biology* [en línea], 1994, 4(2), pp. 246-255. [Consulta: 15 de agosto 2014]. Disponible en: [http://doi.org/10.1016/S0959-440X\(94\)90316-6](http://doi.org/10.1016/S0959-440X(94)90316-6)
- [5] **BURGE, S. y otros.** "Survey and Summary Quadruplex DNA: sequence, topology and structure". *Nucleic Acids Research*, vol. 34, No. 19 (2006), pp. 5402-5415.
- [6] **CAI, W.** *NPT and Grand Canonical Ensembles* [en línea], 2011  
[Consulta: 14 de agosto 2014]. Disponible en:  
[http://micro.stanford.edu/~caiwei/me334/Chap9\\_NPT\\_Grand\\_Canonical\\_Ensemble\\_v04.pdf](http://micro.stanford.edu/~caiwei/me334/Chap9_NPT_Grand_Canonical_Ensemble_v04.pdf).
- [7] **CHEMIE.DE.** *Ácido desoxirribonucleico* [en línea], 1997.  
[Consulta: 13 de agosto 2014] Disponible en:  
[http://www.quimica.es/enciclopedia/%C3%81cido\\_desoxirribonucleico.html#\\_note-Greider/](http://www.quimica.es/enciclopedia/%C3%81cido_desoxirribonucleico.html#_note-Greider/).
- [8] **DAHM, R.** "Friedrich Miescher and the discovery of DNA". *Developmental Biology*, vol. 278, No. 2 (2005), pp. 274-288.

- [9] **FADRNA, E. y otros.** "Single Stranded Loops of Quadruplex DNA as Key Benchmark for Testing Nucleic Acids Force Field". *Journal of Chemical Theory and Computation*, vol. 5, No. 9 (2009), pp. 2514-2530.
- [10] **FADRNA, E. y otros.** "Molecular Dynamics Simulations of Guanine Quadruplex Loops: Advances and Force Field Limitations". *Biophysics* [en línea], 2004, 87(1), pp. 227-242. [Consulta: 15 febrero 2014]. Disponible en: <http://doi.org/10.1529/biophysj.103.034751>
- [11] **FOGOLARI, F. y otros.** "Molecular models for intrastrand DNA G-quadruplexes". *BMC Structural Biology*, vol. 9, No. 64 (2009).
- [12] **FORESTER, T. & SMITH, W.** "SHAKE, Rattle, and Roll: Efficient Constraint Algorithms for linked Rigid Bodies". *Journal of Computational Chemistry* [en línea], 1998, 19(1), pp. 102-111. [Consulta: 12 Febrero 2014].  
Disponible en: [http://doi.org/10.1002/\(SICI\)1096-987X\(19980115\)19:1<102::AID-JCC9>3.0.CO;2-T](http://doi.org/10.1002/(SICI)1096-987X(19980115)19:1<102::AID-JCC9>3.0.CO;2-T)
- [13] **FRIEDMAN, R. & HONIG, B.** "A Free Energy Analysis of Nucleic Acid Base Stacking in Aqueous Solution". *Biophysical Journal* [en línea], 1995, 69 (4), 1995, pp. 1528-1535. [Consulta: 18 Enero 2014].  
Disponible en: [http://doi.org/10.1016/S0006-3495\(95\)80023-8](http://doi.org/10.1016/S0006-3495(95)80023-8)
- [14] **GARGALLO, R.** Investigación [en línea], Universitat de Barcelona, 2008. [Consulta: 18 Enero 2014]. Disponible en: <http://www.ub.edu/gesq/raimon/investigacion.html>.
- [15] **HAIDER, S. & NEIDLE, S.** "Molecular Modeling and Simulation of G-Quadruplexes and quadruplex ligand complexes". *Methods in Molecular Biology*, vol. 608, 2010, pp. 17-37.
- [16] **HALL, H.** *The structures of DNA and RNA* [en línea], 2002. [Consulta: 18 julio 2014]  
Disponible en: [http://biology.kenyon.edu/courses/biol63/watson\\_06.pdf](http://biology.kenyon.edu/courses/biol63/watson_06.pdf).
- [17] **HOFFORT, A.** *A structural study of m-dna* (tesis de pregrado). University of Saskatchewan, Saskatoon, Saskatchewan. 2006

- [18] **HUMPHREY, W., DALKE, A. & SCHULTEN, K.** "VMD: Visual Molecular Dynamics". *Journal of Molecular Graphics* [en línea], 1996, 14(1), pp. 27-38. [Consulta: 20 septiembre 2013]. Disponible en: [http://doi.org/10.1016/0263-7855\(96\)00018-5](http://doi.org/10.1016/0263-7855(96)00018-5)
- [19] **HUPPERT, J. L.** "Four stranded DNA: cancer, gene regulation and drug development". *Philosophical Transactions. Series A, mathematical, physical and engineering sciences* [en línea], 2007, 365(1861), pp. 2969-2984. [Consulta: 20 septiembre 2013]. Disponible en: <http://doi.org/10.1098/rsta.2007.0011>
- [20] **JAQUET, R.** *Introduction to potential energy surfaces and graphical interpretation* [en línea], Siegen, 2002. [Consulta: 20 septiembre 2013]. Disponible en: <http://www.tc.chemie.uni-siegen.de/jaquet/Copy-of-introPEStexpowerlight.pdf>.
- [21] **JISSY, A. K., ASHIK, U. & DATTA, A.** "Nucleic Acid G-quartets: Insights into Diverse Patterns and Optical Properties". *The Journal of Physical Chemistry*, vol. 115, No. 25 (2011), pp. 12530-12546.
- [22] **LARRAONA, M. & AREVALO, A.** *Resonancia Magnética Nuclear* [en línea]. 2008. [Consulta: 24 septiembre 2013].  
Disponible en:  
<http://nuclear.fis.ucm.es/CDTEAM/articulos%20pdf/PDF%20PUBLICACIONES%20IMAGEN%20MEDICA/NMR-larraona-2008.pdf>.
- [23] **LAVERY, R.** "Modelling the DNA double helix: techniques and results". *Center de Physique des houches*, vol. 2, (1995), pp. 57-82.
- [24] **LE, H. y otros.** "An Investigation of G Quadruplex Structural Polymorphism in the human telomere using a combined approach of hydrodynamic bead modeling and molecular dynamics simulation". *The Journal of Physical Chemistry*, vol. 118, No 20 (2014), pp. 5390-5405.
- [25] **LIANG, L.** *Computational studies of DNA sequencing with graphene nanopores* (tesis doctoral), Royal Institute of Technology, Stockholm, Sweden, 2014.

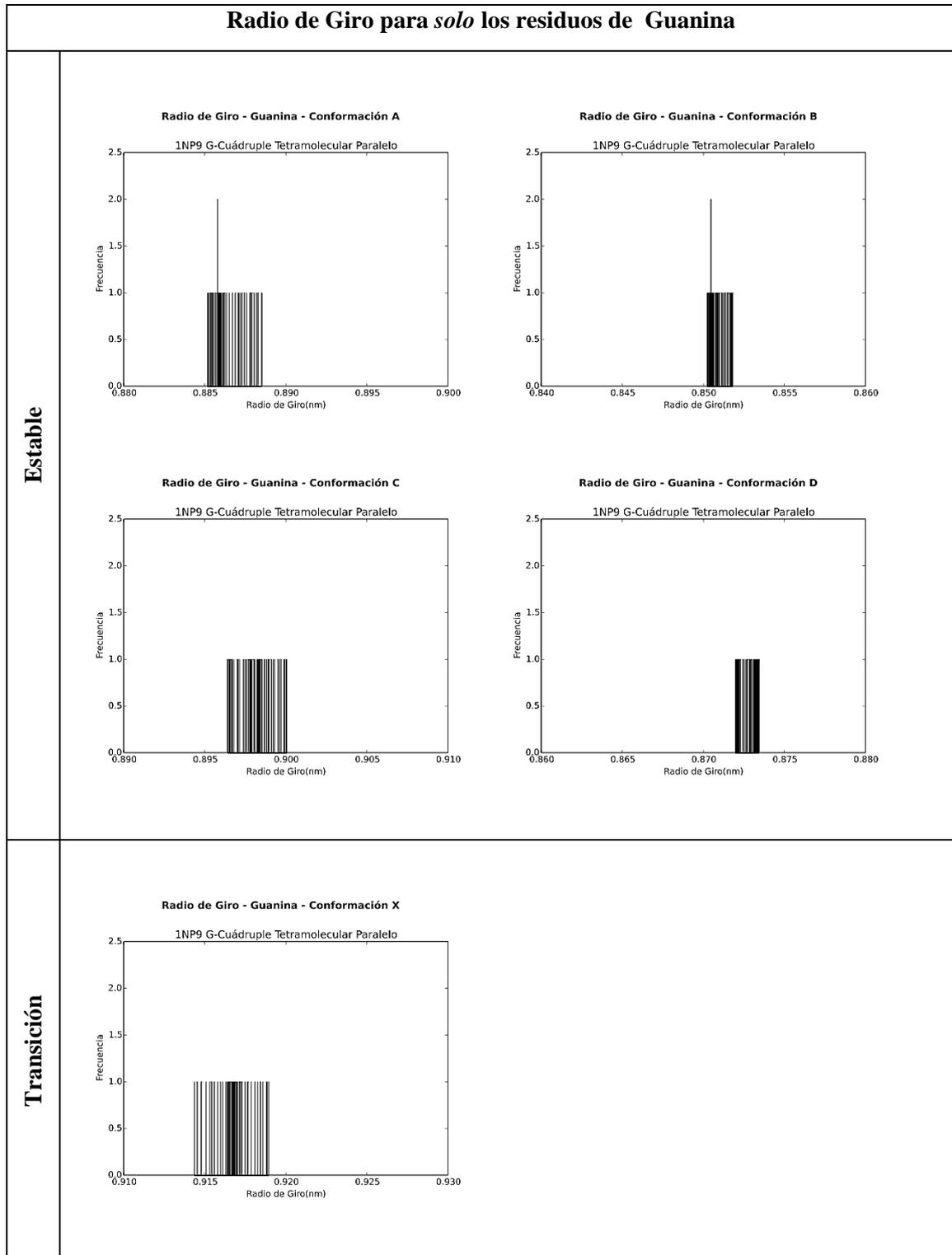
- [26] **MABRAHAM**. *Gromacs* [en línea], 2009. [Consulta: 25 de noviembre del 2015]  
Disponible en: <http://www.gromacs.org/>.
- [27] **MACKERELL, A., BANAVALI, N. & FOLOPPE, N.** "Development and current status of the CHARMM force field for nucleic acids". *Biopolymers*, vol. 56, No. 4 (2001), pp. 257-265.
- [28] **MACKERELL, A. & NILSSON, L.**, 2006. "Theoretical studies of nucleic acids and nucleic acid-protein complexes using CHARMM". *Computational Studies of RNA and DNA*, vol. 2, pp. 73-94.
- [29] **MAFFEO, C. y otros.** "Close encounters with DNA". *Journal of Physics: Condensed Matter: An Institute of Physics Journal*, vol. 26, No. 41 (2014), p. 36.
- [30] **MICHAEL, A., BRAHIM, H. & ANH, P.** "NMR spectroscopy of G-quadruplexes". *Elsevier*, vol. 57, No. 1 (2012), pp. 11-24.
- [31] **NEIDLE, S. & BALASUBRAMANIAN, S.** "Quadruplex Nucleic Acids". *Reims: The Royal Society of Chemistry*, 2007.
- [32] **PARKINSON, G. N., H, M. P. & NEIDLE, S.** "Crystal structure of parallel quadruplexes from human telomeric DNA". *Nature*, vol. 417, No. 6891 (2002), pp. 876-880.
- [33] **PHILLIPS, J. y otros.** "Scalable Molecular Dynamics with NAMD". *Journal of Computational Chemistry* [en línea], 2005, 26(16), pp. 1781-1802. [Consulta: 13 de agosto 2013]. Disponible en: [www.ncbi.nlm.nih.gov/pubmed/1622265](http://www.ncbi.nlm.nih.gov/pubmed/1622265)
- [34] **ROMERO, M. I.** *Efecto de la Composición de los gases de combustión Eb la captura de Co2 por Nanotubos De Carbono: Un estudio de simulación* (tesis de maestría). Escuela Técnica Superior de Ingenieros, 2001.
- [35] **Rueda, M.** *Estudio teórico sobre la influencia del solvente en la estructura y dinámica del ADN* (tesis maestría), Universidad de Barcelona, 2005
- [36] **SANCHO, F.** *Clasificación Supervisada y No Supervisada* [en líena]. 2014.  
[Consulta: 12 de noviembre 2013]. Disponible en: <http://www.cs.us.es/~fsancho/?e=77>.

- [37] **SPONER, J., CANG, X. & CHEATHAM, T.** "Molecular dynamics simulations of G-DNA and perspectives on the simulation of nucleic acid structures". *Methods*, vol.57, No.1 (2012), pp. 25-39.
- [38] **STEGLE, O. y otros.** "Predicting and understanding the stability of G-quadruplexes". *Bioinformatics* [en línea], 2009, 25(12,) pp. 374–382. [Consulta: 15 septiembre 2014]. Disponible en: <http://doi.org/10.1093/bioinformatics/btp210>
- [39] **THAO, P. y otros.** "Tetramolecular Quadruplex Stability and Assembly". *Springer*, vol. 128, 2012, pp. 2012-2334.
- [40] **TRISTÁN, M.** *Hélice cuádruple de ADN en células humanas* [en línea], 2013. [Consulta: 15 noviembre 2014]. Disponible en: <http://biotecmur.com/2013/02/27/dn-quadruplex/>.
- [41] **VALDÉS, A.** *Superficies de energía potencial y dinámica de agregados de Van der Waals triatómicos y tetraatómicos formados por gases nobles y moléculas dihalógenas* (tesis doctoral), Universidad Complutense Madrid, 2006.
- [42] **VALENCIA.** *Química Física III* [en línea], 2012. [Consulta: 21 noviembre 2014]. Disponible en: [http://www.uv.es/tunon/QFIII/tema\\_1.pdf](http://www.uv.es/tunon/QFIII/tema_1.pdf) 2013-10-20
- [43] **VANDIVORT, K. y otros.** 2011. [Consulta: 21 agosto 2014]. Disponible en: <https://web.stanford.edu/~jenhsin/HsinPubs/VAND2008/VAND2008.pdf>.
- [44] **VANOMMESLAEGHE, K. y otros.** "CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields". *Journal Computational Chemical*, vol.31, No. 4 (2010), pp. 671-690.
- [45] **VASUDEVAN, D., S, S. & VAIDYANATHAN, K.** *Texto de Bioquímica*. Sexta ed. Panamá: Jaypee Highlights Medical Publishers. 2011
- [46] **VESELY, F.** *Microcanonical ensemble* [en línea], Franz Vesely, 2005. [Consulta: 21 noviembre 2014]. Disponible en: [http://homepage.univie.ac.at/franz.vesely/sp\\_english/sp/node18.html](http://homepage.univie.ac.at/franz.vesely/sp_english/sp/node18.html).

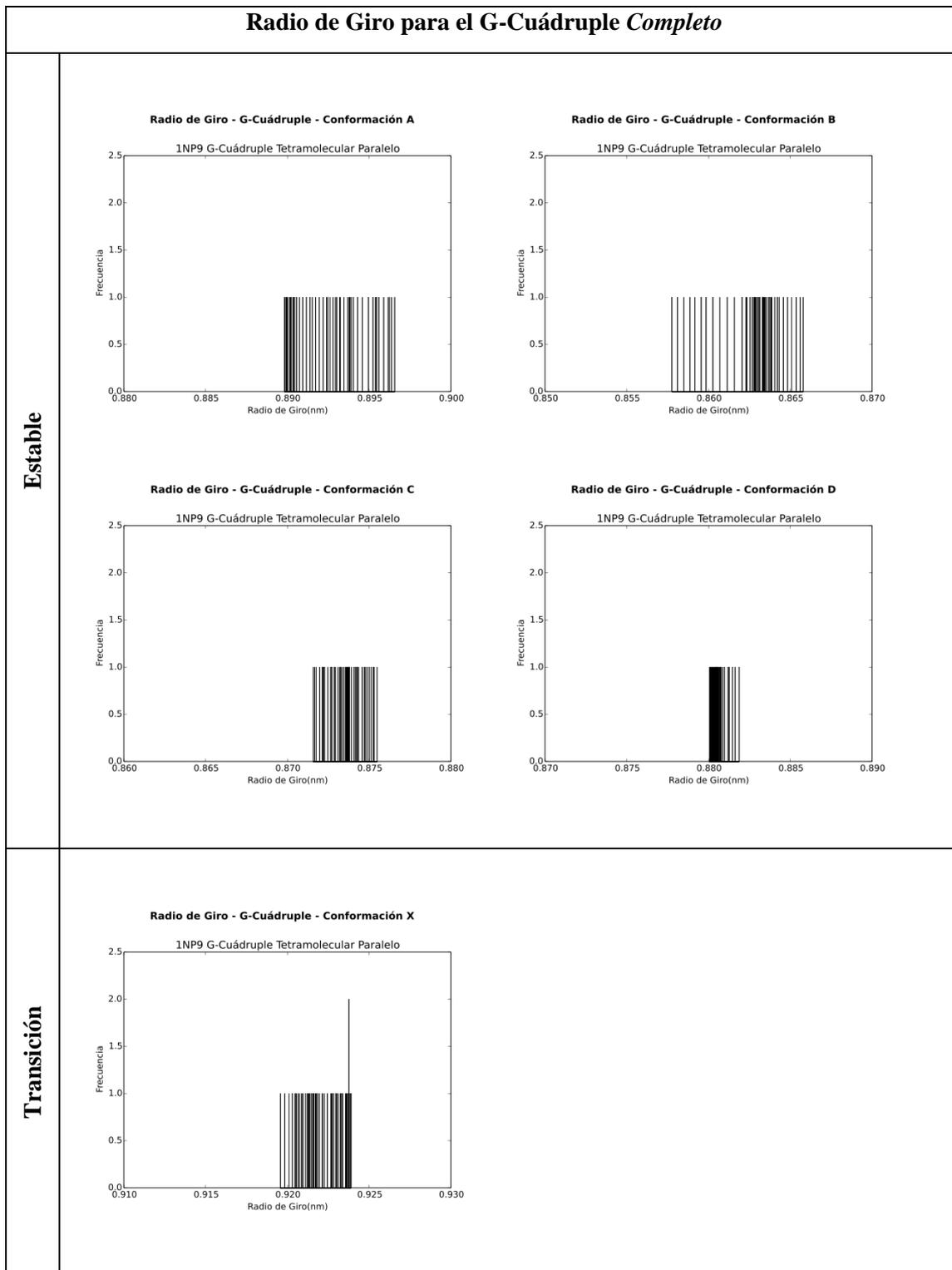
- [47] **VIRGILIO, A. y otros.** "The insertion of two 8-methyl-2'-deoxyguanosine residues in tetramolecular quadruplex structures: trying to orientate the strands". *Nucleic Acids Research*, vol. 40, No. 1, pp. 461-475.
- [48] **WILEY, J. & SONS.** "Physical Chemistry of Macromolecules". Segunda ed. Canada: A Wiley-Interscience Publication.1922.
- [49] **WOLFRAM, S.** *La historia de Mathematica* [en línea], WOLFRAM, 1987. [Consulta: 21 noviembre 2014].  
Disponible en: <http://www.wolfram.com/company/mathematica-history.es.html>.
- [50] **YANG, D. & OKAMOTO, K.** "Structural insights into G-quadruplexes: towards new anticancer drugs". *Future Medicinal Chemistry*, vol. 2, No. 4 (2010), pp. 619-646.
- [51] **ZHAO, G. y otros.** "Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics". *Nature*, vol. 497, No. 7451 (2013), pp. 643-646.
- [52] **ZHOU, J. y otros.** "Assembly of chemically modified G-rich sequences into tetramolecular DNA G-quadruplexes and higher order structures". *Methods*, vol. 67, No. 2 (2014), pp. 159-168.
- [53] **ZHU, X., LOPES, P. & MACKERELL, A.** "Recent Developments and Applications of the CHARMM force fields". *Wiley Interdisciplinary Reviews: Computational Molecular Science* [en línea], 2012, 2(1), pp. 167-185. [Consulta: 18 de septiembre 2014]. Disponible en: <http://doi.org/10.1002/wcms.74>

# ANEXOS

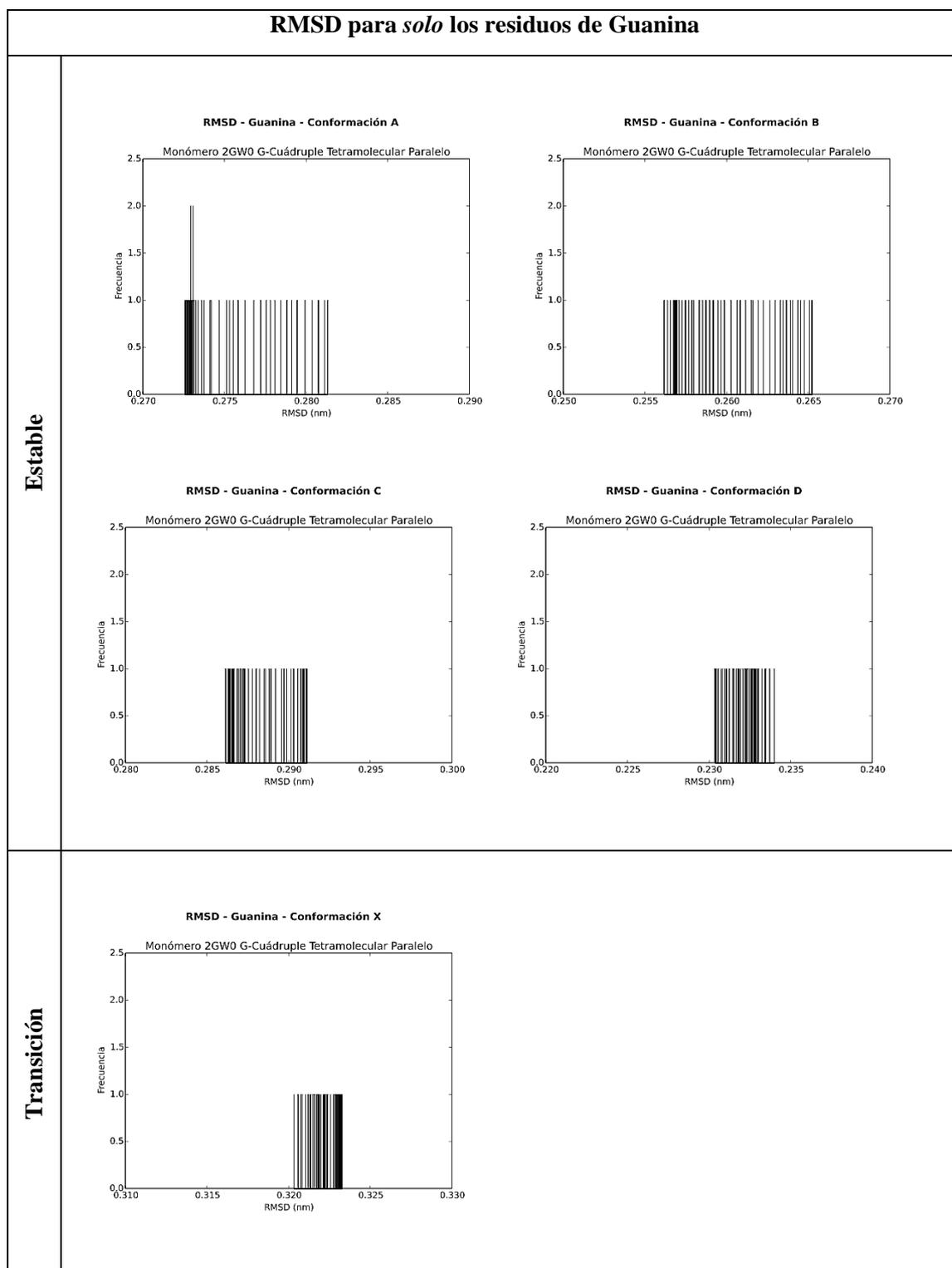
**Anexo A.** Histograma de Conformaciones seleccionadas por Radio de Giro para *solo* los residuos de Guaninas del 1NP9 G-Cuádruple Tetramolecular Paralelo



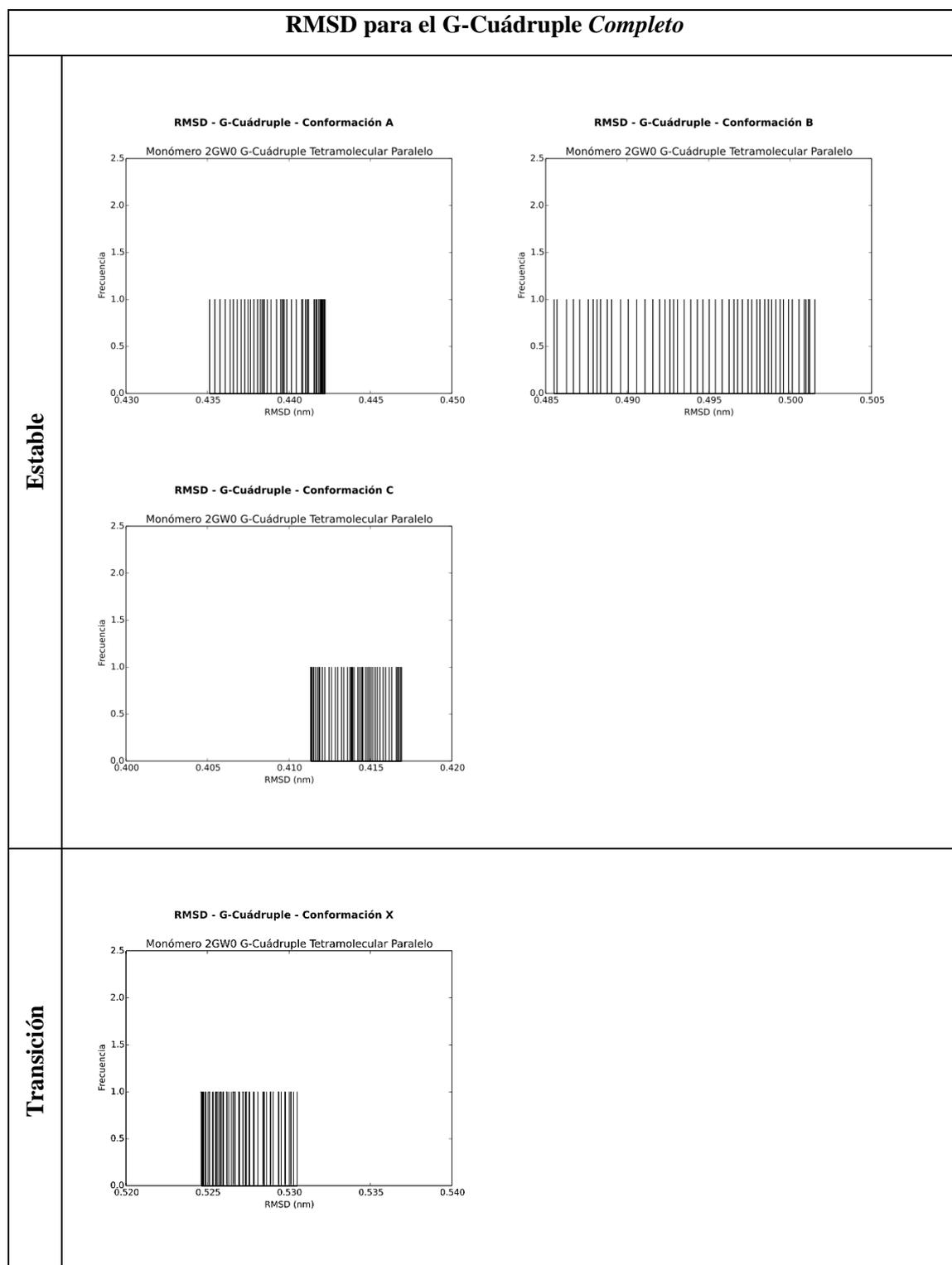
**Anexo B.** Histograma de Conformaciones seleccionadas por Radio de Giro para el G-Cuádruple *completo* 1NP9 G-Cuádruple Tetramolecular Paralelo



**Anexo C.** Histograma de Conformaciones seleccionadas por RMSD para *solo* los residuos de Guanina del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo



**Anexo D.** Histograma de Conformaciones seleccionadas por RMSD para el G-Cuádruple *completo* del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo



## Anexo E. Archivo PSF

### ARCHIVO PSF 1NP9

#### 13333 !NATOM

1 N1	1	THY C4' CN7	0.160000	12.0110
2 N1	1	THY H4' HN7	0.090000	1.0080
3 N1	1	THY O4' ON6	-0.500000	15.9994
4 N1	1	THY C1' CN7B	0.160000	12.0110
5 N1	1	THY H1' HN7	0.090000	1.0080
6 N1	1	THY C2' CN8	-0.180000	12.0110
7 N1	1	THY H2' HN8	0.090000	1.0080
8 N1	1	THY H2" HN8	0.090000	1.0080
9 N1	1	THY H5T HN5	0.430000	1.0080
10 N1	1	THY O5' ON5	-0.660000	15.9994
11 N1	1	THY C5' CN8B	0.050000	12.0110
12 N1	1	THY H5' HN8	0.090000	1.0080
13 N1	1	THY H5" HN8	0.090000	1.0080
14 N1	1	THY N1 NN2B	-0.340000	14.0070
15 N1	1	THY C6 CN3	0.170000	12.0110
16 N1	1	THY H6 HN3	0.170000	1.0080
17 N1	1	THY C2 CN1T	0.510000	12.0110
18 N1	1	THY O2 ON1	-0.410000	15.9994
19 N1	1	THY N3 NN2U	-0.460000	14.0070
20 N1	1	THY H3 HN2	0.360000	1.0080
21 N1	1	THY C4 CN1	0.500000	12.0110
22 N1	1	THY O4 ON1	-0.450000	15.9994
23 N1	1	THY C5 CN3T	-0.150000	12.0110
24 N1	1	THY C5M CN9	-0.110000	12.0110
25 N1	1	THY H51 HN9	0.070000	1.0080
26 N1	1	THY H52 HN9	0.070000	1.0080
27 N1	1	THY H53 HN9	0.070000	1.0080
28 N1	1	THY C3' CN7	0.010000	12.0110

### ARCHIV PSF 2GW0

#### 12060 NATOM

1 N1	1	THY C4' CN7	0.160000	12.0110
0				
2 N1	1	THY H4' HN7	0.090000	1.0080
3 N1	1	THY O4' ON6	-0.500000	15.9994
4 N1	1	THY C1' CN7B	0.160000	12.0110
5 N1	1	THY H1' HN7	0.090000	1.0080
6 N1	1	THY C2' CN8	-0.180000	12.0110
7 N1	1	THY H2' HN8	0.090000	1.0080
8 N1	1	THY H2" HN8	0.090000	1.0080
9 N1	1	THY H5T HN5	0.430000	1.0080
10 N1	1	THY O5' ON5	-0.660000	15.9994
11 N1	1	THY C5' CN8B	0.050000	12.0110
12 N1	1	THY H5' HN8	0.090000	1.0080
13 N1	1	THY H5" HN8	0.090000	1.0080
14 N1	1	THY N1 NN2B	-0.340000	14.0070
15 N1	1	THY C6 CN3	0.170000	12.0110
16 N1	1	THY H6 HN3	0.170000	1.0080
17 N1	1	THY C2 CN1T	0.510000	12.0110
18 N1	1	THY O2 ON1	-0.410000	15.9994
19 N1	1	THY N3 NN2U	-0.460000	14.0070
20 N1	1	THY H3 HN2	0.360000	1.0080
21 N1	1	THY C4 CN1	0.500000	12.0110
22 N1	1	THY O4 ON1	-0.450000	15.9994
23 N1	1	THY C5 CN3T	-0.150000	12.0110
24 N1	1	THY C5M CN9	-0.110000	12.0110
25 N1	1	THY H51 HN9	0.070000	1.0080
26 N1	1	THY H52 HN9	0.070000	1.0080
27 N1	1	THY H53 HN9	0.070000	1.0080
28 N1	1	THY C3' CN7	0.010000	12.0110

## Anexo F. Archivo NAMD

```

## JOB DESCRIPTION ##
#####
# Minimization 1NP9
#####
## ADJUSTABLE PARAMETERS ##
##
structure atomselect_1NP9.psf
coordinates atomselect_1NP9.pdb
firsttimestep 0
numsteps 3000
set temperature 298
set outputname 1NP9_mini
## SIMULATION PARAMETERS ##
##
# Input
paraTypeCharmm on
parameters par_all27_prot_lipid_na.inp
temperature $temperature
# Force-Field Parameters
exclude scaled1-4
1-4scaling 1.0
cutoff 12.0
switching on
switchdist 10.0
pairlistdist 14.0
#margin 5.0
# Integrator Parameters
timestep 2.0 ;# 2fs/step
rigidBonds all ;# needed for 2fs steps
nonbondedFreq 1
fullElectFrequency 2
stepspercycle 10
# Constant Temperature Control
langevin on ;# do langevin dynamics
langevinDamping 1 ;# damping coefficient (gamma)
of 1/ps
langevinTemp $temperature
langevinHydrogen off ;# don't couple langevin bath to
hydrogens
# Periodic Boundary Conditions
cellBasisVector1 50.89 0.0 0.0
cellBasisVector2 0.0 51.19 0.0
cellBasisVector3 0.0 0.0 56.35
cellOrigin 21.15 25.66 28.24
wrapAll on
# PME (for full-system periodic electrostatics)
PME yes
PMEGridSpacing 2.0
#manual grid definition
PMEGridSizeX 52
PMEGridSizeY 52
PMEGridSizeZ 58
# Constant Pressure Control (variable volume)
useGroupPressure yes ;# needed for rigidBonds
useFlexibleCell no
useConstantArea no
# Output
outputName $outputname
restartfreq 500 ;# 500steps =
every 1ps
dcdfreq 250
xstFreq 250
outputEnergies 100
outputPressure 100
#####
fixedAtoms on
fixedAtomsFile atomselect_1NP9_fix.pdb
fixedAtomsCol B
#####
## EXECUTION SCRIPT # Minimization
minimize 3000

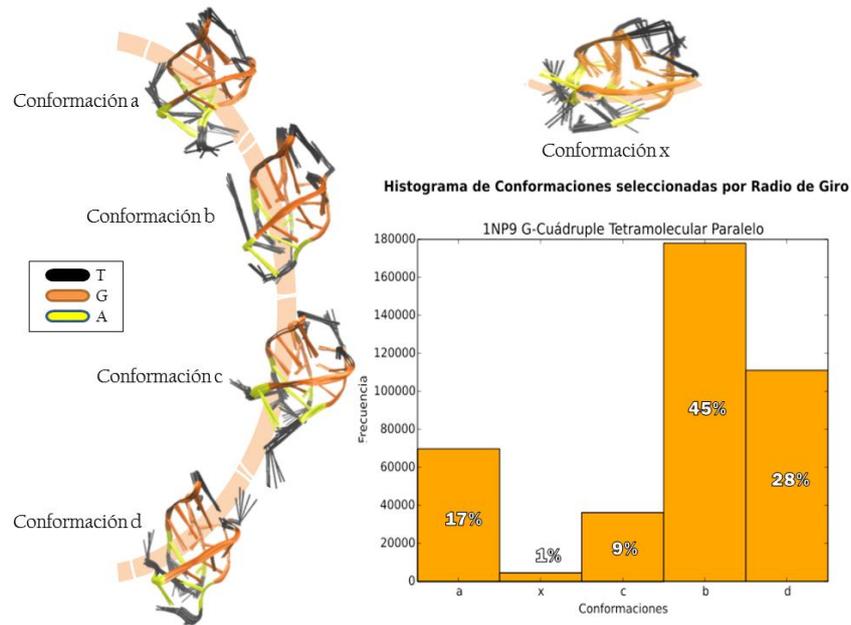
```

```

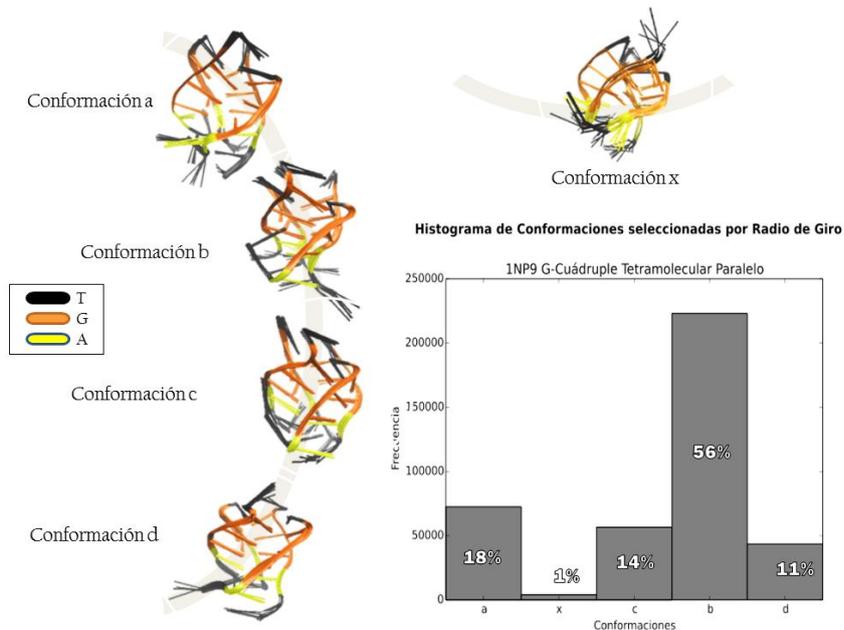
## JOB DESCRIPTION ##
#####
# Minimization Monómero 2GW0
#####
## ADJUSTABLE PARAMETERS ##
##
structure atomselect_2GW0.psf
coordinates atomselect_2GW0.pdb
firsttimestep 0
numsteps 3000
set temperature 298
set outputname 2GW0_mini
## SIMULATION PARAMETERS ##
##
# Input
paraTypeCharmm on
parameters par_all27_prot_lipid_na.inp
temperature $temperature
# Force-Field Parameters
exclude scaled1-4
1-4scaling 1.0
cutoff 12.0
switching on
switchdist 10.0
pairlistdist 14.0
#margin 5.0
# Integrator Parameters
timestep 2.0 ;# 2fs/step
rigidBonds all ;# needed for 2fs steps
nonbondedFreq 1
fullElectFrequency 2
stepspercycle 10
# Constant Temperature Control
langevin on ;# do langevin dynamics
langevinDamping 1 ;# damping coefficient (gamma) of 1/ps
langevinTemp $temperature
langevinHydrogen off ;# don't couple langevin bath to
hydrogens
# Periodic Boundary Conditions
cellBasisVector1 54.3 0.0 0.0
cellBasisVector2 0.0 54.29 0.0
cellBasisVector3 0.0 0.0 45.44
cellOrigin 25.30 3.71 -2.50
wrapAll on
# PME (for full-system periodic electrostatics)
PME yes
PMEGridSpacing 2.0
#manual grid definition
PMEGridSizeX 55
PMEGridSizeY 55
PMEGridSizeZ 46
# Constant Pressure Control (variable volume)
useGroupPressure yes ;# needed for rigidBonds
useFlexibleCell no
useConstantArea no
# Output
outputName $outputname
restartfreq 500 ;# 500steps = every 1ps
dcdfreq 250
xstFreq 250
outputEnergies 100
outputPressure 100
#####
fixedAtoms on
fixedAtomsFile atomselect_1NP9_fix.pdb
fixedAtomsCol B
#####
## EXECUTION SCRIPT # Minimization
minimize 3000

```

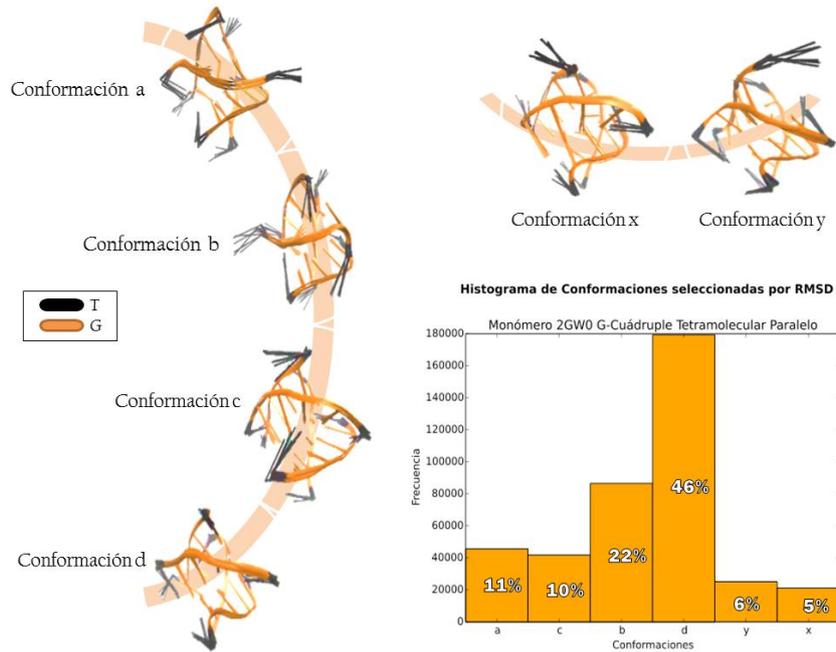
**Anexo G.** Histograma de Clasificación de Conformaciones seleccionadas por Radio de Giro para *solo* los residuos de Guaninas del 1NP9 G-Cuádruple Tetramolecular Paralelo



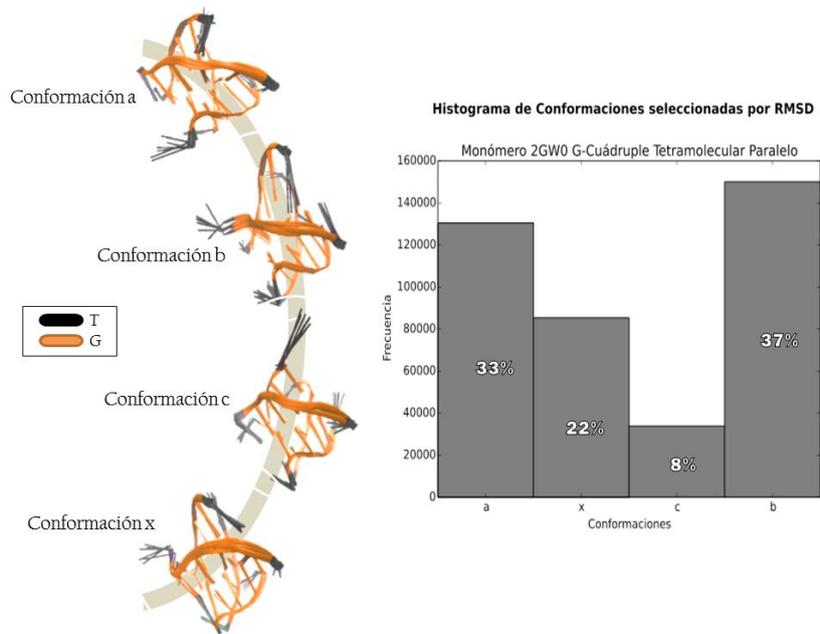
**Anexo H.** Histograma de Clasificación de Conformaciones seleccionadas por Radio de Giro para el G-Cuádruple *completo* 1NP9 G-Cuádruple Tetramolecular Paralelo



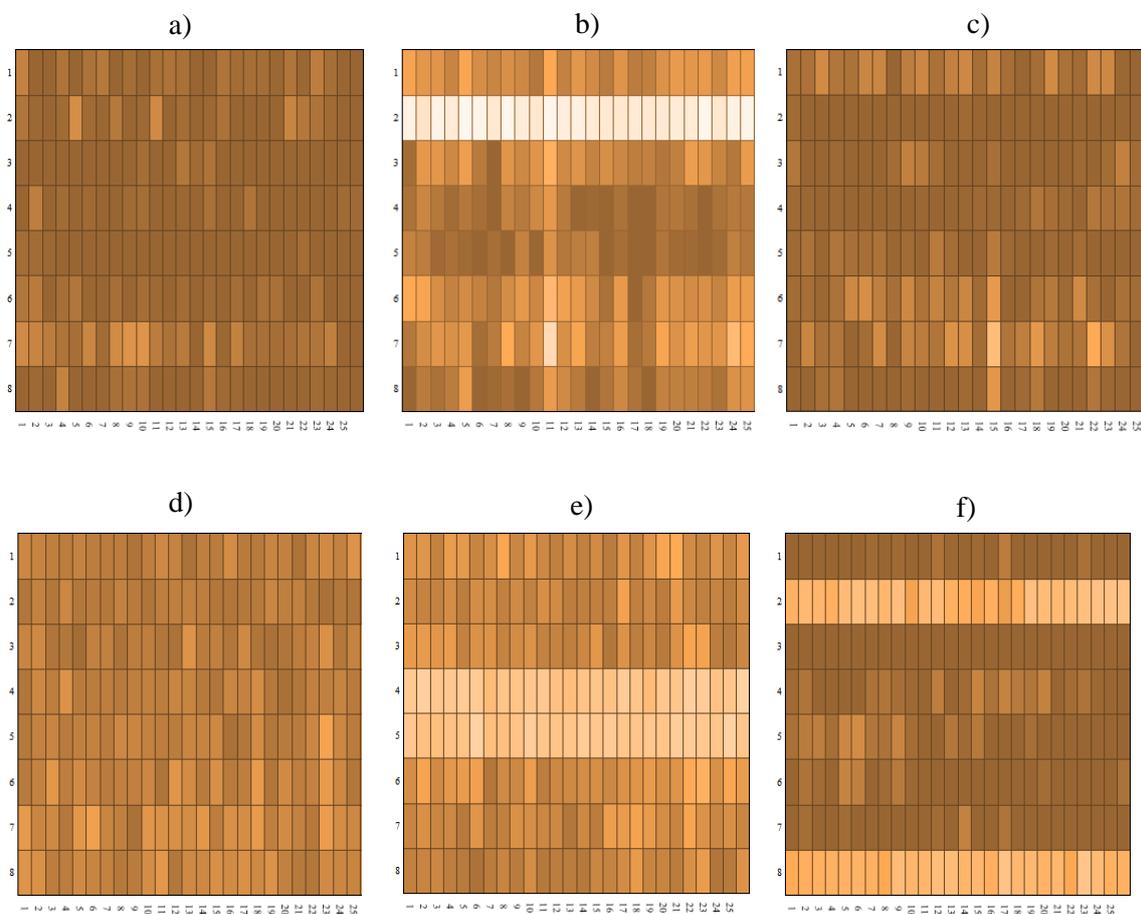
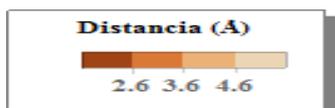
**Anexo I.** Histograma de Clasificación seleccionadas por RMSD para los residuos de Guanina del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo



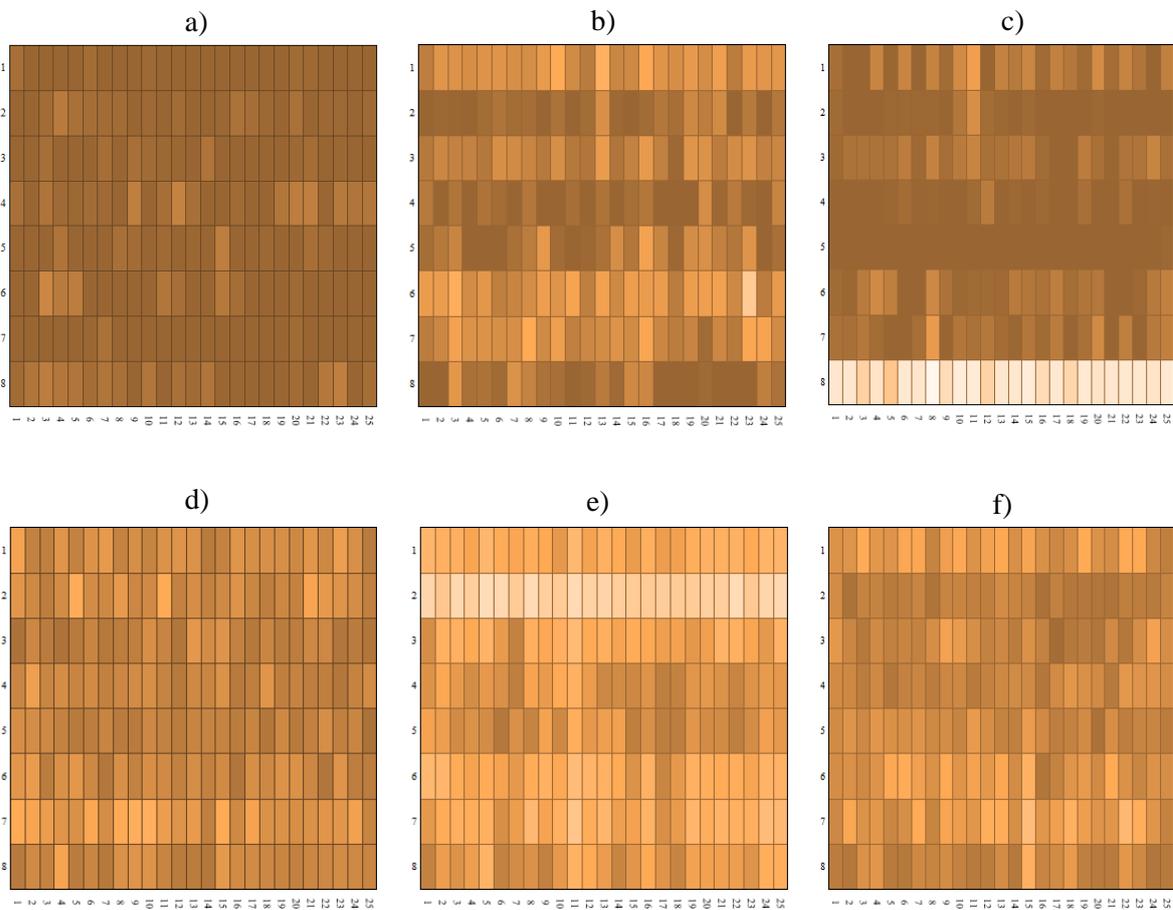
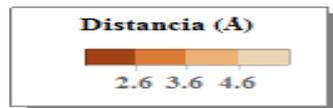
**Anexo J.** Histograma de Clasificación seleccionadas por RMSD para el G-cuádruple *completo* del Monómero 2GW0 G-Cuádruple Tetramolecular Paralelo



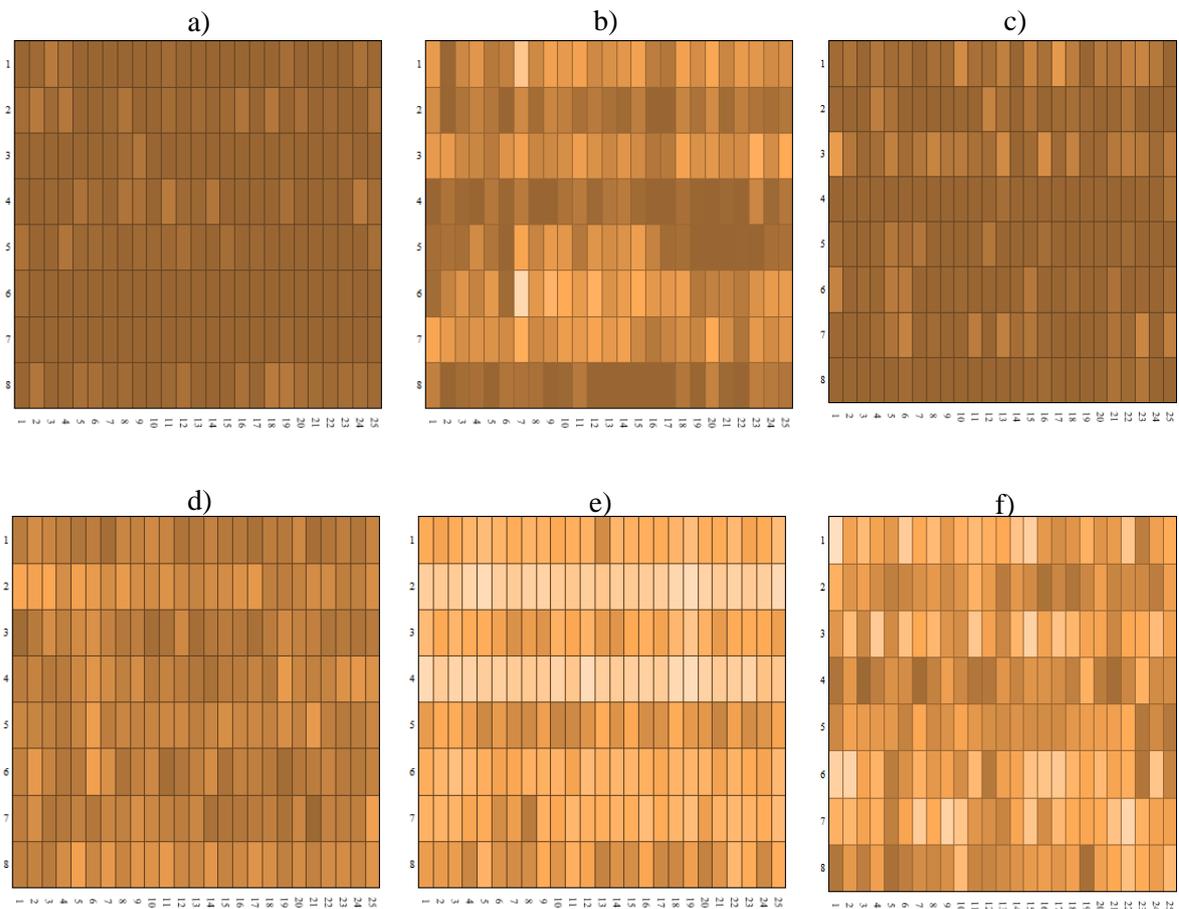
En el **Anexo K**. Mapas de Calor de la estructura 1NP9 G-Cuádruple a) Cuarteto 1 de la Conformación estable “a” b) Cuarteto 2 de la Conformación estable “a” c) Cuarteto 3 de la Conformación estable “a” d) Cuarteto 1 de la Conformación transición “x” e) Cuarteto 2 de la Conformación transición “x” f) Cuarteto 3 de la Conformación transición “x”. Conformaciones pertenecen al parámetro RMSD del G-Cuádruple *completo*



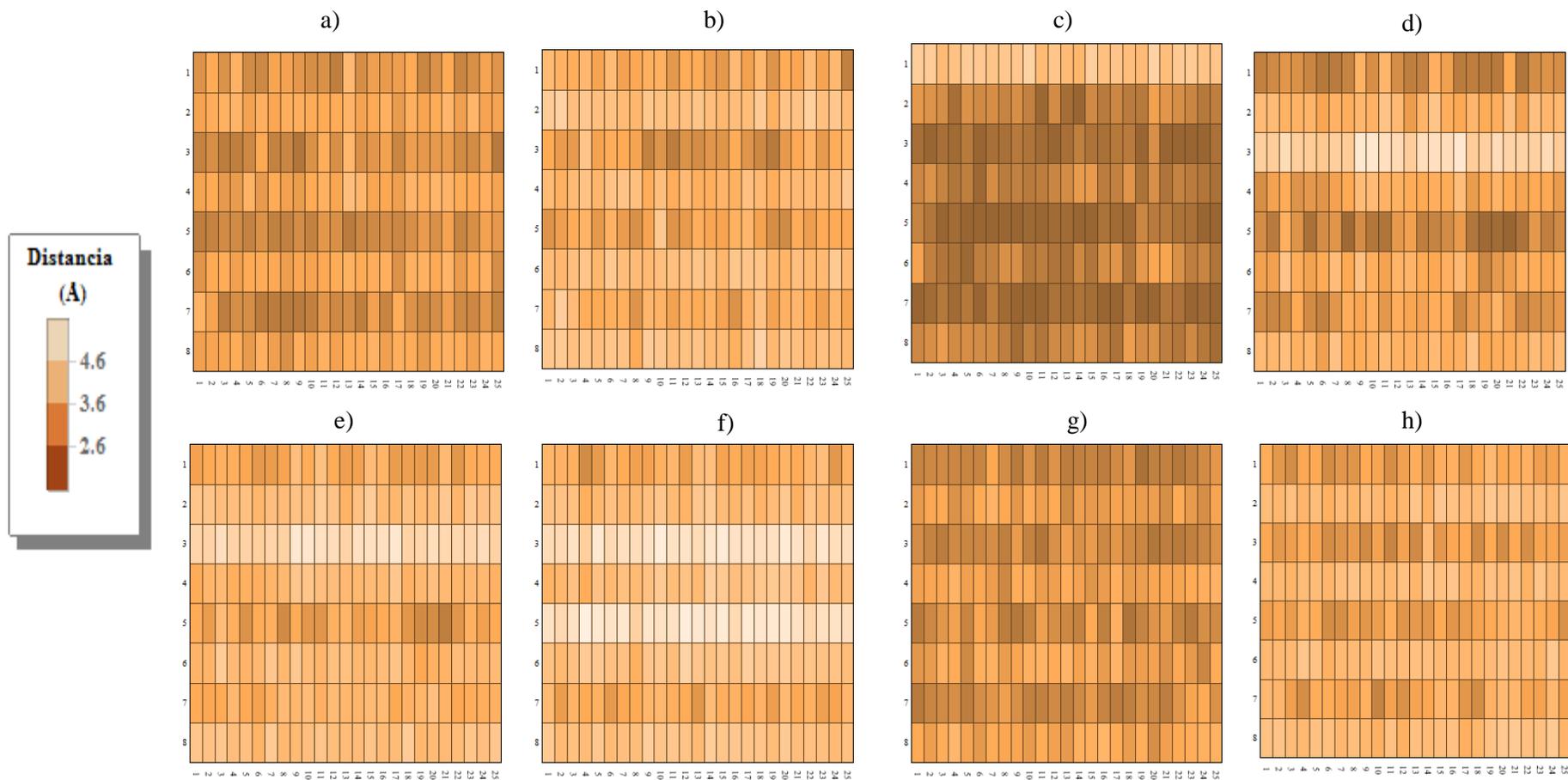
En el **Anexo L**. Mapas de Calor de la estructura 1NP9 G-Cuádruple a) Cuarteto 1 de la Conformación estable “b” b) Cuarteto 2 de la Conformación estable “b” c) Cuarteto 3 de la Conformación estable “b” d) Cuarteto 1 de la Conformación transición “x” e) Cuarteto 2 de la Conformación transición “x” f) Cuarteto 3 de la Conformación transición “x”. Conformaciones pertenecen al parámetro Radio de Giro de *solo* los Residuos de Guanina



En el **Anexo M.** Mapas de Calor de la estructura 1NP9 G-Cuádruple a) Cuarteto 1 de la Conformación estable “b” b) Cuarteto 2 de la Conformación estable “b” c) Cuarteto 3 de la Conformación estable “b” d) Cuarteto 1 de la Conformación transición “x” e) Cuarteto 2 de la Conformación transición “x” f) Cuarteto 3 de la Conformación transición “x”. Conformaciones pertenecen al parámetro Radio de Giro del G-Cuádruple *completo*



En el **Anexo N.** Mapas de Calor del Monómero 2GW0 G-Cuádruple a) Cuarteto 1 de la Conformación estable “b” b) Cuarteto 2 de la Conformación estable “b” c) Cuarteto 3 de la Conformación estable “b” d) Cuarteto 4 de la Conformación estable “b” e) Cuarteto 1 de la Conformación transición “x” f) Cuarteto 2 de la Conformación transición “x” g) Cuarteto 3 de la Conformación transición “x” h) Cuarteto 4 de la Conformación transición “x” Conformaciones pertenecen al parámetro RMSD del G-Cuádruple *completo*



En el **Anexo O**. Mapas de Calor del Monómero 2GW0 G-Cuádruple a) Cuarteto 1 de la Conformación estable “d” b) Cuarteto 2 de la Conformación estable “d” c) Cuarteto 3 de la Conformación estable “d” d) Cuarteto 4 de la Conformación estable “d” e) Cuarteto 1 de la Conformación transición “x” f) Cuarteto 2 de la Conformación transición “x” g) Cuarteto 3 de la Conformación transición “x” h) Cuarteto 4 de la Conformación transición “x” Conformaciones pertenecen al parámetro Radio de Giro de *solo* los residuos de Guanina.

