



BUAP

BENEMÉRITA UNIVERSIDAD
AUTÓNOMA DE PUEBLA

Facultad de Ciencias Fisico Matemáticas

**Estudio Mecánico Cuántico de regularidades de
formación de estructuras de hélices dobles
diferentes del modelo de Watson-Crick**

Tesis presentada al

Posgrado en Ciencias Física Aplicada

como requisito parcial para la obtención del grado de

**MAESTRA EN CIENCIAS
(Física Aplicada)**

por

Lic. Verónica Domínguez Benítez

asesorada por

**Dr. Valeri Poltev (FCFM-BUAP)
Dra. Alexandra Deriabina (FCFM-BUAP)**

Julio 2017

Este documento está preparado para ser impreso a doble cara.

BENEMÉRITA UNIVERSIDAD
AUTÓNOMA DE PUEBLA

Facultad de Ciencias Fisico Matemáticas

**Estudio Mecánico Cuántico de regularidades de
formación de estructuras de hélices dobles
diferentes del modelo de Watson-Crick**

Tesis presentada al

Posgrado en Ciencias Física Aplicada

como requisito parcial para la obtención del grado de

**MAESTRA EN CIENCIAS
(Física Aplicada)**

por

Lic. Verónica Domínguez Benítez

asesorada por

Dr. Valeri Poltev (FCFM-BUAP)
Dra. Alexandra Deriabina (FCFM-BUAP)

Julio 2017

Copyright © Verónica Domínguez Benítez

Título: Estudio Mecánico Cuántico de regularidades de formación de estructuras de hélices dobles diferentes del modelo de Watson-Crick

Estudiante: Lic. Verónica Domínguez Benítez

COMITÉ

Dr. Eduardo González Jiménez
PRESIDENTE

Dra. Dolores García Toral
SECRETARIO

Dra. Olga Leticia Fuchs Gómez
VOCAL

Dr. Roberto Ramírez Sánchez
SUPLENTE

Dr. Valeri Poltev
ASESOR

Dra. Alexandra Deriabina
ASESORA

*A ese bello y cálido ángel
que ahora me cuida “desde la Luna”;
quien me dió el más bello amor
llenando mi vida de cariño, risas,
juegos, locuras y travesuras,
recuerdos que atesoro y
están latentes en mí.
A mi abuelito Chago.*

*Con todo mi amor a mis papás,
Nelly y Santiago,
quienes me han apoyado total e
incondicionalmente en todos y cada uno
de los aspectos de mi vida,
siendo ambos mi ejemplo a seguir,
su fortaleza es digna de admirar.
Ambos le han dado alas a mis sueños
y eso es algo que atesoraré por siempre en mí.
Y a mis compañeras de chocaventuras,
risas, frustraciones y lágrimas;
con quienes sé que siempre podré contar,
Nelly y Karen,
aunque podemos llegar a discutir,
mi vida no sería buena si
no estuviesen aquí.
A todos ustedes porque han estado siempre,
porque cuando mi mundo parecía derrumbarse
no me dejaron hundir y me ayudaron a salir
motivándome a luchar y no desistir.*

*A esa pequeña personita que
llegó a mi vida para iluminarla
con el más puro amor,
mi luz en la oscuridad.
Ese pequeño angelito que es lo más
valioso y bello que es parte de mí.
Por quien quiero dar lo mejor
y algún día ser un ejemplo a seguir.
Quien me regala los más dulces
besos y abrazos que llenan de
alegría mi ser.
A tí mi hermoso lucero,
mi pequeño y hermoso bebé,
Mateo.
Tuyo es por siempre mi corazón.*

Agradecimientos

*A veces en la vida
encuentras alguien que
cambia tu vida
simplemente siendo parte de tí.*

El camino recorrido para llegar a esta altura, la finalización de mi tesis de maestría, está plagado de momentos y lugares rodeada de mi familia, amigos, compañeros, y gente muy querida a la que me gustaría agradecer.

En primer lugar, a mis padres, Nelly y Santiago, quienes me han apoyado incondicionalmente durante todo este tiempo, han sido mi apoyo en todos los sentidos, me han escuchado y aconsejado en mis momentos de angustia recordándome que ya he salido adelante en muchos aspectos y puedo volver a hacerlo, sólo no tengo que darme por vencida. A mis hermanas, Nelly y Karen, porque en esta etapa me han apoyado más que en ninguna otra, han estado ahí para mi goldito y le han brindado amor y, tiempo que a veces yo no he tenido. A ustedes porque el comienzo de esta etapa fue algo dura para mí, me vieron caer y me ayudaron a levantarme, sentía que me hundía pero me mantuvieron a flote hasta que me dí cuenta de que no era el final y había aún muchas cosas hermosas por vivir.

A mi pequeño bebé Mateo, por llenar mi vida de luz, por darme todo su amor de la manera más linda y pura que he podido conocer, en esos pequeños abrazos he encontrado muchas veces la paz que mi ser ha necesitado, por sus ocurrencias, juegos y risas que alegran mis días. A tí mi angelito porque, a pesar de tu muy corta edad, has entendido que mamá tenía trabajos que terminar, porque hemos sacrificado momentos de juego y aún así me ves con esa carita llena de amor que me motiva a seguir adelante, a luchar por los dos.

Al Dr. Valeri Poltev, por todo el conocimiento que ha compartido conmigo, siento que nunca dejaré de aprender de él, por esa pasión que tiene en lo que hace mostrándome que se puede trabajar en cosas que realmente te hagan feliz, por su calidez, recordándome a alguien muy querido; y a su esposa la señora Nina por su

cariño y risas compartidas en el cubo. A la Dra. Alexandra Deriabina por todo su apoyo brindado y esa empatía que siempre muestra. A nuestro pequeño grupo de trabajo porque he encontrado en ustedes una pequeña familia y, un lugar ameno de trabajo.

A mis amigos, con aquéllos que comencé y con los que terminé este camino, porque hicieron que esta etapa se llenara de risas, chocoaventuras, nuevas experiencias y gratos momentos para recordar, porque sin saberlo también han sido un gran apoyo para mí y han hecho que en mis días tristes pueda sonreír.

Finalmente, al Laboratorio Nacional de Supercómputo del Sureste de México perteneciente a la red de laboratorios nacionales CONACYT, por los recursos computacionales, el apoyo y la asistencia técnica, de los cuales este proyecto de investigación se benefició ampliamente.

Índice

Agradecimientos	XIII
Resumen	XXV
Objetivos	XXVII
1. Introducción	1
1.1. Breve historia sobre el descubrimiento de la estructura del ADN [4]	3
1.2. Importancia biológica de las conformaciones de ADN diferentes del modelo de Watson-Crick	6
1.2.1. Estructuras de la Familia Conformacional Z	8
1.2.2. Dúplex con el par Ade:Thy de Hoogsteen	8
1.2.3. Conformaciones de cadenas paralelas	9
1.2.4. Dúplex de Watson-Crick con pares no correctos	10
2. Estructura del ADN con conformación diferente del modelo de Watson-Crick	13
2.1. Fragmentos de dobles hélices de ADN	15
2.1.1. Fragmento de ADN de la Familia Conformacional Z	16
2.1.2. Dúplex con pares de bases Ade:Thy de Hoogsteen	17
2.1.3. Modelo de conformaciones de ADN de cadenas paralelas con el par reverso de WC Ade:Thy	19
2.1.4. Dúplex de ADN del modelo de WC con pares no correctos	20
2.2. Parámetros geométricos para el estudio del ADN	22
3. Metodología para el estudio Teórico-Computacional del ADN	27
3.1. Mecánica Cuántica	28
3.1.1. Teoría del Funcional de la Densidad (DFT)	34
3.2. Mecánica Molecular	40
3.2.1. Contribuciones Intramoleculares a la Energía del Sistema Molecular [33]	41

3.2.2.	Interacciones intermoleculares e intramoleculares no enlazantes [33]	43
3.2.3.	Campo de Fuerza AMBER (Assisted Model Building with Energy Refinement)	45
3.3.	Metodología empleada para estudiar los dúplex de ADN diferentes del modelo de Watson-Crick	45
3.3.1.	Especificación de las estructuras empleadas para los cálculos .	49
3.3.2.	Medición de parámetros	49
4.	Resultados	51
4.1.	cdDMPs de la Familia Estructural Z	52
4.2.	Características conformacionales de dúplex con pares de bases Ade(<i>syn</i>):Thy de Hoogsteen	56
4.3.	Conformaciones de cadenas paralelas con el par reverso de Watson-Crick Ade:Thy	64
4.4.	Dúplex de Watson-Crick con pares no correctos	71
4.5.	Estructuras de cadena única	82
5.	Conclusiones	87
A.	Distribución de ángulos de torsión y de pseudorotación para las familias estructurales AI, AII, BI y BII	89
	Bibliografía	91

Índice de figuras

1.1. <i>ADN sintetizado, cristalizado y microfotografiado por Linden Gledhill.</i>	2
1.2. <i>“Foto 51”, tomada por Franklin, que muestra un patrón de difracción de rayos X de ADN de la forma B</i>	4
1.3. <i>Modelo de Watson y Crick de la molécula de ADN.</i>	5
1.4. <i>Estructura de la hélice doble del ADN. Fragmento de ADN de la familia B1, vista aérea (imagen a la izq.) y frontalmente con diferentes efectos.</i>	6
1.5. <i>Clasificación de las cuatro bases nitrogenadas en Pirimidinas y Purinas.</i>	7
1.6. <i>Visitante de un museo en Nueva York observando una representación digital del genoma humano en el 2001. (Cada color en la secuenciación del genoma humano representa una de las 4 bases nitrogenadas del ADN, identificando así los genes en cada cromosoma.)</i>	8
1.7. <i>Meiosis. (Profase I: a, b, c, d y e. Metafase I: f. Anafase I: g. Telofase I: h. Profase II: i. Metafase II: j. Anafase II: k. Telofase II: l.)</i>	9
1.8. <i>Cromosomas homólogos y Cromátidas hermanas. Sinapsis y Crossing-over.</i>	9
1.9. <i>Mutación Puntual.</i>	10
1.10. <i>Fragmento mínimo de ADN, ejemplo de dDMPs complementarios neutralizados con Na⁺, con los pares de bases Gua:Cyt y Ade:Thy.</i>	11
2.1. <i>Estructura de un nucleótido</i>	14
2.2. <i>(IZQ.) Modelo de Watson-Crick con cadenas antiparalelas. (DER.) Pares de bases de WC. a) Par Thy:Ade. b) Par Cyt:Gua.</i>	14
2.3. <i>Vistas laterales y superiores que ilustran los rasgos característicos de las estructuras helicoidales regulares A, B y ADN-Z deducidas a partir de los modelos de difracción de rayos X de fibras representativas. Las cintas rastrean la progresión de la columna vertebral definida por los átomos de fósforo y las líneas negras gruesas (cajas) representan los ejes helicoidales [18].</i>	15

2.4.	(IZQ.) <i>Fragmento de ADN de la Familia Conformacional Z. (DER.) Conformación de nucleósidos en ADN-Z.</i>	17
2.5.	(IZQ.) <i>Fragmento de ADN con pares de Hoogsteen. (DER.) Par Thy:Ade de Hoogsteen.</i>	18
2.6.	(IZQ.) <i>Modelo de dúplex de ADN de cadenas paralelas. (DER.) Par reverso de WC Thy:Ade, presente en las estructuras de cadenas paralelas.</i>	19
2.7.	(IZQ.) <i>Fragmento de ADN con dos pares de bases no correctos, Gua:Ade. (DER.) Pares de bases no correctos. a) Par Gua:Ade. b) Par Ino:Ade. c) Par Thy:Gua.</i>	21
2.8.	<i>Designación de los ángulos de torsión de un dDMP.</i>	22
2.9.	<i>Empaquetamiento del azúcar.</i>	23
2.10.	<i>Descripciones gráficas de los parámetros Buckle, Propeller, Tilt, Roll, h-Rise y h-Twist, empleados para describir la geometría de pares de bases complementarios y secuencias de pares de bases.</i>	24
2.11.	<i>Arreglo de bases G-A apiladas. La base con color azul en sus anillos se encuentra detrás de la base con color amarillo en sus anillos, con color verde se muestra el área de superposición de los anillos.</i>	24
3.1.	<i>Unidad mínima de la hélice doble del ADN. Desoxidinucleósidos monofosfatos complementarios.</i>	46
4.1.	<i>Dúplex de (dCpdG)₂ de la Familia Conformacional Z, optimizados con PBE y posteriormente con AMBER. Vista superior y lateral, mostrando superposición de las bases apiladas y paralelismo entre los pares de bases respectivamente.</i>	55
4.2.	<i>Dúplex de (dGpdC)₂ de la Familia Conformacional Z, optimizados con PBE y posteriormente con AMBER. Vista superior y lateral, mostrando superposición de las bases apiladas y posición mutua entre los pares de bases respectivamente.</i>	55
4.3.	<i>Conformaciones de dúplex con pares de Hoogsteen Ade(syn):Thy mostrando la superposición de sus bases apiladas. (ARRIBA) dTpdT:dApdA. (CENTRO) (dTpdA)₂. (ABAJO) (dApdT)₂.</i>	63
4.4.	<i>Conformaciones de dúplex de cadenas paralelas con el par reverso de Watson-Crick Ade:Thy mostrando la superposición de sus bases apiladas. (ARRIBA) dTpdA:dApdT. (ABAJO) dTpdT:dApdA.</i>	70
4.5.	<i>Dúplex optimizado con DFT del dTpdT:dGpdA con el par erróneo Gua:Thy, mostrando la superposición de sus bases apiladas y la posición mutua de sus pares de bases.</i>	80

4.6.	<i>Dúplex optimizado con DFT del dGpdA:dTpdA con el par erróneo Gua:Ade(syn), mostrando la superposición de sus bases apiladas y la posición mutual de sus pares de bases.</i>	80
4.7.	<i>Conformaciones de dúplex con pares erróneos mostrando la superposición de sus bases apiladas. (ARRIBA) dCpdG:dTpdG (Gua:Thy). (CENTRO) dApdG:dCpdG (Gua:Ade(syn)). (ABAJO) dApdG:dCpdI (Ino:Ade(syn)).</i>	81
4.8.	<i>Conformaciones de cdDMPs optimizados con DFT que presentaron sustancial traslape entre Pyr-Pyr, mostrando la posición mutual de sus bases. (ARRIBA-Izq.) Dúplex de cadenas paralelas, dTpdT:dApdA, con superposición entre las bases Thy. (ARRIBA-Der.) Estructura de cadena única optimizada con DFT presentando bases no paralelas, dTpdT, tomada del dúplex con cadenas paralelas mencionado. (ABAJO-Izq.) Dúplex con par no correcto Gua:Thy, dTpdT:dGpdA, con superposición entre las bases Thy. (ABAJO-Der.) Estructura de cadena única optimizada con DFT presentando bases no paralelas, dTpdT, tomada del dúplex con par erróneo Gua:Thy mencionado.</i>	84
4.9.	<i>Conformación de cdDMPs optimizados con DFT que presentaron sustancial traslape entre Pyr-Pur(syn), mostrando la posición mutual de sus bases. (IZQ.) Dúplex con el par no correcto Gua:Ade(syn), dTpdA:dGpdA, con superposición entre las bases Thy-Ade. (DER.) Estructura de cadena única optimizada con DFT presentando gran ángulo entre sus bases, dTpdA, tomada del dúplex con par erróneo Gua:Ade(syn) mencionado.</i>	84
4.10.	<i>Estructuras de cadena única, dDMPs, tomadas de dúplex que presentaron una gran superposición entre Pyr-Pyr. (IZQUIERDA) dTpdT con gran traslape entre sus bases apiladas, optimizado con DFT. (DERECHA) dTpdT con gran traslape entre sus bases apiladas, optimizado con AMBER.</i>	85
A.1.	<i>Distribución del ángulo de pseudorotación, P, del anillo de azúcar en las estructuras optimizadas de dDMPs, correspondientes a dúplex diestros. Cada secuencia de bases se localiza en una circunferencia punteada. Cuatro grupos de secuencias están separados por círculos sólidos. Estos son Pur-Pur: 1-dApdA, 2-dGpdG, 3-dAdpG, 4-dGdpA; Pyr-Pyr: 1-dCpdC, 2-dTpdC, 3-dCpdT, 4-dTpdT; Pur-Pyr: 1-dApdC, 2-dApdT, 3-dGpdT, 4-dGpdC; Pyr-Pur: 1-dCpdA, 2-dTpdG, 3-dCpdG, 4-dTpdA.</i>	89
A.2.	<i>Distribución de los ángulos de torsión del esqueleto azúcar-fosfato para estructuras optimizadas de dDMPs. Los valores promedio de los ángulos de torsión para cada una de cuatro familias de dúplex en cristales se indican mediante líneas rectas. Otras designaciones son explicadas en la leyenda de la Fig. A.1.</i>	90

Índice de Tablas

4.1. <i>Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ, (véase designación en la fig. 2.8), para los mínimos de dDMP-Na+ complementarios correspondientes a la familia estructural Z.</i>	53
4.2. <i>Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios correspondientes a la familia estructural Z.</i>	54
4.3. <i>Parámetros de las bases y pares de bases de los dDMPs complementarios de la familia estructural Z (véase designación en la fig. 2.10).</i>	54
4.4. <i>Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ, (véase designación en la fig. 2.8), para los mínimos de dDMP-Na+ complementarios con pares de Hoogsteen [22].</i>	57
4.5. <i>Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios con pares de Hoogsteen [22].</i>	58
4.6. <i>Parámetros de las bases y pares de bases de los dDMPs complementarios con pares de Hoogsteen [22] (véase designación en la fig. 2.10).</i>	59
4.7. <i>Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ, (véase designación en la fig. 2.8), para los mínimos de dDMP-Na+ complementarios con pares de Hoogsteen [21].</i>	60
4.8. <i>Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios con pares de Hoogsteen [21].</i>	61
4.9. <i>Parámetros de las bases y pares de bases de los dDMPs complementarios con pares de Hoogsteen [21] (véase designación en la fig. 2.10).</i>	62
4.10. <i>Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ, (véase designación en la fig. 2.8), para los mínimos de dApdT:dTpdA-Na+ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy.</i>	65

4.11. <i>Empaquetamiento del anillo de azúcar para los mínimos de dApdT:dTpdA-Na⁺ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy.</i>	65
4.12. <i>Parámetros de las bases y pares de bases de los dApdT:dTpdA-Na⁺ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy (véase designación en la fig. 2.10).</i>	66
4.13. <i>Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ, (véase designación en la fig. 2.8), para los mínimos de dApdA:dTpdT-Na⁺ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy.</i>	67
4.14. <i>Empaquetamiento del anillo de azúcar para los mínimos de dApdA:dTpdT-Na⁺ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy.</i>	68
4.15. <i>Parámetros de las bases y pares de bases de los dApdA:dTpdT-Na⁺ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy (véase designación en la fig. 2.10).</i>	69
4.16. <i>Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ, (véase designación en la fig. 2.8), para los mínimos de dDMP-Na⁺ complementarios con el par erróneo Gua:Thy [25].</i>	72
4.17. <i>Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na⁺ complementarios con el par erróneo Gua:Thy [25].</i>	72
4.18. <i>Parámetros de las bases y pares de bases de los dDMP-Na⁺ complementarios con el par erróneo Gua:Thy [25] (véase designación en la fig. 2.10).</i>	73
4.19. <i>Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ, (véase designación en la fig. 2.8), para los mínimos de dDMP-Na⁺ complementarios con el par erróneo Gua:Ade(syn) [26].</i>	74
4.20. <i>Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na⁺ complementarios con el par erróneo Gua:Ade(syn) [26].</i>	74
4.21. <i>Parámetros de las bases y pares de bases de los dDMP-Na⁺ complementarios con el par erróneo Gua:Ade(syn)[26] (véase designación en la fig. 2.10).</i>	75
4.22. <i>Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ, (véase designación en la fig. 2.8), para los mínimos de dDMP-Na⁺ complementarios con el par erróneo Ino:Ade(syn) [27].</i>	77
4.23. <i>Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na⁺ complementarios con el par erróneo Ino:Ade(syn) [27].</i>	78
4.24. <i>Parámetros de las bases y pares de bases de los dDMP-Na⁺ complementarios con el par erróneo Ino:Ade(syn)[27] (véase designación en la fig. 2.10).</i>	79

-
- 4.25. *Ángulos de torsión del esqueleto azúcar-fosfato, ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8) y valores promedio del ángulo entre los planos de las bases apiladas, ϕ , para los mínimos de dTpdT-Na+ con gran superposición entre Pyr-Pyr. 83*
- 4.26. *Empaquetamiento del anillo de azúcar y área de superposición entre los anillos de las bases apiladas para los mínimos de dTpdT-Na+ con gran superposición entre Pyr-Pyr. 83*

Resumen

«No sé lo que puedo parecer al mundo; pero a mí me parece que he sido sólo como un niño jugando en la orilla del mar, y yo en tanto encontrando un guijarro más liso o una concha más bonita que ordinaria, mientras el gran océano de la verdad pone todo lo que hay por descubrir delante de mí»

NEWTON (1642-1726)

El ADN es un biopolímero cuya estructura molecular-química es el resultado de un proceso selectivo natural durante millones de años. Es considerada la molécula biológica más importante para la vida, pues toda la información genética necesaria para el funcionamiento de un ser vivo se encuentra almacenada en ésta. Durante las últimas décadas, dada su vital importancia, se han estudiado sus características estructurales para comprender su funcionamiento y sus mecanismos de regulación.

En particular, se han investigado las estructuras de ADN de hélices dobles con cadenas antiparalelas del modelo de Watson-Crick (Familias Estructurales A y B). Sin embargo, éstas no son las únicas posibles conformaciones de dúplex de ADN que pueden participar en los procesos de funcionamiento de los ácidos nucleicos en los seres vivos. Por ello, en éste trabajo de tesis se sometieron a cálculos de optimización y se analizaron fragmentos mínimos de hélices dobles de ADN diferentes del modelo establecido por Watson-Crick; es decir, se estudiaron conformaciones de cadenas paralelas con el par reverso de WC Ade:Thy, dúplex de ADN con pares de bases Ade:Thy de Hoogsteen, dúplex de WC con los pares no correctos Gua:Thy, Gua:Ade(*syn*) e Ino:Ade(*syn*), así como estructuras de la Familia Conformacional Z. Lo anterior con objeto de obtener regularidades de las características conformacionales de dichas estructuras; las cuales no presentaron las mismas regularidades que los WCDs, tienen distintas interacciones que las estabilizan y dan pie a otras características conformacionales.

Así mismo, para analizar las diferencias en las características conformacionales obtenidas por los distintos métodos y funcionales, se hicieron optimizaciones mediante el método de la Teoría del Funcional de Densidad (DFT) con los funcionales PBE y PW91, y con Mecánica Molecular (empleando el campo de fuerzas *AMBER*).

Objetivos

*Algunas cosas son fáciles, sólo son
difíciles de hacer.*

A. EINSTEIN a J. O. HIRSCHFELDER

1. Objetivo General

Estudio de las características conformacionales de fragmentos mínimos de hélices dobles de ADN diferentes del modelo de Watson-Crick, por el método DFT con distintos funcionales, y comparación con resultados obtenidos por Mecánica Molecular.

2. Objetivos Particulares

- Preparar fragmentos mínimos de dúplex de ADN obtenidos del *Nucleic Acid Data Bank* [1], con cadenas paralelas, con pares de Hoogsteen y con pares no correctos; además de estructuras de la Familia Conformacional Z, para su estudio por método DFT.
- Optimizar la geometría de los fragmentos mínimos obtenidos, mediante cálculos de Mecánica Cuántica y Mecánica Molecular.
- Comparar los parámetros conformacionales y las diferencias entre las estructuras obtenidas por DFT (empleando los funcionales PBE y PW91) y Mecánica Molecular (con el campo de fuerzas AMBER).
- Obtener regularidades de las características conformacionales para las estructuras diferentes del modelo establecido por Watson-Crick. Comparar con características de familias de hélices dobles de Watson-Crick.

Capítulo 1

Introducción

*Si nada se hace de la mano de épocas
pasadas, el mundo debe permanecer siempre
en la infancia del conocimiento.*

CICERO

Hoy en día, la doble hélice del ADN es probablemente la molécula biológica más emblemática de todas. La primera descripción de la estructura de doble hélice del ADN con cadenas antiparalelas y pares de bases complementarias fue hecha por Watson y Crick, estableciendo así su modelo y publicándose el 25 de abril de 1953 en la revista científica *Nature* [2]. Además de este modelo, la estructura molecular del ADN facilita la formación de una serie de otras estructuras espaciales que son necesarias para el funcionamiento de esta biomolécula y que tienen importancia para la herencia.

La construcción del modelo de la doble hélice del ADN atrajo la atención de investigadores matemáticos, físicos y biólogos, iniciándose así un intenso desarrollo de la Biofísica Molecular. A partir de allí se sucedieron numerosos e importantes avances en el estudio de la biología y la medicina, entre ellos el diseño de tratamientos para enfermedades producidas por mutaciones en el ADN y la capacidad de detectar enfermedades hereditarias.

El entendimiento de los problemas genéticos a nivel molecular tiene como punto de partida el estudio de la estructura y la dinámica de los ácidos nucleicos; pero, no es posible observar los detalles estructurales del ADN directamente en las células, no obstante, cristales de ADN sintetizado presentan varias de las características que tiene éste en complejos con proteínas. Obtener las coordenadas de los átomos que componen a fragmentos de ADN es posible mediante la difracción de rayos X en cristales de dichos fragmentos y, de este modo, visualizar estructuras de dobles héli-



Figura 1.1: ADN sintetizado, cristalizado y microfotografiado por Linden Gledhill.

ces de ADN. Sin embargo, para comprender más profundamente la dinámica de los ácidos nucleicos son necesarios estudios más detallados, es por ello que se recurre a los cálculos con métodos de la Mecánica Cuántica (MC) y de la Mecánica Molecular (MM), para realizar optimizaciones de dichas estructuras y hacer un análisis más detallado de éstas (con cálculos matemáticos se busca poner a las estructuras en un estado de mínima energía, pues la naturaleza es tan sabia que busca estar en un estado en el que gaste la menor energía posible y eso es lo que las optimizaciones tratan de reproducir).

El análisis y comparación de las características geométricas y energéticas de las estructuras de ADN tomadas de los bancos de datos (en calidad de estructuras iniciales) y de las estructuras ya optimizadas mediante los diferentes métodos computacionales (con el uso de diversos programas computacionales) han permitido sacar conclusiones sobre las regularidades generales de la estructura y su relación con la función de la molécula principal de la vida. Algunas características de las estructuras de fragmentos de ADN obtenidos en investigaciones experimentales prácticamente coinciden con las que se obtienen por diferentes métodos de optimización sin embargo, otras difieren notablemente. El entender las razones de estas diferencias permitirán aproximarse a la construcción de los modelos de estructuras espaciales del ADN más parecidos a los existentes en la célula viva.

Los trabajos anteriores de cálculos utilizando el método de DFT de complejos de desoxidinucleósidos monofosfatos (dDMPs) con iones de sodios (Na^+) han demostrado que las características principales de las familias dúplex derechas de Watson-Crick (WC) están predefinidas en los mínimos locales de energía de los dDMPs [3]. Sin embargo, es poco lo que se conoce sobre estructuras de hélices dobles de ADN diferentes del modelo de Watson-Crick, y, es fundamental conocer todas las posibles conformaciones y los mecanismos de transformaciones entre ellas. Es por ello que en este trabajo de tesis se hizo un estudio de fragmentos mínimos de dúplex de ADN de cadenas paralelas con el par reverso de WC Ade:Thy, de confor-

maciones con pares de bases Ade:Thy de Hoogsteen, de dúplex de WC con los pares no correctos Gua:Thy, Gua:Ade(*syn*) e Ino:Ade(*syn*), así como de estructuras de la Familia Conformacional Z; esto, con objeto de obtener regularidades de las características conformacionales de dichas estructuras. Para analizar las diferencias entre las características conformacionales obtenidas por los distintos métodos y funcionales, se realizaron optimizaciones mediante el método de la Teoría del Funcional de Densidad (DFT) (con los funcionales PBE y PW91) y con Mecánica Molecular (empleando el campo de fuerzas *AMBER*).

Nota (Fig. 1.1): Estas fotografías fueron capturadas por el bioquímico y bioartista Linden Gledhill para un nuevo proyecto llamado *MSSNG*, un ambicioso programa lanzado por la organización de apoyo Autism Speaks. A medida que los científicos secuencian el ADN de 10.000 familias afectadas por el autismo, todos los datos recopilados estarán disponibles como fuente abierta a otros investigadores de todo el mundo en un intento de rellenar las piezas faltantes que rodean dicha condición. En su estudio, Gledhill utiliza un microscopio de investigación que puede ampliar hasta 1000 veces. Los bellos colores son el resultado de la utilización de la luz polarizante que luego es torcida por los cristales de ADN, causando una interferencia en los espectros de luz. Los patrones que se ven son el resultado visual de esa interferencia, capturada por una cámara instalada en un puerto extra en el microscopio. A medida que el agua se evapora de los bordes de las muestras de ADN colocadas entre dos placas de vidrio, la estructura cristaliza gradualmente. Las áreas oscuras son donde hay líquido y no hay estructura, pero a medida que las moléculas se alinean mejor vemos estos colores vivos.

1.1. Breve historia sobre el descubrimiento de la estructura del ADN [4]

La “victoria” de determinar la estructura del ADN correspondió a un cuarteto inverosímil de científicos de Inglaterra que no trabajaban como equipo, se enfadaban a menudo, no se hablaban y eran mayoritariamente novatos en ese campo: Maurice Wilkins y Rosalind Franklin del Colegio King de Londres, y Francis Crick y James Watson del Laboratorio Cavendish, de Cambridge (a éstos dos últimos es a quienes se les atribuye el descubrimiento).

La cuestión es que supusieron (correctamente, como se demostraría) que, si se podía determinar la forma de la molécula de ADN, se podría ver cómo hacía lo que hacía. Franklin tenía las mejores imágenes que existían de la posible estructura del ADN, conseguidas por medio de la cristalografía de rayos X, la técnica perfeccionada por Linus Pauling, pero, para constante irritación de Wilkins, se negaba a compartir sus descubrimientos.

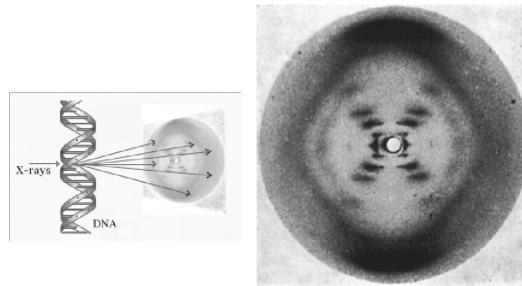


Figura 1.2: “Foto 51”, tomada por Franklin, que muestra un patrón de difracción de rayos X de ADN de la forma B

Que Wilkins y Franklin no congeniasen fue un hecho que Watson y Crick parece ser que explotaron en beneficio propio. Tras un “problema” que tuvieron¹, en enero de 1953, Wilkins mostró a Watson las imágenes de Franklin, al parecer sin que ella lo supiese ni lo consintiese² (figura 1.2). Watson y Crick, armados con el conocimiento de la forma básica de la molécula de ADN y algunos elementos importantes de sus dimensiones, redoblaron sus esfuerzos.

Se sabía que el ADN tenía cuatro componentes químicos (llamados adenina, guanina, citosina y timina) y que esos componentes se emparejaban de formas determinadas. Así que, jugando con piezas de cartón cortadas según la forma de las moléculas, Watson y Crick consiguieron determinar cómo encajaban las piezas (figura 1.3). A partir de ahí construyeron un modelo tipo Mecano, que consistía en placas metálicas atornilladas en espiral.

La edición del 25 de Abril de 1953 de *Nature* incluía un artículo de 900 palabras de Watson y Crick, titulado «Una estructura para el ácido desoxirribonucléico»³. Iba acompañado de artículos independientes de Wilkins y Franklin.

El descubrimiento de Watson y Crick no se confirmó, en realidad, hasta la década de los años ochenta. Como dijo Crick en uno de sus libros: «Hicieron falta veinticinco años para que nuestro modelo de ADN pasase de ser sólo bastante plausible a ser muy plausible... y, de ahí, a ser casi con seguridad correcto»⁴. En realidad casi no se

¹White, *Rivals*, p.257; Maddox, Brenda, *Rosalind Franklin: The Dark Lady of DNA*, London: HarperCollins, 2002, p. 185.

²Portal internet PBS, «A Science Odyssey», n.d.

³De Duve, *A Guided Tour of the Living Cell*, vol. 2, p.290

⁴Crick, Francis, *What Mad Pursuit: A Personal View of Scientific Discovery*, London: Penguin Press, p. 73-74

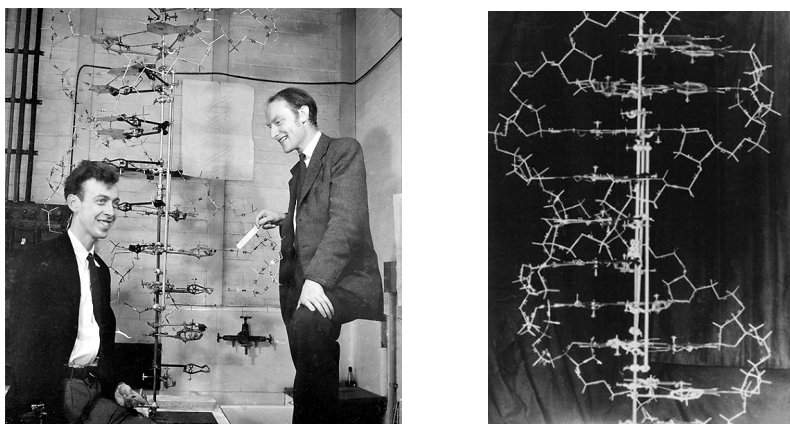


Figura 1.3: *Modelo de Watson y Crick de la molécula de ADN.*

había hecho más que empezar. Hoy día incluso hay muchas peculiaridades del ADN que apenas entendemos. Gran porcentaje del ADN consiste de largas extensiones de materia extraña aparentemente sin sentido... “basura” o “ADN sin código”. Sólo aquí y allá, a lo largo de cada hebra, encuentras secciones que controlan y organizan funciones vitales. Se trata de los curiosos genes, tan esquivos y escurridizos durante mucho tiempo.

Los genes son fragmentos de ADN y estos fragmentos corresponden a ARN *mensajero*, *ribosómico* o *de transferencia*, cuando los genes corresponden a ARNm tenemos instrucciones para hacer proteínas, por otro lado, el ARNr sólo se encarga del transporte de proteínas. Las fabrican con una fidelidad monótona y segura. En este sentido, son más bien como las teclas de un piano, que cada una de ellas da sólo una nota y nada más⁵, lo que es evidentemente un poco monótono. Pero, si combinas los genes, igual que haces con las notas del piano, puedes crear acordes y melodías de infinita variedad. Junta todos esos genes y tendrás la gran sinfonía de la existencia, conocida como el genoma humano.

⁵ *National Geographic*, «Secrets of the Gene», octubre de 1995, p. 55.

1.2. Importancia biológica de las conformaciones de ADN diferentes del modelo de Watson-Crick

La biomolécula más importante es el ADN ya que en ella se encuentra toda la información genética necesaria para que se lleve a cabo la vida y la reproducción en los sistemas biológicos. Ésta es una doble hélice que consta de dos cadenas antiparalelas complementarias de polinucleótidos (figura 1.4); a pesar de la gran importancia del ADN, su estructura molecular es relativamente sencilla: cada una de las cadenas antiparalelas consta de una secuencia de monómeros llamados nucleótidos —donde cada uno se compone de una base nitrogenada, un azúcar desoxirribosa y un grupo fosfato— que pueden ser de cuatro tipos, la diferencia entre ellos es la base nitrogenada que los conforma (purinas: Adenina, A, y Guanina, G; pirimidinas: Timina, T, y Citosina, C; ver Fig. 1.5); sin embargo, esto es suficiente para codificar la gran diversidad de la vida. Cada secuencia de tres nucleótidos constituye una palabra del código genético o codón, que especifica un aminoácido determinado.

Aunque la estructura del ADN es conservativa la hélice no es regular, las variaciones en los parámetros conformacionales a lo largo de los ejes de la hélice son responsables de la formación de diferentes regiones de ADN con funciones biológicas específicas. Entender el papel de los ácidos nucleicos como bloques de construcción en la formación, la dependencia de la secuencia de las bases y las preferencias conformacionales de la estructura tridimensional del ADN es un paso importante hacia el entendimiento de los procesos genéticos y la comprensión de la vida misma.

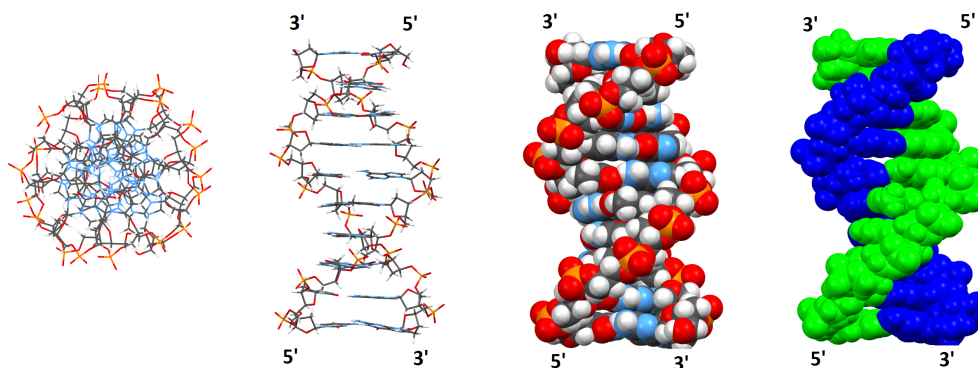


Figura 1.4: Estructura de la hélice doble del ADN. Fragmento de ADN de la familia B1, vista aérea (imagen a la izq.) y frontalmente con diferentes efectos.

Han sido revelados, en experimentos estructurales, varios patrones de la estructura tridimensional de los fragmentos de ADN, incluyendo las bien conocidas estructuras del modelo de Watson-Crick (familias de dobles hélices A y B), confor-

maciones con secuencias específicas (familia Z) y estructuras diferentes del modelo establecido por Watson-Crick (WC). Ésta importante información estructural obtenida ha sido acumulada en el *Nucleic Acid Data Bank* [1] y ha sido analizada por Svoil y col. [5]. Las familias y los fragmentos específicos de las secuencias de polinucleósidos dentro de cada familia WC se distinguen por los ángulos de torsión del esqueleto azúcar-fosfato y por el arreglo mutal de las bases.

Trabajos anteriores [3, 6, 7, 8] mostraron que la contribución del esqueleto azúcar-fosfato al paralelismo de las bases es dependiente de la secuencia de las bases, desoxidinucléósidos monofosfatos (dDMPs) con secuencias *purina-pirimidina* y *purina-purina* muestran una amplia superposición en los anillos de las bases, mientras que dDMPs con secuencias *pirimidina-purina* y *pirimidina-pirimidina* tienen superposición de los anillos de las bases despreciable o nula; además se encontró que las características principales de las familias de dúplex de Watson-Crick están predefinidas en la energía local mínima de dDMPs [3].

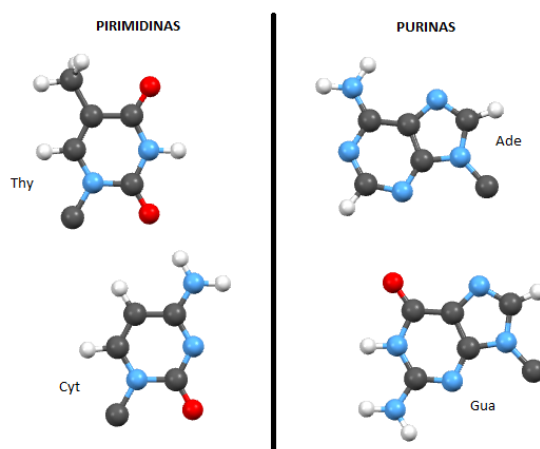


Figura 1.5: Clasificación de las cuatro bases nitrogenadas en Pirimidinas y Purinas.

Como ya se mencionó antes, se han realizado muchos estudios sobre dúplex de ADN del modelo de WC; sin embargo, dichas estructuras no son las únicas posibles estructuras que pueden participar en los procesos de funcionamiento de los ácidos nucleicos en los seres vivos. Hay resultados experimentales de dúplex de cadenas paralelas con el par reverso de WC Ade:Thy y de conformaciones con pares Ade-Thy de Hoogsteen, además, se tiene evidencia de regiones del genoma donde sólo pueden encontrarse dichas bases nitrogenadas; por otro lado, al tener mutaciones en el ADN se tiene la formación de pares no correctos, como Gua:Thy, Gua:Ade e Ino:Ade. Asimismo, se han encontrado estructuras de cadenas antiparalelas formando una doble hélice *zig-zagueada*.

1.2.1. Estructuras de la Familia Conformacional Z

Se han encontrado estructuras de la Familia Conformacional Z en solución y cristales, sobre su importancia biológica no está claro ya que no hay evidencias de éstas en complejos con proteínas, pero, es importante analizar dichas estructuras para ver si son posibles o sólo es un resultado *artificial*. Es posible distinguir a las estructuras con esta conformación de otras por sus características ópticas (luz polarizada), pues son distintas para cada Familia Conformacional del ADN.

1.2.2. Dúplex con el par Ade:Thy de Hoogsteen

Después de ser propuesto el modelo de ADN hubo muchos intentos para demostrar que el emparejamiento de WC también se daba en bases aisladas, en su lugar fue descubierto el par Ade:Thy de Hoogsteen, en 1959, en una estructura cristalina de complejo de 1-metil-Timina con 9-metil-Adenina [9].

Pares de bases con ésta conformación han sido reportados en algunos complejos de proteínas y ADN [10]. Estructuras de hélices dobles con pares de Hoogsteen se han encontrado en solución, por lo que es posible encontrarlos en células.

Téngase en cuenta que, los fragmentos de ADN con pares de Hoogsteen (y los de cadenas paralelas), tomados del banco de datos para nuestros estudios, contienen únicamente bases Adeninas y Timinas y es conocido que dihas bases predominan en regiones no codificantes del genoma, así que ADN de este tipo quizá juega un rol estructural en el genoma.



Figura 1.6: *Visitante de un museo en Nueva York observando una representación digital del genoma humano en el 2001.* (Cada color en la secuenciación del genoma humano representa una de las 4 bases nitrogenadas del ADN, identificando así los genes en cada cromosoma.)

1.2.3. Conformaciones de cadenas paralelas

Durante la Meiosis, en la Profase I (figura 1.7), se presenta la *Sinapsis* que es cuando dos cromosomas homólogos se aparean, lo cual permite el intercambio de material genético entre ambos cromosomas; al tener dos cromosomas se tienen dos hélices dobles de cadenas antiparalelas entre sí, por lo que se puede tener la formación de complejos con hélices dobles de cadenas paralelas; además, en la Sinapsis se puede tener el fenómeno *crossing-over*, Fig. 1.8, el cual nos da la posibilidad de formación de un cromosoma de madre con pedazo de cromosoma de padre y viceversa, que es lo que nos otorga mayor variabilidad genética como individuos.

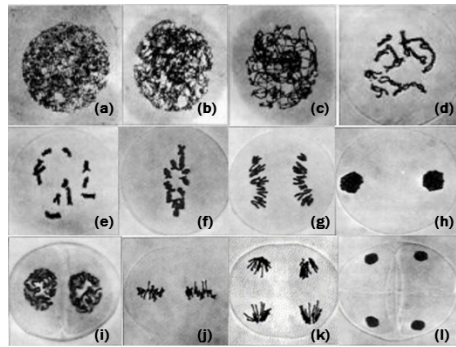


Figura 1.7: *Meiosis*. (Profase I: a, b, c, d y e. Metafase I: f. Anafase I: g. Telofase I: h. Profase II: i. Metafase II: j. Anafase II: k. Telofase II: l.)

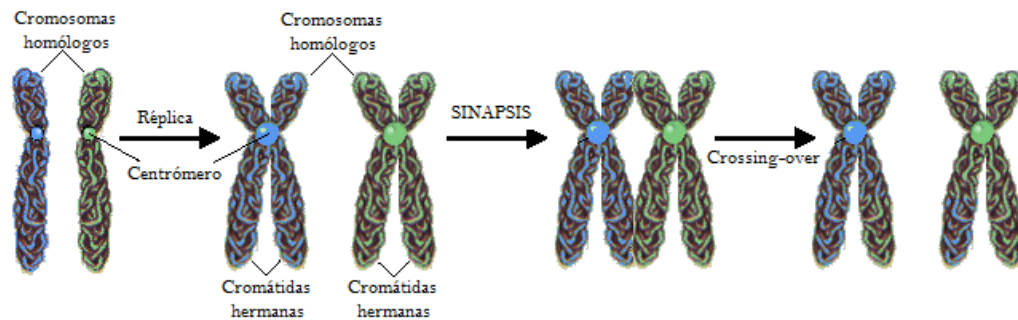


Figura 1.8: *Cromosomas homólogos y Cromátidas hermanas. Sinapsis y Crossing-over.*

Hay reportes de cadenas complementarias paralelas de ADN en diferentes genomas, por ejemplo, Tchurikov y colaboradores reportan una región complementaria paralela entre dos secuencias de ADN de la mosca *Drosophila* [11].

1.2.4. Dúplex de Watson-Crick con pares no correctos

Pares de bases no correctos se tienen como rutas de mutaciones puntuales (figura 1.9); esto es, durante la replicación del ADN se cambia una base por otra teniendo pares de bases como Gua:Thy, Gua:Ade(*syn*) e Ino:Ade(*syn*). Tales errores son generalmente detectados y eliminados por las enzimas de corrección de pruebas y de reparación de desajustes. Sin embargo, ocasionalmente, el error no se corrige dando lugar a mutaciones de sustitución y posiblemente a cambios evolutivos. Los pares de bases de *purina-purina* están específicamente implicados en mutaciones de transversión.

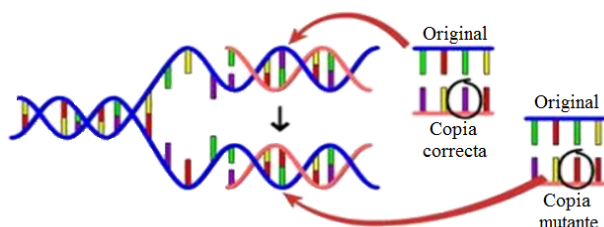


Figura 1.9: *Mutación Puntual.*

Se sabe que en el cáncer la replicación del ADN es muy rápida, lo cual da la posibilidad de formación de más mutaciones, por lo que es importante estudiar éstos mecanismos. El par Gua:Thy fue mostrado por primera vez en un artículo de Crick, formado en proceso de síntesis de proteínas (*Traducción*) [12]. El par de bases Gua:Ade(*syn*) fue propuesto por Traub y Sussman para ARNs ribosómicos [13]. Por otro lado, la desoxiinosina (Ino) es eficaz como base universal en sondas sintéticas de hibridación cuando se clonan genes para proteínas que contienen aminoácidos con codones degenerados [14, 15]. La Inosina forma desajustes estables, suele posicionarse en el extremo 5' de algunos anticodones de ARNt, donde se sabe que forma enlaces de hidrógeno con Adenina y Citosina [16]; también es capaz de ocupar la posición media del anticodón y formar par con Adenina.

Cabe mencionar que, anteriormente, se comprobó que las estructuras de dDMPs complementarios, Fig. 1.10, son los mínimos fragmentos de la hélice doble de ADN que nos aportan información estructural sobre ésta, y, debido al alto costo computacional de someter a cálculos estructuras de ADN, dichos fragmentos son los idóneos para realizar estudios teóricos computacionales de la estructura de dúplex de ADN [17].

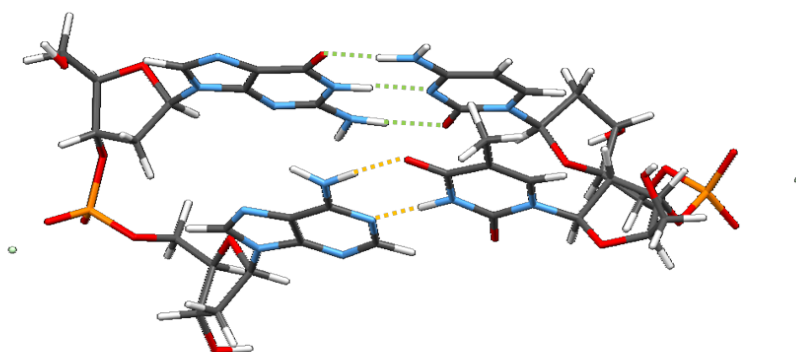


Figura 1.10: *Fragmento mínimo de ADN, ejemplo de dDMPs complementarios neutralizados con Na^+ , con los pares de bases Gua:Cyt y Ade:Thy.*

Debido a la ya mencionada importancia biológica de estas estructuras diferentes del modelo de WC, se sometieron a optimización fragmentos mínimos de dúplex con estas conformaciones, utilizando la Teoría del Funcional de la Densidad (DFT) empleando los funcionales PBE y PW91, y con cálculos de Mecánica Molecular con el campo de fuerzas AMBER. Las estructuras experimentales analizadas con pares Ade:Thy de Hoogsteen, con los pares no correctos Gua:Thy, Gua:Ade(*syn*) e Ino:Ade(*syn*), y las conformaciones de cadenas paralelas con el par reverso de WC Ade:Thy (obtenidas de resultados de Resonancia Nucleomagnética), presentan algunos de los ángulos de torsión del esqueleto azúcar-fosfato fuera de las regiones en que se encuentran éstos para los dúplex de WC (WCDs), después de ser optimizadas en algunas se encontraron todos sus ángulos dentro de las regiones esperadas para los WCDs, lo cual nos mostró la viabilidad de su estructura; además, se encontró que estas conformaciones presentan distintas regularidades a las de los WCDs, pues la superposición de sus bases apiladas varió, esto, debido a que presentan diferentes interacciones que las estabilizan con dichas regularidades. Por otro lado, los estudios realizados a los cdDMPs de la Familia Conformacional Z nos permitió corroborar los resultados obtenidos anteriormente para fragmentos mínimos de cadenas únicas de esta Familia, es decir, las estructuras de la Familia Conformacional Z se dan, y estabilizan, por las interacciones con el medio.

Capítulo 2

Estructura del ADN con conformación diferente del modelo de Watson-Crick

*La naturaleza le dirigirá una mentira directa
si puede.*

CHARLES DARWIN

Los ácidos nucleicos, las biomoléculas más importantes, contienen la información genética en instrucciones codificadas por la secuenciación de las bases nitrogenadas que los conforman. De su expresión depende la secuencia aminoacídica de las proteínas, de cuya síntesis son responsables inmediatos, y por ende, de la totalidad de la síntesis y el control de los procesos celulares. Cada tres bases nitrogenadas forman una letra del código genético, también llamadas triplete o *codón*, las cuales codifican un aminoácido.

El estudio de su estructura, funcionamiento y mecanismos de regulación, de la expresión genética, es esencial para comprender la integración de los procesos que sostienen la vida tanto a nivel celular como en organismos pluricelulares.

El ADN es una doble hélice que consta de dos cadenas antiparalelas complementarias de polinucleótidos, cada nucleótido está formado por tres unidades, Fig. 2.1: una molécula de azúcar llamada desoxirribosa, diferente de la ribosa del ARN por carecer de oxígeno en el carbono 2', lo que contribuye a una mayor estabilidad en la cadena, un grupo fosfato y una de las cuatro posibles bases nitrogenadas, donde dos son púricas (Adenina y Guanina) y dos son pirimidínicas (Timina y Citosina), ver Fig. 1.5, las cuales usualmente forman los pares Gua:Cyt y Ade:Thy. Cabe mencionar que un nucleósido sólo está formado por la desoxirribosa y una de las cuatro bases nitrogenadas.

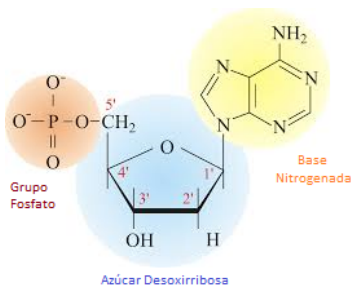


Figura 2.1: Estructura de un nucleótido

Por naturaleza, los ácidos nucleicos usualmente adoptan una conformación dúplex de cadenas antiparalelas con una cadena en la orientación $5' \rightarrow 3'$ y la otra en la orientación $3' \rightarrow 5'$. Además suelen formar los pares de bases establecidos en el modelo de Watson-Crick (figura 2.2).

Sin embargo, también tienen la capacidad de adoptar una conformación dúplex con cadenas paralelas (ambas cadenas con la orientación $5' \rightarrow 3'$), conformaciones dúplex de cadenas antiparalelas con pares de Hoogsteen o con pares no correctos, así como estructuras de doble hélice zig-zagueada conocidas como ADN-Z.

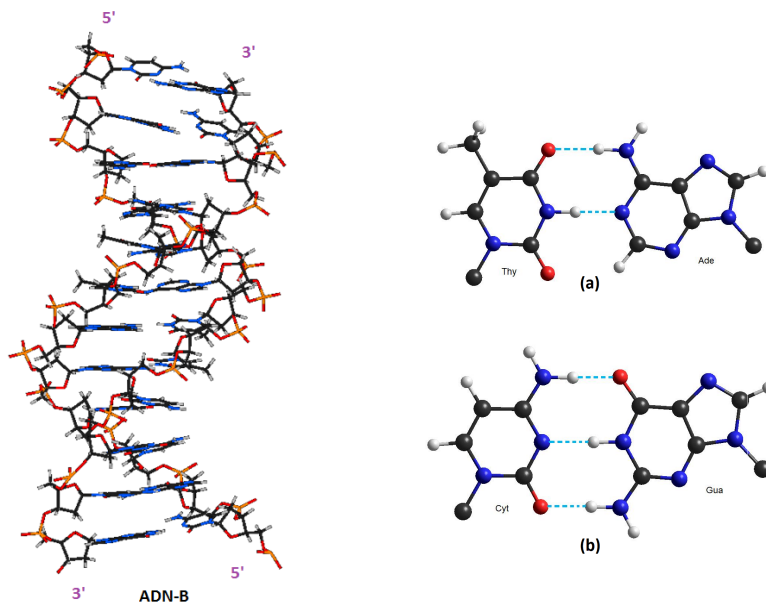


Figura 2.2: (IZQ.) Modelo de Watson-Crick con cadenas antiparalelas. (DER.) Pares de bases de WC. a) Par Thy:Ade. b) Par Cyt:Gua.

2.1. Fragmentos de dobles hélices de ADN

Experimentalmente, gracias a la difracción de rayos X en fragmentos de ADN cristalizados, se ha podido clasificar a las estructuras (secundarias) de ADN en 3 grandes Familias Conformacionales: Familia A, Familia B (ambas Dextrógiras) y la Familia Z (Levógira); ver Fig. 2.3. Nótese que el ADN de tipo B es el más abundante y es el descubierto por Watson y Crick; además, las estructuras pertenecientes a las Familias estructurales AI, AII, BI y BII son consideradas como estructuras del modelo de Watson-Crick. Pero, como ya hemos mencionado antes, éstas no son las únicas posibles conformaciones dúplex de ADN, se pueden tener estructuras de cadenas paralelas, estructuras con pares de Hoogsteen o con pares no correctos.

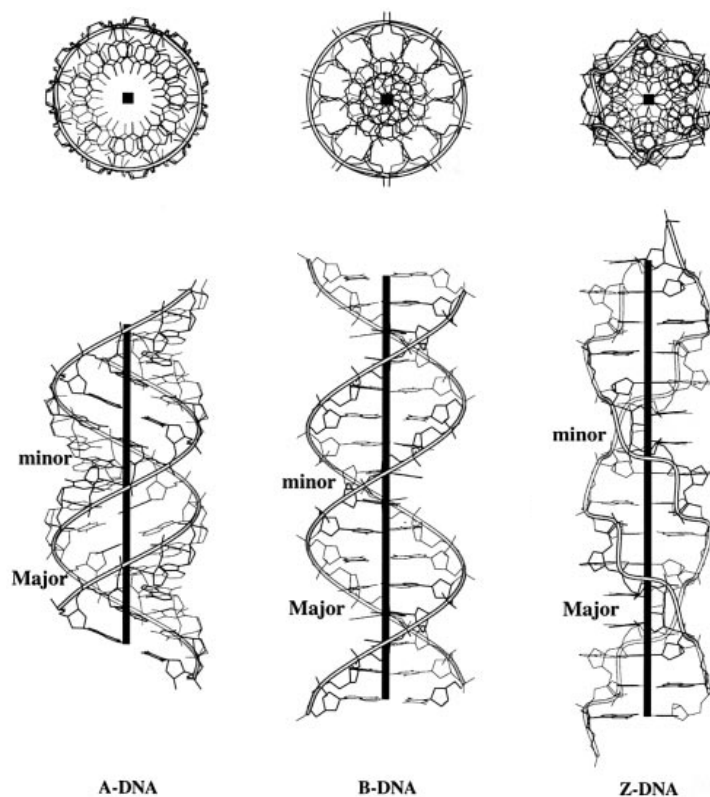


Figura 2.3: Vistas laterales y superiores que ilustran los rasgos característicos de las estructuras helicoidales regulares A, B y ADN-Z deducidas a partir de los modelos de difracción de rayos X de fibras representativas. Las cintas rastrean la progresión de la columna vertebral definida por los átomos de fósforo y las líneas negras gruesas (cajas) representan los ejes helicoidales [18].

FAMILIAS ESTRUCTURALES DEL ADN [19]:

- **ADN-A:** Tenemos ésta familia conformacional cuando el ADN, fibra de ADN, se encuentra en un medio con escasa agua. Ésta es una gran hélice, rechoncha con bases fuertemente inclinadas, dejando de ser perpendiculares al eje de la hélice, tienen un surco mayor cavernoso y un surco menor que es casi demasiado superficial para ser llamado un surco en absoluto.
- **ADN-B:** Ésta familia conformacional la obtenemos cuando tenemos el ADN en solución de cloruro de sodio con suficiente agua. Es más delgada y más alargada (comparando con ADN-A), con planos de las bases esencialmente paralelos entre sí y perpendiculares al eje de la hélice, y con un surco menor estrecho y surco mayor ancho de profundidad comparable. La mayor capacidad informativa se encuentra en el surco mayor como consecuencia de la mayor variabilidad de los grupos químicos que presenta: en el surco mayor, los grupos químicos presentes especifican la identidad del par de base, pudiendo distinguir hasta los pares de bases Ade:Thy de los pares Thy:Ade, permitiendo así que las proteínas reconozcan sin equivocación las secuencias de ADN sin tener que abrir la doble hélice.
- **ADN-Z:** Dicha familia conformacional es posible para secuencias específicas de bases, por ejemplo, para G-C-G-C-G (purina-pirimidina-purina...). La hélice Z es más delgada y más alargada para el mismo número de pares de bases que las otras dos familias conformacionales. En ésta el surco menor es cavernosamente profundo, y el surco mayor está completamente aplanado sobre la superficie de la molécula. Los dos rasgos característicos del ADN-Z son la inversión de la dirección de la hélice y el uso de dos pares de bases equivalentes no consecutivos como la unidad de repetición helicoidal, este último atribuye en el esqueleto de azúcar-fosfato una conformación en *zig-zag* que es responsable de la designación Z.

2.1.1. Fragmento de ADN de la Familia Conformacional Z

Aunque el ADN-Z es una de las tres Familias Estructurales del ADN, éste no pertenece al modelo establecido por Watson-Crick. En el ADN-Z se tiene a las bases Purinas en conformación *syn*, mientras que las Pirimidinas se encuentran en conformación *anti*, esto y la alternación entre Pirimidinas-Purinas es lo que le da la conformación *zig-zagueada* a la doble hélice; el par que se forma en estas estructuras es un par normal de WC (figura 2.4).

Para analizar la viabilidad de dichas estructuras, en este trabajo, se sometieron a cálculos de optimización fragmentos mínimos de doble hélice del dúplex hexámero

$d(\text{CGCGCG})_2$, tomado del banco de datos [1], determinado a una ultra alta resolución [20]; cabe mencionar que, el cristal de dicha estructura contiene espermina formando algunos enlaces de hidrógeno con átomos del fragmento de ADN-Z.

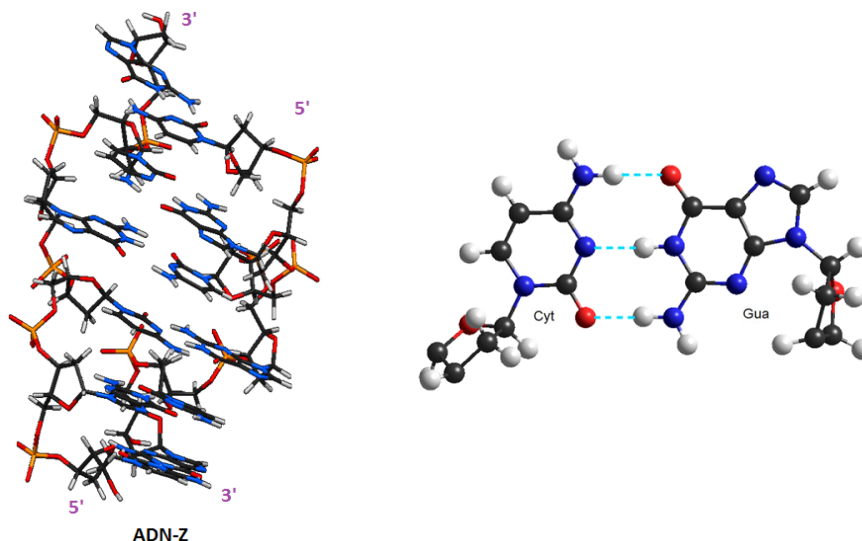


Figura 2.4: (IZQ.) Fragmento de ADN de la Familia Conformacional Z. (DER.) Conformación de nucleósidos en ADN-Z.

En la estructura de la molécula de ADN-Z estudiada, todos los pares de bases se presentan paralelos entre sí. Esta molécula es altamente rígida y ordenada, contradictoriamente al modelo de “ADN flexible”. La rigidez de este ADN-Z puede estar vinculada con su compacta estructura, reflejada en su pequeño diámetro (18\AA vs 20\AA en ADN-B), lo cual conduce a un empaquetamiento más apretado de los fosfatos y bases contra los azúcares.

2.1.2. Dúplex con pares de bases Ade:Thy de Hoogsteen

Después de ser propuesto el modelo de ADN fueron hechos muchos intentos para demostrar que el emparejamiento WC era favorecido por bases aisladas. Sin embargo, un esquema diferente de enlaces de hidrógeno entre Ade:Thy fue descubierto, primero descrito por Hoogsteen usando rayos X de cristalografía [9].

Pares de Hoogsteen han sido observados en varias hélices triples incluyendo aquéllas de importancia biológica. Pueden ser vistos en estructura cristalina que representa reconocimiento de ADN por la *proteína p53* supresora de tumores. Además, ocasionalmente, estos pares son encontrados en algunos complejos de ADN con

drogas y proteínas. También es probable que algunas regiones del genoma puedan tener esta conformación *in vivo* bajo las condiciones apropiadas, ya que la transición entre ambas formas (WC y Hoogsteen) tiene un muy bajo costo energético.

Comparando con el par Ade:Thy de WC, en el par de Hoogsteen la base Adenina está rotada 180° teniéndola así en conformación *syn*, mientras que en el par de WC se encuentra en conformación *anti* (figura 2.5), con lo cual se tiene un distinto conjunto de enlaces de H; además, una de las diferencias más significativas entre los pares de WC y los de Hoogsteen es la distancia entre $C1' - C1'$ del azúcar a través de cada par de base, ya que en los pares de Hoogsteen se acorta aproximadamente 2\AA .

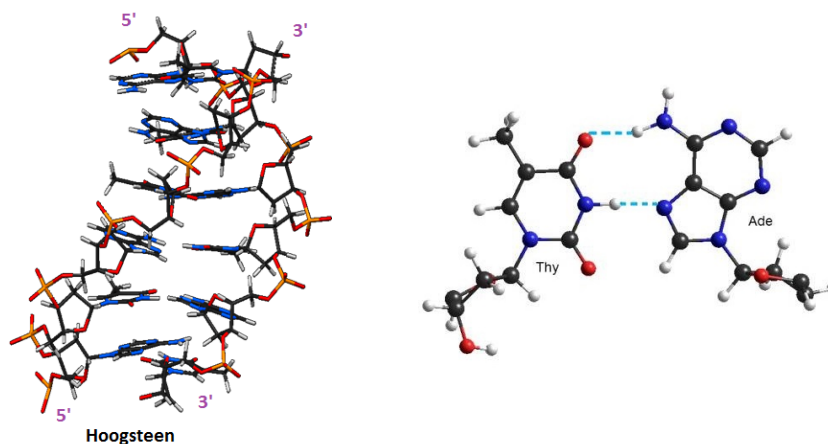


Figura 2.5: (IZQ.) Fragmento de ADN con pares de Hoogsteen. (DER.) Par Thy:Ade de Hoogsteen.

Las estructuras analizadas en este trabajo, con pares de Hoogsteen, contienen sólo pares Ade:Thy y, es conocido que las secuencias promotoras de muchos genes estructurales eucariotas se componen de tramos de Adenina y Timina. Dichas estructuras experimentales, tomadas del banco de datos, contienen las secuencias $d(ATATAT)_2$ [21] y $d(ATTAAT)_2$ [22].

2.1.3. Modelo de conformaciones de ADN de cadenas paralelas con el par reverso de WC Ade:Thy

La estructura dextrógira de hélices dobles de ADN, más conocida como ADN-B, ha jugado un rol central en la explicación de varios procesos biológicos. Estructuras alternativas pueden ser estabilizadas por interacciones favorables por apilamiento y enlaces de hidrógeno diferentes a los que se tienen en pares de bases del modelo de Watson-Crick.

Como se ha mencionado anteriormente, los ácidos nucleicos suelen adoptar una conformación dúplex de cadenas antiparalelas con una cadena en la orientación $5' \rightarrow 3'$ y la otra en la orientación $3' \rightarrow 5'$, como sea, también tienen la capacidad de adoptar una conformación dúplex con cadenas paralelas (parallel-stranded, *ps*), ver Fig. 2.6. Tal *ps* ADN podría desempeñar un importante rol en la regulación de replicación y transcripción, recombinación genética, plegamiento de cromosomas, procesos de mutación y empalme de ARN. Hay varios reportes teóricos y experimentales de la formación de dúplex con *ps* ADN.

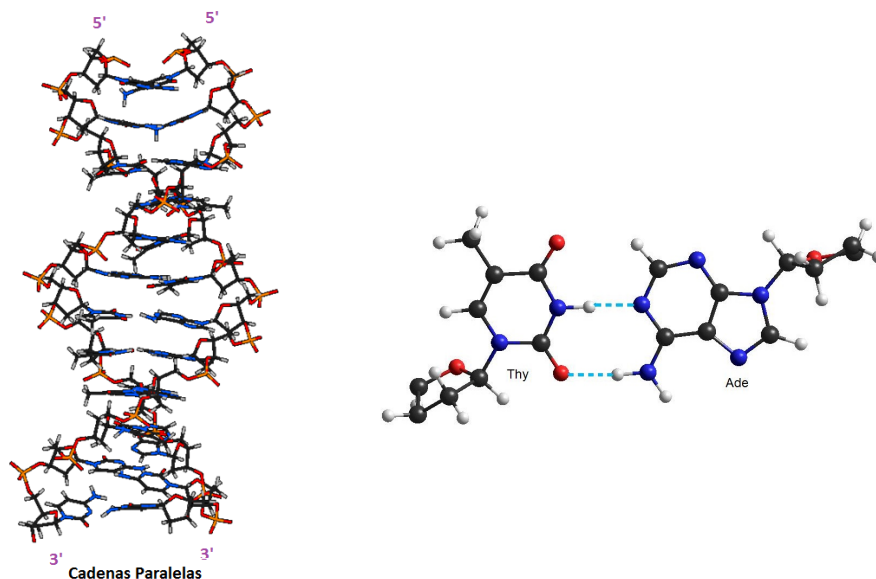


Figura 2.6: (IZQ.) Modelo de dúplex de ADN de cadenas paralelas. (DER.) Par reverso de WC Thy:Ade, presente en las estructuras de cadenas paralelas.

Para realizar el estudio teórico-computacional de éstas, se tomaron del banco de datos estructuras simuladas de *ps* dúplex. Las estructuras tridimensionales de dúplex de cadenas paralelas fueron obtenidas usando simulaciones de dinámica molecular bajo restricciones de Resonancia Nucleomagnética (NMR) [23]. Dichas

estructuras contienen dos pares de bases C:C⁺ en los dos extremos del dúplex paralelo, donde cada par tiene 3 enlaces de H, lo cual estabiliza la formación del *ps* ADN. Además, la superposición de las bases provee una fuente adicional de estabilidad en los dúplex de ADN. Cabe mencionar que, se tienen 20 modelos de *ps* ADN con longitudes de enlace de H y ángulos aceptables, y viabilidad estereocímica de los ángulos de torsión.

Los ángulos de torsión en éstas estructuras son semejantes a los que se tiene en el ADN-B, pero los parámetros helicoidales y la superposición de las bases apiladas son significativamente diferentes. Por otro lado, forman pares de bases de WC reversos (Ade:Thy), donde ambas bases se encuentran en conformación *anti* (figura 2.6).

2.1.4. Dúplex de ADN del modelo de WC con pares no correctos

Una de las tareas más importantes de la célula es la síntesis de proteínas, moléculas que intervienen en la mayoría de las funciones celulares, éstas están formadas por aminoácidos que determinan su estructura y función. El ADN es quien contiene la información necesaria para dirigir la fabricación de proteínas, ya que la secuencia de aminoácidos está determinada por la secuencia de bases de los nucleótidos del ADN. Cada secuencia de tres bases nitrogenadas, dTDPs, constituye una palabra del código genético o *codón*, que especifica un aminoácido determinado.

Un gen es una secuencia de nucleótidos de ADN que especifica el orden de aminoácidos de una proteína por medio de una molécula intermediaria de ARNm. La sustitución de un nucleótido de ADN por otro que contiene una base distinta hace que todas las células o virus descendientes contengan esa misma secuencia de bases alterada. Como resultado de la sustitución, también puede cambiar la secuencia de aminoácidos de la proteína resultante. A esta alteración de una molécula de ADN es a lo que llamamos *mutación*. Casi todas las mutaciones son resultado de errores durante el proceso de replicación del ADN.

Cuando se tienen mutaciones puntuales, se puede tener la formación de pares no correctos ya que se cambia una base por otra, y dada la importancia de la secuencia de nucleótidos en el ADN, es indispensable estudiar estructuras con dichos pares. La importancia de los pares de bases no complementarias y su posible implicación en la mutación espontánea fue reconocida por Watson y Crick en una de sus primeras publicaciones [2]. Pares de bases no complementarias pueden incorporarse en el ADN durante la replicación, durante la recombinación genética o, en el caso del par erróneo Gua:Thy, por la desaminación de 5-metil citosina [24].

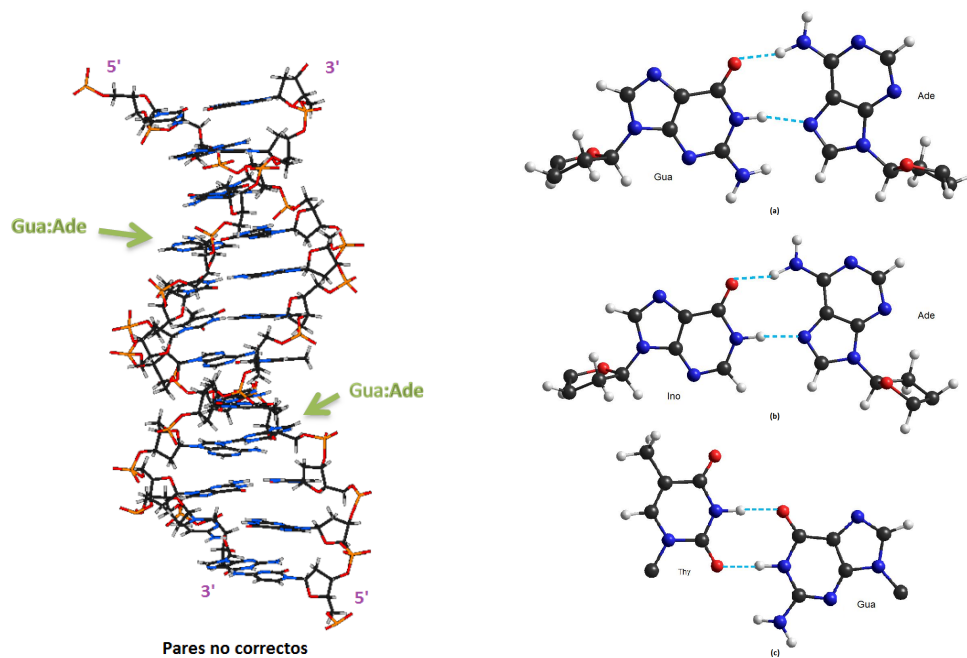


Figura 2.7: (IZQ.) *Fragmento de ADN con dos pares de bases no correctos, Gua:Ade.* (DER.) *Pares de bases no correctos. a) Par Gua:Ade. b) Par Ino:Ade. c) Par Thy:Gua.*

En éste trabajo se analizaron dúplex con tres pares no correctos: Gua:Thy [25], Gua:Ade [26] e Ino:Ade [27]; cada uno de los tres fragmentos de ADN tomados del banco de datos contienen dos pares no correctos del mismo tipo (figura 2.7). Nótese que en el par Gua:Thy ambas bases se encuentran en conformación *anti*, igual que el par Gua:Cyt de WC; mientras que en los pares Gua:Ade e Ino:Ade las Adeninas se encuentran en conformación *syn*, Guanina e Inosina están en conformación *anti*.

2.2. Parámetros geométricos para el estudio del ADN

Para el estudio del ADN nos podemos valer de distintos parámetros geométricos que nos aportan información estructural sobre éste. En este trabajo analizamos los ángulos de torsión del esqueleto azúcar-fosfato (Fig. 2.8), el empaquetamiento del azúcar (Fig. 2.9), además de los parámetros de posición mutua de las bases *h-Twist*, *h-Rise*, *Buckle*, *Propeller*, *Tilt*, *Roll* (Fig. 2.10) y el área de superposición de las bases debido al *Stacking*.

Para los ácidos nucleicos son especificados seis ángulos principales de torsión de la cadena azúcar-fosfato (figura 2.8). Uno de estos ángulos de torsión, δ , es particularmente informativo pues identifica el empaquetamiento del azúcar o conformación del anillo de desoxirribosa, un parámetro que ha jugado un papel importante en estudios anteriores de la estructura del ADN.

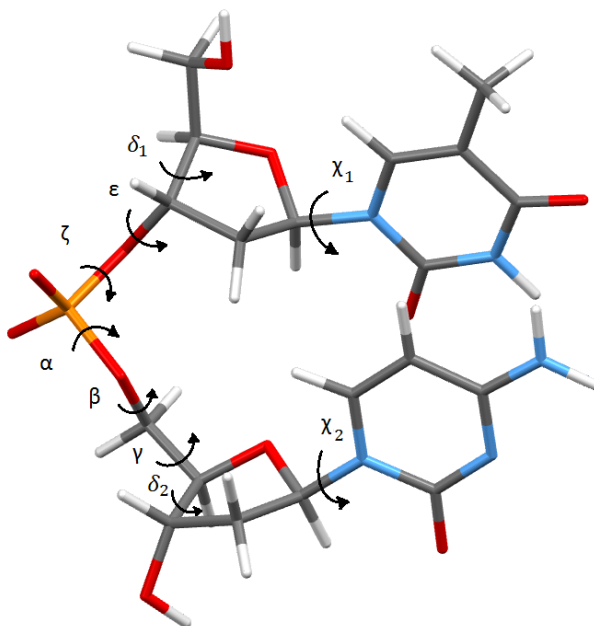


Figura 2.8: Designación de los ángulos de torsión de un ddMP.

En el anillo de desoxirribosa uno de sus átomos (de los cinco átomos que lo conforman) se encuentra fuera del plano, ya sea a un lado o al otro, dependiendo de la conformación se les designa como Xn' -endo o Xn' -exo, donde Xn' identifica el átomo y *endo* y *exo* describen si está fuera de planicidad en la misma dirección que el átomo $C5'$ o en el lado opuesto (figura 2.9). La conformación $C3'$ -endo o $C2'$ -exo tradicionalmente se ha asumido para el ADN-A, y $C2'$ -endo o $C3'$ -exo para el ADN-B, pero esto es una simplificación [19].

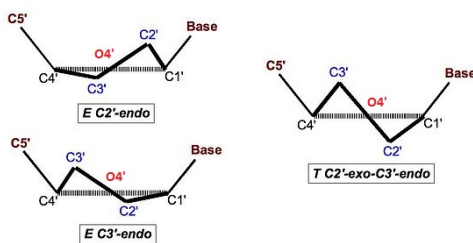


Figura 2.9: *Empaquetamiento del azúcar.*

El otro ángulo de torsión importante o relevante es el ángulo glucosídico χ , entre el átomo $C1'$ del anillo de azúcar y el nitrógeno de su base adjunta (corresponde a la posición mutua base-azúcar, ver figura 2.8). Cuando el anillo se gira de manera que se dobla hacia el surco menor, la conformación se describe como *syn*, y χ tiene un valor alrededor de $+70^\circ$. En contraste, en la conformación *anti*, menos restrictiva, el anillo de azúcar se dobla fuera del surco menor, χ entonces es libre de adoptar una amplia gama de valores, desde -80° hasta -180° . La peculiar cadena con trayectoria *zig-zag* en el ADN-Z surge debido a la alternancia de conformaciones *syn* y *anti* en guaninas y citosinas, respectivamente. En todas las demás hélices de ADN, del modelo de WC, sólo se encuentra la conformación *anti*. Nótese que, en las estructuras analizadas de pares no correctos Purina-Purina, las Adenas se encuentran en conformación *syn*, al igual que en el par Ade:Thy de Hoogsteen.

Los otros cinco principales ángulos de torsión de la cadena azúcar-fosfato también varían en cierta medida entre las tres familias (WC) de hélice, pero no son tan característicos o de diagnóstico como lo son los ángulos δ y χ .

Los parámetros *Buckle* y *Propeller* (ver Fig. 2.10, a la izquierda) describen las imperfecciones, es decir, la no planaridad de determinado par de bases [18]. En términos generales, *Propeller* siempre tiene un valor de aproximadamente -10° en A y B-ADN, mientras que *Buckle* puede ser positivo o negativo hasta 20° .

Se requieren seis parámetros de cuerpo rígido (tres rotaciones y tres traslaciones) para describir la posición y la orientación de un par de bases con relación a otro. Dos conjuntos de dichos parámetros se utilizan comúnmente en la literatura: el conjunto de parámetros locales de secuencias de pares de bases, *Shift*, *Slide*, *Rise*, *Tilt*, *Roll*, (ver Fig. 2.10, en medio), y *Twist*, que describen la geometría del apilamiento de una secuencia de dinucleótidos desde una perspectiva local; y el conjunto de parámetros helicoidales, X-desplazamiento, Y-desplazamiento, *Inclination*, *Tip*, *helical Rise* y *helical Twist* (ver Fig. 2.10, a la derecha), que describen la regularidad de la hélice. Por ejemplo, *h-Twist* es el ángulo de rotación alrededor del eje

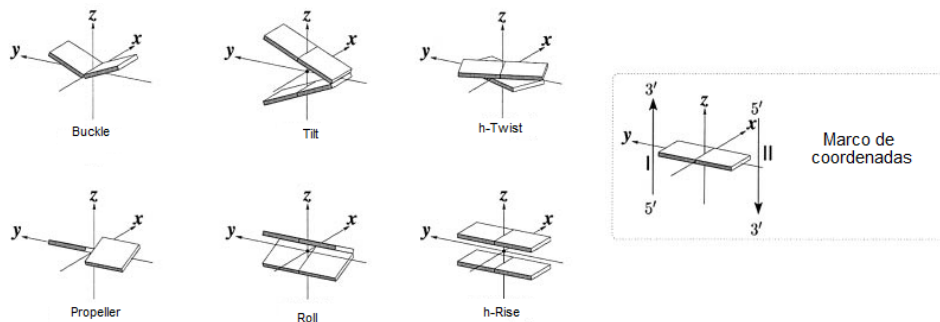


Figura 2.10: Descripciones gráficas de los parámetros *Buckle*, *Propeller*, *Tilt*, *Roll*, *h-Rise* y *h-Twist*, empleados para describir la geometría de pares de bases complementarios y secuencias de pares de bases.

helicoidal que tienen pares de bases sucesivos. Resultados obtenidos por análisis de rayos X de cristales de ADN sugieren valores de *h-Rise* alrededor de $2.83 \pm 0.36 \text{ \AA}$ para A-ADN y $3.29 \pm 0.21 \text{ \AA}$ para B-DNA [18]. En el ADN-B *Tilt* y *Roll* tienen valores nulos, ya que los pares de bases son paralelos entre sí.

Los enlaces de hidrógeno y las interacciones por el *Stacking* de las bases nitrogenadas son fundamentales para la organización del ADN y ARN, dando a ambos sus estructuras tridimensionales y modos de reconocimiento, es lo que le da estabilidad a las dobles hélices. El *Stacking* cara a cara de los anillos aromáticos depende sutilmente de la secuencia de nucleótidos, ver Fig. 2.11. Se ha comprobado anteriormente que, para las estructuras del modelo de WC, el grado de superposición entre *purina-pirimidina* y *purina-purina* es mayor que el que hay entre *pirimidina-purina* y *pirimidina-pirimidina* [28].

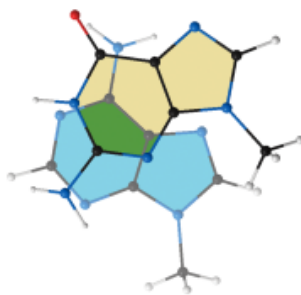


Figura 2.11: Arreglo de bases G-A apiladas. La base con color azul en sus anillos se encuentra detrás de la base con color amarillo en sus anillos, con color verde se muestra el área de superposición de los anillos.

Para cada Familia Estructural del modelo de WC se tienen los ángulos de torsión del esqueleto azúcar-fosfato en regiones específicas [28]. Por ello, con objeto de analizar la viabilidad de la estructura de las conformaciones de cadenas paralelas, de dúplex con pares de Hoogsteen o pares erróneos y de ADN-Z, en el presente trabajo, se estudiaron dichos ángulos comparándolos con las regiones en que se encuentran estos para los dúplex de WC. Mediante los parámetros *Buckle*, *Propeller*, *Tilt* y *Roll* se obtuvo información sobre la posición mutual de sus bases y pares de bases. Además, con el área de superposición de los anillos de las bases debida al *Stacking* se analizó si estas conformaciones presentan o no las mismas regularidades que los dúplex de WC.

Capítulo 3

Metodología para el estudio Teórico-Computacional del ADN

*Cuanto más precisos se hagan los cálculos,
más tienden los conceptos a deshacerse
en el aire fino.*

R. S. MULLIKEN, J.C.P. 43, S2 (1965).

Es indiscutible que, gracias a su riqueza estructural, el ADN juega un papel esencial en los procesos de funcionamiento de los seres vivos. Pero, ¿cuáles son los factores que permiten dicha riqueza estructural?. Para poder responder esta pregunta los científicos se han valido de dos diferentes métodos para el estudio del ADN. El primer método consiste en técnicas experimentales tales como: difracción de rayos X en cristales de fragmentos de ADN y Resonancia Núcleo Magnética (NMR), éstas técnicas son útiles para visualizar la estructura en ciertas condiciones pero no nos permiten conocer detalladamente las interacciones moleculares que existen entre las subunidades del ADN. El segundo método consiste en estudios computacionales, el cual es excelente para conocer las interacciones y analizar las estructuras más favorables de las moléculas.

Existen dos tipos de modelos moleculares empleados para realizar los cálculos computacionales: los modelos Mecánico Moleculares y los modelos Mecánico Cuánticos. El primero permite trabajar con moléculas grandes de una manera bastante sencilla y a costos accesibles, mientras que el segundo sólo se puede aplicar a pequeños fragmentos de moléculas y su costo en tiempo y dinero es muchísimo mayor que el anterior. Sin embargo, dado que los sistemas que estudiamos no son tan complejos (sistemas moleculares de 126 a 129 átomos), es posible hacer uso de ambos tipos de cálculos. En este trabajo se emplearon ambos modelos para optimizar los desoxidinucléosidos monofosfatos complementarios (cdDMPs), que se detallan en la Sección 3.3.1, como unidad mínima de los dúplex de ADN distintos del modelo de Watson-Crick.

3.1. Mecánica Cuántica

El Método de Mecánica Cuántica (MC) se basa principalmente en resolver la ecuación de Schrödinger. Sin embargo, debido a que ésta sólo puede ser resuelta exactamente para el problema de dos cuerpos, como en el caso del átomo de hidrógeno donde sólo interactúan el núcleo y un electrón, han surgido algunos métodos para trabajar con problemas más complejos como lo son los sistemas de moléculas orgánicas que nosotros estudiamos.

Uno de los conjuntos de métodos más conocidos y utilizados son los métodos *ab initio* (de primeros principios), los cuales se derivan directamente de principios teóricos, sin incluir parámetros empíricos en sus ecuaciones, es decir, sin tomar en cuenta datos experimentales. Las aproximaciones en este caso son estrictamente matemáticas, buscan una forma más simple de las funciones empleadas, para obtener soluciones aproximadas a las ecuaciones diferenciales complicadas.

En los métodos *ab initio* se obtiene la estructura electrónica de la molécula resolviendo la ecuación de Schrödinger independiente del tiempo para cada átomo o partícula:

$$\hat{H}\Psi = E\Psi \quad (3.1)$$

donde \hat{H} es el Hamiltoniano del sistema, operador correspondiente a la observable energía (cinética y potencial), esto es, que representa la energía total del sistema (formalmente es definido como un operador autoadjunto que actúa sobre un dominio denso en el espacio de Hilbert del sistema); E es el valor propio de \hat{H} y Ψ es la función de onda que nos habla del comportamiento del sistema (representa a los orbitales atómicos).

En general, para simplificar los cálculos realizados con estos métodos, se utiliza como primera aproximación la representación de *Born-Oppenheimer*. Esta aproximación permite la separación de los movimientos de los electrones y los núcleos, y considera la geometría de las moléculas como un parámetro adiabático, es decir, que se da mediante cambios infinitamente lentos en el Hamiltoniano del sistema.

3.1.0.1. Ecuación de Schrödinger y la aproximación de Born-Oppenheimer

Para determinar las propiedades de un sistema cuántico, como lo son los átomos y moléculas, es necesario emplear la ecuación de Schrödinger:

$$\left[-\frac{\hbar^2}{8\pi^2m} \nabla^2 + V \right] \psi(r, t) = \frac{i\hbar}{2\pi} \frac{\partial \psi(r, t)}{\partial t} \quad (3.2)$$

donde ψ es la función de onda para una partícula de masa m ; \hbar la constante de Planck y V el potencial. ψ no tiene interpretación física pero el producto de ψ con su complejo conjugado (ψ^*) se entiende como la distribución de probabilidad de la partícula a la que está asociada.

Cuando el potencial, V , no es una función del tiempo la ecuación de Schrödinger puede simplificarse. Empleando el método de separación de variables, es decir, escribiendo a la función de onda como una multiplicación de dos funciones, una espacial y otra temporal.

$$\psi(r, t) = \psi(r)\tau(t) \quad (3.3)$$

Sustituyendo esto en la ecuación de Schrödinger, con un potencial independiente del tiempo y aplicando el método de separación de variables llegamos a la ecuación de Schrödinger independiente del tiempo:

$$H\psi(r) = E\psi(r) \quad (3.4)$$

donde E es la energía de la partícula, y H es el operador Hamiltoniano igual a:

$$H = -\frac{\hbar^2}{8\pi^2m}\nabla^2 + V \quad (3.5)$$

Las soluciones a la Ec. 3.4 corresponden a diferentes estados estacionarios de la partícula (o molécula). Aquella con la energía más baja es llamada el estado base.

Para sistemas moleculares o iónicos el cálculo de la función de onda y de la energía se vuelve una tarea tediosa. Para facilitar esta tarea comúnmente se emplea la *aproximación de Born-Oppenheimer* (B-O). En términos básicos esta aproximación nos permite dividir el problema molecular en dos partes: una electrónica y otra nuclear.

$$\psi_{total} = \psi_{ele}\psi_{nuc} \quad (3.6)$$

Esta separación es válida ya que los electrones son ligeros en comparación con los núcleos, por lo que la velocidad de estos últimos es correspondientemente pequeña. Por eso podemos considerar que los electrones se mueven alrededor de los núcleos, mientras estos últimos permanecen estáticos. De este modo podemos escribir un Hamiltoniano puramente electrónico, H^{elec} . Entonces la ecuación de Schrödinger independiente del tiempo para describir el movimiento de electrones en el campo de núcleos fijos queda como:

$$H^{elec}\psi^{elec}(r, R) = E^{eff}\psi^{elec}(r, R) \quad (3.7)$$

donde ψ depende de las coordenadas de los electrones, r , y de las de los núcleos, R . Resolviendo esta ecuación para la función electrónica producirá la función potencial efectiva, E^{eff} , que depende de las coordenadas nucleares, R , y que describe la superficie potencial del sistema.

Por otro lado, si estamos interesados en calcular movimientos nucleares: rotaciones, vibraciones y traslaciones; se utiliza la función E^{eff} para determinar el hamiltoniano nuclear que ha de usarse en la ecuación de Schrödinger correspondiente.

$$H^{nucl} = T^{nucl}(R) + E^{eff}(R) \quad (3.8)$$

Sin embargo, a pesar de que estas consideraciones simplifican la Ec. 3.1, no son suficientes y es necesario hacer algunas otras aproximaciones. Una de ellas es considerar los orbitales moleculares como una combinación lineal de los orbitales atómicos. Esto es, que el conjunto base sea construido como una combinación lineal de los orbitales atómicos del sistema (LCAO-Linear Combination of Atomic Orbitals).

Luego de aplicar estas dos simplificaciones se hace necesario adoptar algún otro método que permita resolver las ecuaciones necesarias. El método más común de los cálculos con *ab initio* para la estructura electrónica es el de Hartree-Fock (HF), que es un método iterativo. En ellos la repulsión electrón-electrón, correspondiente al efecto de *correlación*, no se toma en cuenta directamente, sólo se incluye su efecto como un promedio de todas las interacciones.

El primer paso en este método es la elección de un conjunto de orbitales iniciales aproximados que representen el movimiento de los electrones. En general, para los cálculos de un átomo, estos orbitales o funciones de onda de sus electrones son típicamente los orbitales del átomo de hidrógeno, pero, para cálculos de moléculas las funciones aproximadas iniciales de un electrón son una combinación lineal de los orbitales atómicos, que debido a su naturaleza fermiónica, deben ser anti-simétricos. Esto se logra con el determinante de Slater [29, 30], el cual se define para un sistema de N electrones de la siguiente manera:

$$\Psi(\phi_1, \phi_2, \dots, \phi_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_N(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_N) & \phi_2(x_N) & \cdots & \phi_N(x_N) \end{vmatrix} \quad (3.9)$$

donde $\Psi(\phi_1, \phi_2, \dots, \phi_N)$ es la función de onda molecular; ϕ_i son las funciones de onda de cada electrón en la molécula y son conocidas como espín-orbital; x_i indica la posición y el espín del orbital; $1/\sqrt{N}$ es un factor de normalización. Esta función de onda es asimétrica y no hace distinciones entre electrones. Además, se hace cero si dos funciones o dos electrones son iguales, lo que equivale a satisfacer el principio de exclusión de Pauli.

Sobre estas bases, se construye entonces un nuevo operador Hamiltoniano, llamado operador Fock, para el cual, sus primeros términos son la suma de la energía cinética para cada electrón, la energía de repulsión entre núcleos y una suma de los términos coulombianos de atracción electrón-núcleo. El último término modela la repulsión coulombiana entre cada uno de los electrones de la suma. La suma se compone de una energía de repulsión neta para cada electrón en el sistema, la cual se calcula tratando a todos los demás electrones de la molécula como una distribución suave de la carga negativa.

El operador de Fock, que se construye con la descripción anterior, se utiliza como el nuevo Hamiltoniano en la ecuación de Schrödinger independiente del tiempo, Ec. 3.1. Su solución lleva a un nuevo conjunto de orbitales aproximados para cada electrón, con los cuales se vuelve a construir otro operador de Fock y se sigue el ciclo hasta obtener una solución para la cual el cambio en la energía sea mínimo. En el proceso se calcula un conjunto de orbitales de un electrón llamados auto-consistentes. La función de onda electrónica de HF es igual al determinante de Slater de estas funciones de onda de un electrón aproximadas. El operador de Fock para un electrón se muestra en la Ec. 3.10.

$$\hat{F} = -\frac{1}{2}\nabla_i^2 - \sum_{\alpha} \frac{Z_{\alpha}}{\|\vec{r}_{i\alpha}\|} + \sum_{j=1}^{N/2} [2\hat{J}_j(i) - \hat{K}_j(i)] \quad (3.10)$$

donde el primer término corresponde a la energía cinética del electrón i ; el segundo término es la energía potencial de la atracción entre el electrón i y los núcleos α ; \hat{J} es el operador Coulombiano que define la energía de repulsión electrón-electrón que se debe al electrón j -ésimo y \hat{K} es el operador de *intercambio* que define la energía del intercambio electrónico. Estos dos últimos términos corresponden a un par de integrales que nos dan el promedio de los dos efectos mencionados. El factor de 2 que aparece con el término \hat{J} se debe a que existen 2 electrones en cada orbital espacial y la sumatoria se hace tomando en cuenta los $N/2$ orbitales ocupados.

De este modo, las ecuaciones a resolver son de la forma:

$$\hat{F}\phi_i = E'\phi_i \quad (3.11)$$

donde ϕ_i es un conjunto de funciones de onda de un electrón, llamado los Orbitales Moleculares de Hartree-Fock. Estos orbitales o funciones de onda en un principio tenían la forma de la Ec. 3.12 (tomada de las funciones para el átomo de hidrógeno):

$$s(r) = Ar^l e^{-\alpha|\vec{r}-\vec{a}|} \quad (3.12)$$

Estos orbitales son conocidos como Orbitales de Slater [31]. Sin embargo, con el tiempo se descubrió que estos orbitales pueden ser escritos como combinaciones lineales de funciones Gaussianas, que en un espacio cartesiano son de la forma:

$$g(x, y, z) = N_c x^l y^m z^n e^{-\alpha r^2} \quad (3.13)$$

donde l, m, n son enteros positivos; $r^2 = x^2 + y^2 + z^2$ y N_c es una constante de normalización que puede ser calculada. Estas funciones son llamadas Orbitales Gaussianos y son usadas debido a que facilitan los cálculos de las integrales involucradas en la Ec. 3.10 y por sus propiedades. Una de ellas, que es la más importante, viene dada por el Teorema del Producto de Gaussianas, que garantiza que el producto de dos orbitales del tipo gaussiano (GTO-Gaussian-Type-Orbitals) centrados en dos diferentes átomos es una suma infinita de las Gaussianas centradas en un punto en el eje que los conecta. De esta manera, las integrales de cuatro centros se pueden reducir a una suma finita de las integrales de dos centros y estas en el siguiente paso se reducen a una suma finita de las integrales de un centro. Con lo que se logra la aceleración del cálculo 4-5 veces en comparación con los orbitales de Slater, aunque se genera una cola extra en los orbitales generados durante el cálculo.

Existen cientos de conjuntos base compuestos de orbitales tipo Gaussiano. Los más cortos de estos tienen el nombre de conjunto base mínimo y típicamente se componen del número mínimo de funciones base requeridas para representar a todos los electrones en cada átomo. Sin embargo, los conjuntos más grandes pueden contener docenas y hasta cientos de funciones base para cada átomo.

La adición más común a los conjuntos mínimos de base son las funciones de polarización, que se denotan por el asterisco (*), o doble asterisco (**), para indicar que las funciones de polarización se agregan también a los átomos ligeros (hidrógeno y helio). Esta adición en las funciones provee una flexibilidad adicional necesaria dentro del conjunto base, que permite a los orbitales moleculares envolver al átomo de hidrógeno de la manera más asimétrica alrededor del núcleo. Este es un resultado importante al considerar la representación exacta de la formación de enlaces

entre los átomos, pues la sola presencia de los cambios de enlace en un átomo hace el entorno energético de los electrones esféricamente asimétrico. Existe también la adición de funciones tipo d que se agregan al conjunto de base con los orbitales de valencia p , y las funciones f que se agregan al conjunto con d-orbitales, etc. Existe una notación para indicar cuántas y cuáles de estas funciones son utilizadas como función base, pero de estas se hablará más adelante.

Otra adición usual al conjunto de funciones base es la adición de funciones difusas, denotadas por el signo más (+), o doble signo más (++) para indicar que estas funciones también se agregan a los átomos ligeros (hidrógeno y helio). Estas funciones Gaussianas representan con mayor precisión la porción de la «cola» de los orbitales atómicos que se encuentran distantes a los núcleos atómicos y pueden ser importantes cuando se consideran aniones y otros sistemas moleculares «suaves».

La convención común de nombres para los conjuntos de base mínimos es STO-XG, donde X es un entero y representa el número de Gaussianas primitivas que componen una función base. En estos conjuntos de base el mismo número de Gaussianas primitivas compone los orbitales cerrados de los electrones cubriendo el núcleo y los orbitales de valencia. Algunos ejemplos de estos conjuntos son: STO-2G, STO-3G, STO-6G y STO-3G* (versión polarizada de STO-3G).

Sin embargo, este tipo de conjuntos mínimos de base no proveen resultados con un buen grado de aproximación, por lo que es necesario hacer uso de los conjuntos base con valencia dividida, de los cuales se hablará a continuación.

Dado que en la mayoría de los enlaces moleculares principalmente los electrones de valencia son los que forman parte del enlace, es común representar a los orbitales de valencia con más de una función base. Esto se hace al escribir a cada una de ellas como una combinación lineal fija de Gaussianas primitivas. A este tipo de conjuntos se les llama Conjuntos de Valencia Dividida y se denotan de la siguiente forma: X-YZg. Donde X representa el número de las Gaussianas primitivas que componen cada función base del orbital atómico del núcleo, Y y Z indican que los orbitales de valencia se componen de dos funciones base cada una, la primera compuesta de una combinación lineal de Y-gaussianas primitivas y la otra de Z-gaussianas primitivas. En este caso, la presencia de dos números después del guión indica que este conjunto base es de *valencia dividida doble-zeta*. Los conjuntos base de valencia dividida triple y cuádruple-zeta se denotan como X-YZWg y X-YZWWg. Algunos ejemplos de bases doble-zeta son: 3-21g, 3-21g*, 3-21+g, 3-21+g* y 6-31g**.

Como el método HF no es tan exacto, en lo que respecta a los efectos de correlación, este puede ser complementado con otros métodos que ayudan a resolver

este problema. Uno de ellos es el método proveniente de la Teoría del Funcional de la Densidad (DFT), del cual hablaremos a continuación.

3.1.1. Teoría del Funcional de la Densidad (DFT)

Los métodos de primeros principios, a diferencia de Mecánica Molecular, no utilizan parámetros experimentales en el cálculo, están basados únicamente en las leyes de la Mecánica Cuántica (MC), teoría que hasta ahora describe mejor los fenómenos que ocurren a escalas atómicas y moleculares, y utiliza los valores de las constantes físicas fundamentales, por ejemplo, velocidad de la luz, masas y cargas de electrones y núcleos, constante de Plank.

Son evidentes las ventajas de estos métodos, ya que emplean una teoría más general y por lo tanto podemos obtener resultados inaccesibles para otros modelos empíricos y semi-empíricos. Sin embargo, la desventaja más clara es la gran cantidad de recursos informáticos que son necesarios para poder utilizar satisfactoriamente este modelo.

Es por ello que, para trabajar con modelos *ab initio* debemos seleccionar un sistema simple que nos proporcione la mayor información posible. En este trabajo, dicho sistema son los desoxidinucléosidos monofosfatos complementarios, cdDMPs, unidad mínima de la doble hélice del ADN.

Para poder realizar todos los cálculos que se requieren al hacer uso de este modelo, nos vemos en la necesidad de utilizar aproximaciones adicionales. La Teoría del Funcional de Densidad (DFT, por sus siglas en inglés) ha demostrado ser eficiente para algunos sistemas de interés práctico. Ésta se basa en los *dos teoremas de Hohenberg-Kohn* los cuales demuestran que las propiedades de los sistemas moleculares están determinadas por la densidad electrónica. Con la ventaja de que la densidad electrónica es una función de sólo tres variables. En particular, Hohenberg y Kohn mostraron que la energía puede expresarse como un funcional de la densidad y que además la densidad del sistema minimiza este funcional. El problema, es que no se conoce la forma completa de este funcional, por lo que debe ser aproximado. Es por ello que surgen distintos funcionales aplicables a distintos sistemas moleculares.

3.1.1.1. Primer Teorema de Hohenberg-Kohn

“Toda observable de un sistema mecánico cuántico estacionario es determinada por la densidad electrónica del estado base”

En otras palabras, se encuentra la energía del sistema como una función de la densidad, sin referencia a la función de onda. Como caso particular en un sistema

de n-electrones, la energía como observable será, según este primer teorema, función de la densidad electrónica. La ecuación de Schrödinger para tal sistema es:

$$\hat{H}\Psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots) = E\Psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots) \quad (3.14)$$

En el Hamiltoniano se incluyen: la energía cinética y la potencial. En la energía potencial son considerados dos términos: el correspondiente a la interacción núcleo-electrón y el correspondiente a la interacción electrón-electrón:

$$\hat{H} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee} \quad (3.15)$$

donde el operador de energía cinética es:

$$\hat{T} = \sum_{i=q}^N \frac{-\nabla_i^2}{2m_i} \quad (3.16)$$

La energía potencial de la interacción núcleo-electrón es:

$$\hat{V}_{ne} = \sum_i v\vec{r}_i = - \sum_{i=1}^N \sum_{\alpha} \frac{Z_{\alpha}}{r_{\alpha i}} \quad (3.17)$$

donde Z_{α} es el número de protones del núcleo α que interactúan con el electrón i-ésimo; y $r_{\alpha i}$ es la distancia de separación entre el núcleo α y el i-ésimo electrón.

Por otro lado, la energía potencial de interacción electrón-electrón es:

$$\hat{V}_{ee} = \sum_{i < j}^N \left(\frac{1}{r_{ij}} \right) \quad (3.18)$$

donde r_{ij} es la distancia de separación entre los electrones i-ésimo y j-ésimo.

Una vez que contamos con la solución a la ecuación de Schrödinger es posible hacer el siguiente mapeo uno a uno con la densidad electrónica $\rho(\mathbf{r})$:

$$\rho(\vec{r}_i) = \sum_{\sigma=\uparrow\downarrow} \int |\psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots)| dx_1 dx_2 \dots dx_N \quad (3.19)$$

donde la sumatoria se realiza tomando en cuenta a los electrones con spin hacia arriba y a los electrones con spin hacia abajo.

Cuando exponemos al sistema a un potencial externo es posible mostrar la relación que existe entre la densidad y la función de onda; y también la relación entre la función de onda y el potencial.

$$\hat{V} \leftrightarrow \psi \leftrightarrow \rho \quad (3.20)$$

Comprobemos esta relación, en primer lugar notemos que dado el potencial \hat{V} es posible, usando la ecuación de Schrödinger, encontrar la función de onda ψ ; así, con la función de onda y empleando la Ec. 3.19 podemos encontrar la densidad electrónica. Es decir:

$$\hat{V} \rightarrow \psi \rightarrow \rho \quad (3.21)$$

Ahora analicemos si se satisface $\hat{V} \leftarrow \psi$. Como ya hemos probado que $\hat{V} \rightarrow \psi$ sólo resta verificar que para dos potenciales, \hat{V} y \hat{V}' diferentes, se tendrán dos funciones de onda, ψ y ψ' , respectivamente. *"Si \hat{V} y \hat{V}' difieren por más de una constante C , no podrán conducir a la misma función de onda"*.

Para comprobar lo anterior supongamos que para un estado base no degenerado dos potenciales conducen a la misma función de onda:

$$(\hat{T} + \hat{V} + \hat{V}_{ee}) |\psi\rangle = E_{gs} |\psi\rangle \quad (3.22)$$

Y

$$(\hat{T} + \hat{V}' + \hat{V}_{ee}) |\psi\rangle = E'_{gs} |\psi\rangle \quad (3.23)$$

Restando las dos ecuaciones anteriores tenemos que:

$$(\hat{V} - \hat{V}') |\psi\rangle = (E_{gs} - E'_{gs}) |\psi\rangle \quad (3.24)$$

Lo cual implica que $V = V' + C$, entonces si dos potenciales conducen a la misma función de onda éstos sólo difieren por una constante.

Ya hemos probado que $\psi \rightarrow \rho$. Para probar que además $\psi \leftarrow \rho$ usaremos la siguiente afirmación: *"Si el estado base de ψ para H y ψ' para H' son diferentes, no podrán conducir a la misma densidad electrónica"*.

Para comprobar esto, apliquemos el *principio variacional* a los estados base de ψ y ψ' , cuyos hamiltonianos sólo difieren en la energía potencial. Y también para los cuales, en general, las densidades electrónicas son diferentes:

$$E_{gs} = \langle \psi | \hat{H} | \psi \rangle < \langle \psi' | \hat{H} | \psi' \rangle = \langle \psi' | \hat{H}' + \hat{V} - \hat{V}' | \psi' \rangle = E'_{gs} + \int \rho'(\vec{r}) [v(\vec{r}) - v'(\vec{r})] \quad (3.25)$$

$$E'_{gs} = \langle \psi' | \hat{H} | \psi' \rangle < \langle \psi | \hat{H}' | \psi \rangle = \langle \psi | \hat{H} + \hat{V}' - \hat{V} | \psi \rangle = E_{gs} + \int \rho(\vec{r}) [v'(\vec{r}) - v(\vec{r})] \quad (3.26)$$

Asumiendo que $\rho = \rho'$ y combinando estas dos ecuaciones llegamos a la contradicción:

$$E_{gs} + E'_{gs} < E_{gs} + E'_{gs} \quad (3.27)$$

Por lo tanto, con esto hemos probado que la densidad electrónica queda unívocamente determinada por la función de onda:

$$\psi \longleftrightarrow \rho(\vec{r}) \quad (3.28)$$

Entonces todas las observables del sistema son función de la densidad electrónica.

Ahora que sabemos que la densidad electrónica determina las propiedades del sistema, sólo resta saber cómo calcularla.

3.1.1.2. Segundo Teorema de Hohenberg-Kohn

"La densidad exacta para el estado base de un sistema con un potencial particular externo puede encontrarse por la minimización del funcional de energía"

Sin embargo el segundo teorema de Hohenberg-Kohn sólo nos afirma que es posible encontrar la densidad exacta pero no es posible aplicarlo directamente de forma práctica.

Para calcular la densidad es preferible usar las ecuaciones de Kohn-Sham. La ecuación de Schrödinger para un sistema de electrones que no interactúan ($\hat{V}_{ee} = 0$) y que se encuentran bajo un potencial externo (\hat{V}_{so}) es:

$$\left(-\frac{\nabla^2}{2} + \hat{V}_{so}(\vec{r}) \right) \phi_i(\vec{r}) = \epsilon_i \phi_i(\vec{r}) \quad (3.29)$$

donde la densidad del sistema se puede calcular por la sumatoria:

$$\rho_s(\vec{r}) = \sum_{i=1} |\phi_i(\vec{r})| \quad (3.30)$$

Y la energía cinética no interactiva del sistema es:

$$T_s[\rho] = \sum_{i=1}^N \left\langle \phi_i \left| -\frac{\nabla^2}{2} \right| \phi_i \right\rangle \quad (3.31)$$

Por otro lado, la energía total de interacción de un sistema puede escribirse como:

$$E[\rho] = T_s[\rho] + \int V_0(\vec{r})\rho(\vec{r})d\vec{r} + \frac{1}{2} \int \int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}d\vec{r}' + E_{XC}(\rho) \quad (3.32)$$

donde el funcional de energía de intercambio-correlación, $E_{XC}(\rho)$, se define como:

$$E_{XC}[\rho] = F_{HK}[\rho] - T_s\rho - \frac{1}{2} \int \int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}d\vec{r}' \quad (3.33)$$

Pero recordemos que el *primer Teorema de HK* asegura que el funcional de energía $E[\rho]$ es estacionario para pequeñas variaciones, $\delta\rho(r)$, alrededor de la densidad base exacta $\rho_0(r)$. Entonces:

$$\frac{\delta E[\rho]}{\delta \rho(\vec{r})} \Big|_{\rho=\rho_0} = 0 \quad (3.34)$$

Así, empleando las Ecuaciones 3.32, 3.33 y 3.34 tenemos:

$$\frac{\delta F_{HK}[\rho]}{\delta \rho(\vec{r})} \Big|_{\rho=\rho_0} = -V_0(\vec{r}) \quad (3.35)$$

y aplicando este resultado al sistema no interactivo de nuestro interés tenemos:

$$\frac{\delta F_{HK}[\rho]}{\delta \rho(\vec{r})} \Big|_{\rho=\rho_0} = -V_{s,o}(\vec{r}) \quad (3.36)$$

Al resolver la derivada parcial indicada obtenemos:

$$V_{s,o}(\vec{r}) = V_0(\vec{r}) + V_H(\vec{r}) + V_{X,C}(\vec{r}) \quad (3.37)$$

Y combinando esta última ecuación con la ecuación de Schrödinger, finalmente se concluyen las **ecuaciones de Kohn-Sham**:

$$\left[-\frac{\nabla^2}{2} + V_0(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d(\vec{r}') + \frac{\delta E_{XC}}{\delta \rho(\vec{r})} \right] \phi_i(\vec{r}) = \epsilon_i \phi_i(\vec{r}) \quad (3.38)$$

Donde el potencial de intercambio correlación está dado por:

$$V_{X,C}(\vec{r}) = \frac{\delta E_{XC}}{\delta \rho(\vec{r})} \quad (3.39)$$

Por lo tanto éste potencial de KS depende de la densidad. En resumen las ecuaciones de KS son:

$$V_{S,O}(\vec{r}) = V_0(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \frac{\delta E_{XC}}{\delta \rho(\vec{r})} \quad (3.40)$$

$$V_{X,C}(\vec{r}) = \frac{\delta E_{XC}}{\delta \rho(\vec{r})} \quad (3.41)$$

Y deben de ser resueltas mediante un esquema auto consistente. En la práctica se asigna un valor aproximado a E_{XC} para comenzar un proceso iterativo de correcciones hasta que se logren satisfacer ambas ecuaciones simultáneamente.

3.1.1.3. Intercambio y Correlación

Aún cuando el planteamiento de Kohn y Sham es exacto, hasta el momento el funcional de intercambio y correlación exacto, E_{XC} , es desconocido y por lo tanto son necesarias aproximaciones a este funcional. La clasificación a estas aproximaciones se puede encontrar en la escalera de Jacob definida por John. P. Perdew [32].

La primera aproximación para este funcional se conoce como Aproximación de Densidad Local (LDA) y consiste en suponer que en cada punto, la energía de intercambio y correlación depende sólo de la densidad en ese punto. Este valor se considera como el que tendría un gas de electrones libres de esa densidad. Si bien es una aproximación bastante fuerte, se obtienen resultados sorprendentemente precisos para algunas propiedades, y es en parte a eso que se debe el éxito de esta teoría.

Existen aproximaciones más sofisticadas para el funcional de intercambio y correlación, estas se conocen como Aproximaciones de Gradiente Generalizado, éstas son semilocales ya que consideran en cada punto el valor de la densidad y sus gradientes. Un ejemplo representativo de esta aproximación es el funcional reportado por Perdew, Burke y Ernzerhof, el cual ha motivado a varias revisiones y mejoras. Para algunas propiedades estas aproximaciones dan mejores resultados que LDA, en particular para geometrías moleculares y energías del estado fundamental, aunque para otras no representan una mejora sustancial.

Aún con esta aproximación, no se conoce la forma funcional para las energías de intercambio (también llamado interacción de canje) y correlación electrónicas.

Estas corresponden a la interacción cuántica entre electrones, la primera debido a la parte cuántica de la repulsión coulombiana y la segunda debido al principio de exclusión de Pauli entre electrones del mismo espín. Una solución a este problema es el desarrollo de métodos híbridos como el funcional B3LYP, pero existen otros métodos dentro de la propia Teoría del Funcional de la Densidad.

Una serie de funcionales más sofisticados puede obtenerse al suponer que la energía de intercambio y correlación depende explícitamente de los orbitales de Kohn-Sham. El más común de estos funcionales es el de Intercambio Exacto (EXX), que incluye de manera completa la energía de intercambio electrónico y que puede derivarse desde primeros principios. El problema de este tipo de funcionales es que computacionalmente son más costosos de tratar.

3.2. Mecánica Molecular

El Método de Mecánica Molecular (MM) se utiliza para hacer cálculos de las conformaciones de sistemas orgánicos, incluidos proteínas, ácidos nucleicos, etc. Este método considera los sistemas moleculares como un conjunto de átomos puntuales que interactúan por leyes de la Mecánica Clásica. El enfoque MM utiliza implícitamente la aproximación inicial para los esquemas de química clásica y cuántica, a saber, la aproximación de Born-Oppenheimer (BO), la cual permite la separación de los movimientos electrónico y nuclear. En este escenario se imagina que los electrones se adaptan instantáneamente a cada configuración nuclear y crean, junto con la propia interacción de los núcleos entre sí, una *superficie de energía potencial* (PES) que depende sólo de las posiciones nucleares y que actúa como una función de energía potencial nuclear efectiva. Así, la energía de la molécula es función sólo de las posiciones de los núcleos.

Este método se basa en la idea de que una molécula se puede representar por un conjunto de puntos situados en el núcleo de cada átomo, cada punto cuenta con una carga que depende del tipo de átomo, mientras que los enlaces de las moléculas son considerados como resortes. Debido a sus cargas, los átomos en su conjunto producirán una superficie de energía potencial. A diferencia de los métodos semi-empíricos (aquellos basados tanto en evidencia experimental como en Mecánica Cuántica) la MM es un método basado sólo en la Mecánica Clásica.

En los cálculos se considera que los átomos están dentro de un campo de potencial generado por ellos mismos y por medio de un algoritmo, usualmente el método de gradiente conjugado, se producen los movimientos de los átomos que les permiten llegar a una posición de mínima energía. En general, para una molécula se

pueden obtener muchos mínimos de diferente magnitud sobre la superficie de energía potencial. El más profundo es el *mínimo global* y los demás son *mínimos locales*.

Para realizar los cálculos de la energía potencial se emplea un conjunto de expresiones con las cuales se describe la superficie de energía potencial que se obtiene de forma empírica y cuya expresión matemática se toma de la Mecánica Clásica. A este conjunto de funciones de potencial se le llama *Campo de Fuerza* y está formado por parámetros ajustables que son optimizados para que los resultados de los cálculos realizados, con dichos campos, concuerden de la mejor manera posible con las propiedades experimentales de las moléculas, tales como su geometría. Dentro de este esquema el átomo es considerado como una esfera de radio r y con una carga efectiva que usualmente es calculada con métodos de la Mecánica Cuántica.

En general, la MM describe las moléculas en términos de “átomos enlazados” (átomos en molécula), sus posiciones son distorsionadas de alguna geometría idealizada (de equilibrio) debido a interacciones no enlazantes con los átomos de la misma molécula y de otras moléculas.

La primera aproximación principal del método MM es la aditividad para varios niveles de cálculos, es decir, la aditividad de los términos energéticos responsables de las contribuciones de diferente naturaleza física y la aditividad de las contribuciones de los átomos. La representación simple (“minimalista”) de la energía potencial de una molécula o complejo molecular en función de las coordenadas de átomos, R , puede representarse por una suma de cuatro términos principales, siendo cada término una suma de muchas contribuciones (Ec. 3.42).

$$\Delta E(R) = \sum Eb + \sum Ea + \sum Et + \sum Enb \quad (3.42)$$

donde el primer término se refiere a todos los enlaces químicos, el segundo a los ángulos de valencia, el tercero a los ángulos de torsión y el último a todos los pares de átomos no enlazados entre sí o a el tercer átomo en común. Los términos energéticos dependen de las posiciones mutuales de los átomos y de las constantes ajustables (parámetros); los parámetros son sugeridos para ser transferibles entre átomos y moléculas del mismo tipo. La transferibilidad de los parámetros del campo de fuerza es la segunda suposición importante de MM, y el número de tipos de átomos del mismo elemento químico depende del campo de fuerza particular.

3.2.1. Contribuciones Intramoleculares a la Energía del Sistema Molecular [33]

El primer y segundo término de la Ec. 3.42 se refieren a los cambios de energía debidos a las variaciones de las longitudes de los enlaces (el estiramiento, Ec. 3.43)

y de los ángulos de enlace (flexión, Ec. 3.44), respectivamente. Estos términos son usualmente modelados como potenciales armónicos centrados en valores de equilibrio de longitudes de enlace y ángulos de enlace, respectivamente, es decir, se usan dependencias de la “ley de Hooke simple”.

$$Eb_i = kb_i(l_i - l_i^0)^2 \quad (3.43)$$

$$Ea_i = ka_i(\alpha_i - \alpha_i^0)^2 \quad (3.44)$$

en estas ecuaciones, l_i y α_i son los valores actuales de longitudes de enlace y ángulos; los parámetros l_i^0 y α_i^0 son valores de equilibrio para longitudes de enlace y ángulos; kb_i y ka_i son constantes de fuerza de estiramiento y de flexión, respectivamente. Estos parámetros ajustables dependen de los tipos de átomos que forman el enlace o del ángulo de valencia. Algunos campos de fuerza también pueden contener contribuciones cúbicas y de orden superior a estos términos, a veces se puede utilizar un potencial Morse más flexible; adicionalmente, “términos cruzados” pueden ser incluidos para tener en cuenta las correlaciones entre los componentes de estiramiento y doblado. En este último caso, se añaden los términos que dependen tanto de la longitud de enlace como de las variaciones angulares.

La tercera suma de la expresión 3.42 se refiere a los cambios de energía de torsión; esta es responsable de las interacciones de los depósitos de electrones de dos átomos A y D (y de dos enlaces A-B y C-D) que están conectados a través de un enlace químico intermedio B-C.

$$Et_i = kt_i(1 + \cos(n_i\varphi_i - \delta_i)) \quad (3.45)$$

Esta función periódica (el número entero n_i es la periodicidad) contiene el valor actual φ_i del ángulo i de la rotación alrededor del enlace químico B-C (el ángulo entre los dos planos definidos por los átomos A, B y C y por B, C, y D) y tres parámetros (kt_i , n_i , y δ_i ; n_i es la multiplicidad y δ_i es el ángulo de fase) para cada tipo de torsión, para cada combinación de cuatro átomos vecinos de la molécula. Estos parámetros se pueden estimar a partir de datos experimentales sobre la estructura y propiedades de la molécula considerada y de moléculas relacionadas, entonces, dichos parámetro deben ajustarse por cálculos de prueba (así como l_i^0 , α_i^0 , kb_i y ka_i de las Ecs. 3.43, 3.44).

Muchos campos de fuerza incluyen términos responsables de torsiones “impropias” o curvaturas fuera del plano, es decir, términos relacionados con cuatro átomos que no forman enlaces químicos consecutivos, los cuales funcionan como factores de corrección para desviaciones fuera del plano (por ejemplo, pueden ser utilizados

para mantener planos los anillos aromáticos). Estos términos se pueden expresar a través de potenciales armónicos como los de estiramiento de enlace y flexión de ángulo de valencia. Los términos cruzados dependientes del ángulo de torsión y de la longitud del enlace o del ángulo de valencia se agregan en algunos campos de fuerza. Los tres términos mencionados anteriormente son positivos para prácticamente todos los estados de todos los sistemas moleculares (o igual a cero cuando el término corresponde al estado de equilibrio para contribución de enlace, ángulo o torsión).

3.2.2. Interacciones intermoleculares e intramoleculares no enlazantes [33]

La última suma de la Ec. 3.42 se refiere a las llamadas interacciones no enlazantes, E_{nb} , de todos los pares de átomos que no se enlazan entre sí o con el mismo tercer átomo, Ec. 3.46. Cada término átomo-átomo suele estar representado por una suma de interacciones electrostáticas, de Coulomb (el primer término de la Ec. 3.46) y de van der Waals (el segundo y tercer término de la Ec. 3.46).

$$E_{ij}(r_{ij}) = K \frac{q_i q_j}{r_{ij}} - \frac{A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} \quad (3.46)$$

donde r_{ij} es la distancia actual entre los átomos i y j ; q_i y q_j son cargas atómicas efectivas; A_{ij} y B_{ij} son parámetros ajustables responsables de la atracción de dispersión (London) y de las interacciones de repulsión de corto alcance, respectivamente. Las cargas atómicas se derivan usualmente usando cálculos a través de varios métodos de química cuántica; se puede usar una constante dieléctrica efectiva que implica implícitamente que rodea (este valor puede ser dependiente de la distancia). Los coeficientes A_{ij} y B_{ij} pueden ser estimados preliminarmente a través de la distancia interatómica de equilibrio y los valores de energía en equilibrio para pares neutros de átomos (ρ_{ij} y ϵ_{ij} , respectivamente), y seguido por el ajuste a los datos experimentales de referencia.

La mayor parte de los primeros campos de fuerza usaban para la descripción de las interacciones de van der Waals el potencial de Buckingham (6-exp) en vez del potencial de Lennard-Jones (6-12) en la Ec. 3.46. La expresión total para el término de interacción no enlazante se denomina usualmente como potencial (1-6-12) o (1-6-exp) relacionado con la dependencia de los términos en la distancia entre átomos. El potencial de Buckingham es más flexible (tiene tres parámetros ajustables en lugar de dos para los potenciales 6-12 para cada tipo de par de átomos) y tiene más justificaciones físicas para el rango de distancias observadas experimentalmente (debido a la dependencia exponencial de las funciones de ondas electrónicas en la distancia de los núcleos), pero es menos conveniente para los cálculos. Tiene un

máximo a corta distancia y luego tiende a un valor infinito negativo.

La mayoría de los campos de fuerza modernos utilizan expresiones 6-12 para la descripción de las interacciones de van der Waals, siendo el potencial total de átomo-átomo referido como 1-6-12. Algunos campos de fuerza sustituyen el potencial 6-12 por uno 10-12 para las interacciones de los átomos de hidrógeno de enlaces de hidrógeno con el fin de describir una dependencia en la distancia más aguda en el área más importante de mínimo de energía correspondiente a la formación de enlaces de H (referida como potencial 10-12). Se utilizaron expresiones más complejas (incluyendo aquellas que dependen de los ángulos entre dos líneas rectas que conectan tres átomos de enlace de H) para la descripción del enlace de H en algunos primeros conjuntos de potenciales. Los términos no enlazados de la energía intramolecular relacionados con las interacciones 1-4 (es decir, las interacciones entre átomos en una molécula separada por tres enlaces químicos) se consideran a menudo con un coeficiente menor que 1 (escala 1-4) ya que estas interacciones ya están incluidas en el término de torsión (Ec. 3.45). Para reducir el número de parámetros ajustables A_{ij} y B_{ij} del potencial de Lennard-Jones (y los correspondientes parámetros de otros potenciales), las reglas de combinación para la distancia interatómica de equilibrio, ρ , y los valores mínimos de energía, ϵ , para pares de diferentes átomos se aplican generalmente.

$$\rho_{ij} = \rho_i + \rho_j; \epsilon_{ij} = (\epsilon_i \epsilon_j)^{1/2} \quad (3.47)$$

Algunos campos de fuerza aplican las reglas de combinación directamente a los coeficientes de los términos de van der Waals.

Los cálculos de energía potencial a través de la Ec. 3.42 se utilizan para buscar mínimos *energéticos locales* (posiciones mutuales de átomos correspondientes a posibles configuraciones estables), construir y analizar superficies energéticas potenciales multidimensionales (PSE), seguir trayectorias de movimiento (en MD, simulaciones de dinámica molecular) o estudiar características promedio termodinámicas y geométricas (vía MC, muestreo de Monte Carlo) de los sistemas.

En la actualidad existen un gran número de campos de fuerza diferentes que se utilizan para el estudio de diferentes sistemas de moléculas. Cada uno contiene su conjunto de parámetros cuyos valores han sido elegidos bajo diferentes criterios y siguen sus propias reglas para realizar los cálculos de las energías. Estos campos de fuerza no siempre han sido los mismos, se han ido modificando a través del tiempo con el objetivo de obtener mejores resultados. Uno de ellos es el campo de fuerza AMBER, del cual hablaremos a continuación.

3.2.3. Campo de Fuerza AMBER (Assisted Model Building with Energy Refinement)

Existen diferentes programas que nos permiten hacer cálculos de energía con mecánica molecular. En particular vamos a describir brevemente el funcionamiento de *Assisted Model Building with Energy Refinement* (AMBER, por sus siglas en inglés). AMBER es el nombre de un paquete de simulación de dinámica molecular que implementa campos de fuerza también llamados AMBER. Un campo de fuerza es una expresión analítica dada para los términos de la energía potencial en el modelo de mecánica molecular, para este campo de fuerza dicha expresión está dada de la siguiente manera:

$$E_{Total} = \sum_{Enlace} K_R [R - R_0]^2 + \sum_{Angulos} K_\theta [\theta - \theta_0]^2 + \sum_{Diedros} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right] + \sum_{HB} \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}} \right] \quad (3.48)$$

El paquete AMBER es un conjunto de aplicaciones especializadas. Existen programas preparatorios (como *antechamber* y *xleap*), programas de cálculo de energía (como *sander*) y programas de análisis (como *anal* y *ptraj*). Cada programa juega un papel distinto en el flujo de la información.

Los programas preparatorios nos permiten asignar cargas y parámetros dependiendo del tipo de molécula y campo de fuerza con el que vayamos a trabajar. Además de permitirnos visualizar y editar el sistema bajo estudio. Los programas de cálculo de energía optimizan la geometría de la molécula. Y los programas de análisis nos permiten extraer información de los archivos de resultados.

El paquete AMBER además permite elegir entre distintos campos de fuerzas dependiendo del tipo de moléculas con las que estemos trabajando. Pues recordemos que cada campo de fuerza está ajustado a tipos de moléculas en específico. Entre estos campos existen los siguientes: ff99, ff95 y GLYCAM.

3.3. Metodología empleada para estudiar los dúplex de ADN diferentes del modelo de Watson-Crick

Como ya se ha mencionado, en este trabajo estamos interesados en estudiar estructuras de ADN diferentes del modelo de Watson-Crick. Para ver si es posible que dichas estructuras participen en los procesos de regulación de los ácidos nucleicos en los seres vivos (es decir, para averiguar la viabilidad de su estructura), se analizaron los parámetros conformacionales de fragmentos mínimos de dúplex de

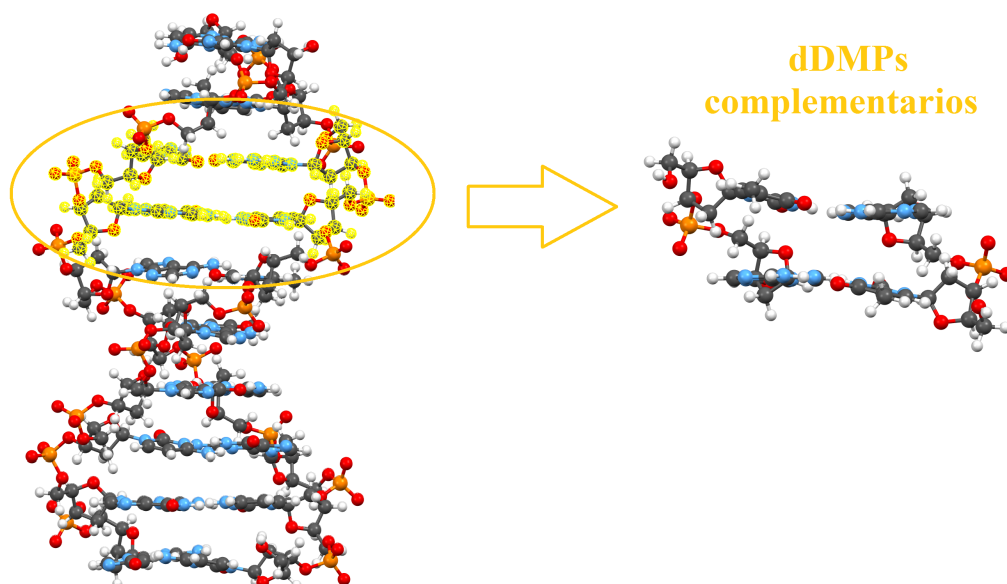


Figura 3.1: *Unidad mínima de la hélice doble del ADN. Desoxidinucleósidos monofosfatos complementarios.*

cadena paralelas, de dúplex con pares Ade(*syn*):Thy de Hoogsteen y con pares no correctos, así como de estructuras de la Familia Z. Para ello, se recurrió a las técnicas teórico-computacionales ya que permiten optimizar nuestros objetos de estudio, es decir, minimizan su energía y esto nos lleva a estructuras físicamente realizables (la naturaleza siempre busca estar en un estado de mínima energía).

Debido al costo computacional, nuestro estudio se basa en los dDMPs complementarios como unidad mínima para estudiar las características conformacionales de las hélices dobles diferentes del modelo de WC, pues, hemos comprobado anteriormente que los cdDMPs son las unidades mínimas de la doble hélice del ADN que nos pueden aportar información estructural de ésta [17]. Para ello, se realizó el corte de dichas estructuras de los dúplex tomados del banco de datos, como se muestra en la Fig. 3.1.

Sin embargo, los dDMPs no son estructuras eléctricamente neutras, por lo que, es necesario estabilizarlos agregando un ión de sodio (Na^+). Dicho ión se añade manteniendo el plano oxígeno-fósforo-oxígeno (O-P-O), del esqueleto azúcar-fosfato, equidistantemente a 2.3Å de ambos oxígenos. Una vez agregado el ión sodio, los dDMP- Na^+ complementarios (ver Fig. 2.8) pueden ser descritos totalmente por los ángulos de torsión del esqueleto azúcar-fosfato (δ , ϵ , ζ , α , β , γ), el empaquetamiento

de los dos anillos de la desoxirribosa, los dos ángulos de torsión glucosídicos (χ), uno por base, y el ángulo promedio formado por las bases nitrogenadas, además de los parámetros de posición mutua de las bases y pares de bases *Buckle*, *Propeller*, *Til*, *Roll*, *h-Twist* y *h-Rise*, y el área de traslape de los anillos de las bases apiladas debida al *Stacking* (ver Fig. 2.10).

Dicho esto, es indispensable mencionar la aproximación de la Mecánica Cuántica que fue empleada para lograr nuestros objetivos.

Para realizar la optimización de las estructuras con el método de Mecánica Cuántica se empleó la Teoría del Funcional de Densidad (DFT), ésta nos presenta un procedimiento variacional alternativo para obtener las propiedades del sistema en las que estamos interesados. Esta teoría trabaja con la densidad electrónica en lugar de con las tres coordenadas de posición de cada electrón. Sin embargo, no existe una expresión exacta del funcional que relaciona la densidad electrónica con la energía del sistema. En la actualidad disponemos de distintos funcionales aproximados, cada uno con sus limitaciones y ventajas. En este trabajo se utilizaron los funcionales *PBEPBE* [34] y *PW91PW91* [35], empleados anteriormente para estructuras de WC, con el fin de compararlos. Ambos métodos arrojaron resultados semejantes tal como se muestra en el Capítulo 4.

El funcional *PBEPBE* contiene *Intercambio* y *Correlación*, lo cual nos da una energía un tanto más cercana a la exacta, a diferencia del funcional *PW91PW91*; éste proporciona la respuesta lineal correcta de el gas uniforme de electrones con la escala uniforme apropiada, y, conduce a pseudopotenciales suaves. Para dichas optimizaciones se utilizó el conjunto base 6 – 31g*, donde cada uno de los orbitales internos son expresados con una *single-z* y esta a su vez es desarrollada por 6 gaussianas, los orbitales de valencia son expresados con *doble-z*, una desarrollada con 3 gaussianas y la otra con sólo una, además, dichas *zetas* están polarizadas.

PW91PW91 está entre los pocos funcionales que describen las interacciones de van der Waals en dímeros de gases raros e hidrocarburos, como también en los compuestos con enlaces de hidrógeno de moléculas simples [36]. Es por ello que dicho funcional es ideal para realizar nuestros estudios de comparación, ya que su alcance está dentro de nuestras necesidades. Además Tsuzuki y Luthi llegaron a la conclusión de que a pesar de las dificultades presentadas para la dispersión en el funcional *PW91*, es una alternativa viable a los costosos métodos *ab initio* [36]. Estudios más recientes con éste funcional demostraron que reproduce, razonablemente, buenos resultados *ab initio* para complejos con enlaces de hidrógeno [37] y para complejos con bases de ADN apiladas [38].

Para realizar los cálculos de minimización, con el funcional *PBE*, se dispuso del programa *Gaussian 09D.1* [39], el cual es considerado uno de los mejores paquetes comerciales para realizar cálculos post Hartree-Fock (HF). Dicho programa, puede realizar cálculos HF, Moller-Plesset (MP), CIS, DFT, ect. Es importante mencionar que, una vez optimizadas las estructuras de cdDMPs, se calcularon sus *frecuencias vibracionales* y se verificó que fuesen positivas, para corroborar que se tenían mínimos reales.

Las optimizaciones realizadas con el funcional *PW91* se llevaron a cabo en el programa ADF (Amsterdam Density Functional) [40]. El paquete ADF usa funciones de Slater y un esquema de integración numérica ajustable con suficiente exactitud, en el cual las coordenadas se adaptan automáticamente a las funciones-base disponibles, lo que reduce el tiempo de evaluación de los potenciales de Coulomb. Este programa cuenta con el conjunto base de valencia dividida (DZV, por sus siglas en inglés) que está especialmente realizado para cálculos DFT, en contraste con otros conjuntos base. Dicho conjunto base toma en consideración la región de valencia, importante químicamente, para la región de los orbitales externos. Para los cálculos realizados en éste programa se utilizó el conjunto base *TZP*, donde cada uno de los orbitales internos son expresados con una *doble-z* y los orbitales de valencia con *triple-z*, además con una función de polarización.

Finalmente, como cálculos preliminares (en algunos casos *Gaussian* no acepta las coordenadas obtenidas directamente de estructuras experimentales) y de comparación, se llevaron a cabo optimizaciones con Mecánica Molecular empleando el paquete comercial *AMBER 9* [41], el cual trabaja con el campo de fuerzas de igual nombre; cabe mencionar que, existen dos campos de fuerzas: AMBER y CHARMM, cada uno de los cuales contiene un conjunto de parámetros experimentales para cada término de la expresión matemática, son diferentes ya que los científicos consideraron algunos parámetros iguales pero otros diferentes, lo cual conduce a resultados distintos. Los varios campos de fuerza AMBER, comúnmente utilizados, son todos derivados del original campo de fuerza de Cornell et al. ff94 [42], éste fue derivado usando una combinación de datos experimentales y cálculos de Mecánica Cuántica. Los campos de fuerza ff98 [43] y ff99 [44] mejoraron el empaquetamiento del azúcar y las torsiones glucosídicas de los ácidos nucleicos en ff94. Posteriormente, en 2007, la corrección ff99bsc0 [45] introdujo una descripción mejorada de las torsiones alfa y gamma del esqueleto azúcar-fosfato. Finalmente, se han derivado dos conjuntos separados de parámetros utilizando la Mecánica Cuántica para describir las torsiones glucosídicas de las cuatro bases. Para nuestras optimizaciones empleamos específicamente el campo de fuerzas AMBER *ff99*.

3.3.1. Especificación de las estructuras empleadas para los cálculos

Como objeto principal de estudio se analizaron estructuras de cdDMPs, haciendo el corte de ellos en fragmentos experimentales de ADN con pares Ade(*syn*):Thy de Hoogsteen [21, 22], con pares no correctos (Gua:Thy, Gua:Ade(*syn*) e Ino:Ade(*syn*)) [25, 26, 27], y de dúplex de cadenas paralelas (obtenidas de resultados de Resonancia Nucleomagnética) [23], además de una estructura de la Familia Conformacional Z [20], todas ellas tomadas del *Nucleic Acid DataBase* ¹.

Para obtener dichos fragmentos mínimos, realizar las optimizaciones y medir sus parámetros estructurales se utilizaron los programas de modelación molecular *HyperChem*, *Avogadro*, *Oxford*, *RasMol*, *ChemCraft* y *Mercury*; es necesario un archivo con extensión *.pdb* que contenga el tipo de átomo, las coordenadas de cada uno, así como los enlaces que forman.

Cabe mencionar que, las estructuras experimentales no contienen átomos de hidrógeno, por lo que fue necesario agregarlos.

3.3.2. Medición de parámetros

Una vez que las estructuras fueron optimizadas, se obtuvo un archivo con extensión *.pdb* conteniendo las nuevas coordenadas de los átomos. Para obtener los parámetros antes mencionados (ángulos de torsión, empaquetamiento del azúcar, área de superposición de las bases debido al *Stacking*, *Buckle*, *Propeller*, *Tilt*, *Roll*, *h-Rise* y *h-Twist*), se empleó el programa en línea *3DNA* ².

Es importante mencionar que los archivos que lee dicho programa tienen un formato específico, por lo que fue necesario cambiar el formato de las estructuras optimizadas. Para ello se empleó el programa *Avogadro*, donde se eliminaron los iones de sodio (Na⁺) y todos los H's.

3.3.2.1. Programa en línea 3DNA

Completo paquete de software para el análisis, la reconstrucción y visualización de las estructuras de ácidos nucleicos tridimensionales. El programa, titulado *3DNA*, se puede aplicar a hélices paralelas y antiparalelas, estructuras de una sola cadena y hélices con multi-cadenas de ADN y ARN. Los análisis se pueden realizar en cualquiera de las estructuras cristalinas individuales, tales como las recopiladas en la base de datos de ácidos nucleicos (NDB), o conjuntos de estructuras generadas en las simulaciones moleculares. Las rutinas de análisis identifican y caracterizan

¹<http://ndbserver.rutgers.edu/>

²http://w3dna.rutgers.edu/index.php/analyze_file

todas las interacciones de las bases y proporcionan una clasificación automática apropiada de secuencias helicoidales dobles. Los patrones de enlaces de hidrógeno se describen en términos del desplazamiento espacial y la orientación de los marcos de referencia estándar en las bases interactuantes, por otro lado, las superposiciones se evalúan directamente de proyecciones planas de los anillos y los átomos de los radicales en bases o pares de bases consecutivos. La disposición espacial y la orientación relativa de secuencias consecutivas de pares de bases también se expresan en términos de parámetros estándar de cuerpo. Parámetros convencionales de torsión y una variedad de distancias virtuales y los ángulos son utilizados para caracterizar la conformación molecular, con las clasificaciones conformacionales automáticas basadas en parámetros derivados conocidos para las diferentes formas helicoidales [18].

Capítulo 4

Resultados

*Todos somos ignorantes, sólo que en
diferentes temas.*

WILL ROGERS

Como ya se ha mencionado, las estructuras de hélice doble del ADN pertenecientes al modelo de Watson-Crick no son las únicas posibles. Se tiene evidencia de estructuras conformadas por cadenas paralelas, además de estructuras con pares Adenina-Timina en su conformación de Hoogsteen; por otro lado, durante la replicación del ADN se puede tener la formación de pares no correctos o mutaciones puntuales, como lo son los pares de bases Guanina-Timina, Guanina-Adenina e Inosina-Adenina. Así mismo, se han encontrado estructuras de cadenas antiparalelas formando una hélice doble *zig-zagueada*, mejor conocidas como Familia Z. Es por ello que, en este trabajo de tesis se tiene como objetivo principal estudiar dichas conformaciones diferentes del modelo de WC por diferentes métodos computacionales, para analizar la viabilidad de su estructura.

Para este trabajo de tesis se minimizaron las estructuras iniciales mediante el método DFT con el funcional *PBE* [34] y la base $6-31g^*$, haciendo uso del programa *Gaussian 09D.1* [39], estas fueron construidas como se menciona en la Sección 3.3.1; en algunos casos se empleó el programa ADF [40] con el funcional *PW91* y la base *TZP*, con objeto de comparar los resultados obtenidos con ambos programas y funcionales. Como ya se ha mencionado antes, para corroborar que se tenía un mínimo real, se calcularon las *frecuencias vibracionales* de las estructuras optimizadas; de ser negativas, se sometían a optimización nuevamente; sólo se muestran los parámetros de las estructuras iniciales y de los mínimos reales. Además, se tienen resultados para más de una estructura «del mismo tipo» y con la misma secuencia de bases, no se muestran todos, sólo los más característicos.

Cabe mencionar que, para las estructuras que fue necesario, se realizaron cálculos de Mecánica Molecular con el campo de fuerzas AMBER *ff99* (empleando el programa de mismo nombre, *AMBER 9*), con objeto de obtener estructuras físicamente realizables, ya que, al tener parámetros experimentales, este tipo de optimización nos arroja buenos resultados y a un bajo costo computacional. Estos cálculos no suelen ser resultados finales, sólo son para comparación y/o como cálculos preliminares.

En las Tablas, los parámetros de las estructuras en **color azul** son correspondientes a las optimizaciones realizadas con el método DFT empleando el funcional *PBE* y la base $6-31g^*$; mientras que los parámetros en **color aqua** son de las estructuras optimizadas mediante cálculos de Mecánica Molecular con el campo de fuerzas AMBER. Se anexan las energías obtenidas para dichas estructuras optimizadas. Los parámetros que no tienen color corresponden a las estructuras iniciales tomadas del banco de datos.

4.1. cdDMPs de la Familia Estructural Z

Las conformaciones de ADN-Z de doble hélice levógira, con secuencia de bases alternando Pur-Pyr, han sido ampliamente estudiadas desde principios de los años setenta. Estas tienen un emparejamiento de bases de Watson-Crick, pero conformación *syn* del nucleósido de purina (figura 2.4), como se puede observar en los valores del ángulo χ en la Tabla 4.1. Estas conformaciones surgen bajo una concentración de sal alta o baja en solución. El primer fragmento macromolecular de ADN que fue obtenido experimentalmente a resolución atómica fue un hexámero alternando dG:dC, con conformación Z [46]. En este contexto, es importante estimar la contribución de sus subunidades en la formación de esta estructura. La pregunta es si las interacciones locales entre las subunidades son suficientes para crear un mínimo de energía local en dDMPs que predisponerían la conformación Z en el dúplex. En anteriores optimizaciones con DFT de dCpdG y dGpdC, como fragmentos mínimos de cadena única de ADN-Z, resultaron en mínimos locales que estaban lejos de la geometría observada experimentalmente para los dúplex de la Familia Estructural Z [8]. Así, se llegó a la conclusión de que en contraste con los dúplex de WC, las conformaciones de ADN-Z se originan no de la energía mínima de los fragmentos de una sola cadena, pero sí de otras interacciones intramoleculares o intermoleculares [3, 8].

Con objeto de estudiar la contribución de los fragmentos mínimos de la doble hélice levógira en la formación de su estructura, analizamos y sometimos a optimización cdDMPs de la estructura cristalina de ultra alta-resolución [20], la cual es una de las estructuras cristalinas con mayor resolución existentes en el banco

de datos [1]. Los ángulos de torsión del esqueleto azúcar-fosfato, el ángulo glucosídico, el empaquetamiento del azúcar, los parámetros de posición mutual de los pares de bases y el área de superposición de los anillos de las bases apiladas, de las estructuras iniciales y las optimizadas, se encuentran en las Tablas 4.1, 4.2 y 4.3.

Tabla 4.1: *Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8), para los mínimos de dDMP-Na+ complementarios correspondientes a la familia estructural Z.*

cdDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2
dCpdG	-154	148	262	72	69	190	181	92	56
dCpdG	-155	147	268	74	65	181	180	98	62
dCpdG	-157	153	264	80	62	187	180	86	67
dCpdG	-159	153	264	80	61	187	180	85	67
dCpdG	-162	131	275	68	69	190	177	94	59
dCpdG	-162	131	275	68	69	190	177	94	59
dGpdC	56	92	182	66	168	167	44	143	-144
dGpdC	67	93	249	283	217	228	56	147	-155
dGpdC	53	80	198	103	152	100	56	127	-130
dGpdC	79	83	209	284	222	260	44	78	-153
dGpdC	28	78	196	64	101	186	69	139	-116
dGpdC	87	74	195	285	232	265	32	88	-149

Ángulos en grados ($^\circ$).

Tal como habíamos mencionado antes, podemos corroborar con los valores para el ángulo glucosídico, χ , en la Tabla 4.1 que, los nucleósidos con purinas se encuentran en conformación *syn*. Nótese además que, los valores de los ángulos de torsión de las estructuras optimizadas son cercanos a los de las estructuras iniciales. Aunque, puede observarse en la Tabla 4.2 que el empaquetamiento del azúcar es variable para ambos cdDMPs optimizados. Cabe mencionar que el (dGpdC)₂ optimizado presentó una energía más favorable, comparado con las optimizaciones realizadas al (dCpdG)₂, para las optimizaciones con DFT y AMBER.

Por otro lado, podemos notar en la Tabla 4.3 que los parámetros de posición mutual de las bases y pares de bases cambió mucho en las estructuras después de las optimizaciones. El (dCpdG)₂ optimizado con DFT, después de someterse a cálculos con AMBER, presentó a sus pares de bases casi paralelos tal como la estructura inicial, Fig. 4.1; sin embargo, no sucedió lo mismo con el (dGpdC)₂ pues no presenta una buena posición mutual entre sus bases, como podemos observar en la Fig. 4.2.

Tabla 4.2: *Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios correspondientes a la familia estructural Z.*

cdDMPs	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
	Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
dCpdG	C2'-endo	159	39	C3'-endo	27	28
dCpdG	C2'-endo	162	35	C3'-endo	18	25
dCpdG	C3'-exo	181	35	C3'-endo	22	28
dCpdG	C3'-exo	182	35	C3'-endo	23	29
dCpdG	C1'-exo	137	43	C3'-endo	29	26
dCpdG	C1'-exo	137	43	C3'-endo	29	26
dGpdC	C3'-endo	27	28	C2'-endo	157	33
dGpdC	C3'-endo	34	30	C2'-endo	162	35
dGpdC	C3'-endo	14	41	C1'-exo	143	45
dGpdC	C3'-endo	19	39	C4'-exo	42	35
dGpdC	C3'-endo	18	40	C2'-endo	159	30
dGpdC	C3'-endo	30	45	C3'-endo	9	32

P y v_{max} en grados ($^{\circ}$).

Tabla 4.3: *Parámetros de las bases y pares de bases de los dDMPs complementarios de la familia estructural Z (véase designación en la fig. 2.10).*

cdDMPs	Traslape		$Buckle_{1,2}$		$Propeller_{1,2}$		$Tilt$	$Roll$
dCpdG	0	0	2.8	-6.2	-3.6	-2.2	-0.8	-3.0
dCpdG	0	0	2.8	-6.2	-3.6	-2.2	-0.8	-3.0
dCpdG	0	0.51	12.1	-12.5	-11.6	-10.6	0.4	-16.9
dCpdG	0	0	12.1	-12.5	-11.6	-10.6	0.4	-16.9
dCpdG	0	0.42	1.7	-1.7	-2.6	-2.6	0	-1.0
dCpdG	0	0	1.7	-1.7	-2.6	-2.6	0	-1.0
dGpdC	1.44	0	-6.2	1.8	-2.2	-1.8	0.9	-1.6
dGpdC	0.71	0	-6.2	1.8	-2.2	-1.8	0.9	-1.6
dGpdC	0	0	22.2	-31.8	5.8	3.9	1.9	-29.2
dGpdC	0	0	22.2	-31.8	5.8	3.9	1.9	-29.2
dGpdC	0.03	0	23.3	-12.7	0.3	15.2	6.3	-22.5
dGpdC	0	0	23.3	-12.7	0.3	15.2	6.3	-22.5

Área de traslape de las bases en Å^2 , $Buckle$, $Propeller$, $Tilt$ y $Roll$ en grados ($^{\circ}$).

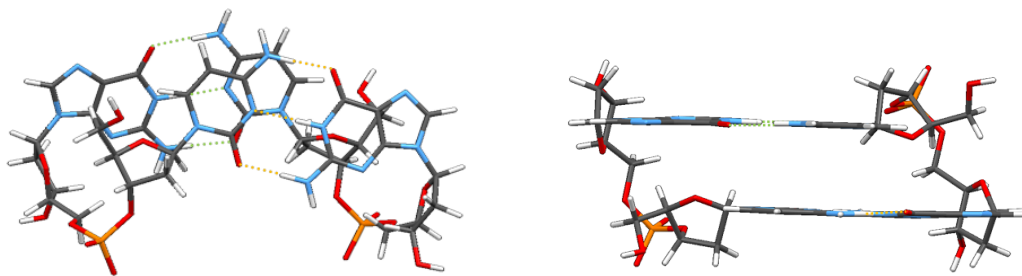


Figura 4.1: *Dúplex de $(dCpdG)_2$ de la Familia Conformacional Z, optimizados con PBE y posteriormente con AMBER. Vista superior y lateral, mostrando superposición de las bases apiladas y paralelismo entre los pares de bases respectivamente.*

Aunque las conformaciones optimizadas (con DFT y AMBER) de $(dCpdG)_2$ y $(dGpdC)_2$, extraídos de la estructura cristalina de ultra-alta resolución de la Familia Conformacional Z [20], retienen los enlaces de Hidrógeno en los pares de bases, y los valores de los ángulos de torsión son cercanos a los de las estructuras iniciales, las disposiciones mutuas de las bases están muy lejos de las del dúplex experimental. Estos resultados refuerzan la conclusión de que la formación de las conformaciones de la Familia Z se origina a partir de las interacciones del ADN con su entorno. Algunas de estas interacciones a través de puentes de agua se sugieren en la literatura [20]. Cabe recalcar que, el ADN-Z no existe en situación normal como el ADN-A y B (en solución), éste se presenta en condiciones no fisiológicas de mucha o poca sal.

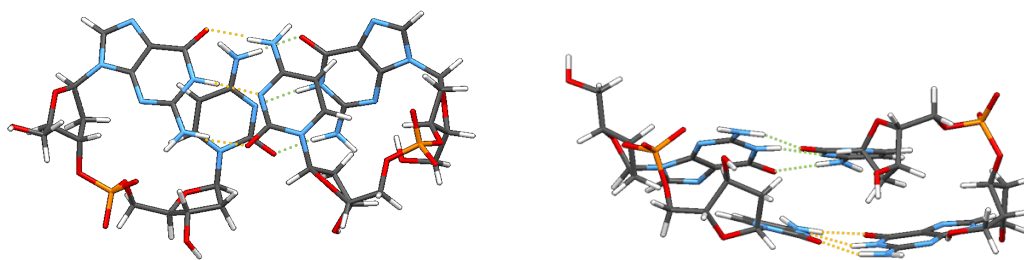


Figura 4.2: *Dúplex de $(dGpdC)_2$ de la Familia Conformacional Z, optimizados con PBE y posteriormente con AMBER. Vista superior y lateral, mostrando superposición de las bases apiladas y posición mutua entre los pares de bases respectivamente.*

4.2. Características conformacionales de dúplex con pares de bases Ade(*syn*):Thy de Hoogsteen

El dúplex de ADN en la forma de Hoogsteen se encontró, por primera vez, en la estructura cristalina del par Ade:Thy en 1959 [9]. Las bases de dicho par forman enlaces de Hidrógeno diferentes a los que se tienen en el par de Watson-Crick, ya que la base Adenina se encuentra en conformación *syn* con respecto al azúcar (figura 2.5). Anteriores optimizaciones de dDMPs correspondientes a dúplex de ADN con pares de Hoogsteen [28] mostraron estructuras con parámetros conformacionales cerca de los del fragmento experimental de dúplex de Hoogsteen con secuencia de bases alternada $d(ATATAT)_2$ [47]. Estos cálculos mostraron que el dúplex de Hoogsteen no reproduce la dependencia de secuencia que se encuentra para los dúplex de WC. Las secuencias de dApdT y dTpdA optimizadas conservaron una superposición sustancial de las bases apiladas, la cual está presente en la estructura cristalina. Tal superposición fue predicha teóricamente en el cálculo de dApdA a partir de la estructura de dApdT Hoogsteen con la base Timina sustituida por una Adenina en la orientación *syn* con respecto al azúcar [3]. Basándose en dichos cálculos, se predijo que todos los dApdT y dTpdA de dúplex con pares de Hoogsteen tendrían notable superposición de bases apiladas, nótese que los dúplex de WC correspondientes a las secuencias de Pyr-Pur muestran una superposición insignificante de los anillos de las bases. Como consecuencia de la amplia superposición de las bases, los dúplex de Hoogsteen, en contraste con los de WC, no serían capaces de doblarse.

En este trabajo, se realizaron optimizaciones con DFT de cdDMPs a partir de las conformaciones de varios fragmentos dúplex mínimos extraídos de estructuras experimentales tomadas del banco de datos, dichas estructuras tienen las secuencias $d(ATATAT)_2$ [21] y $d(ATT AAT)_2$ [22]. Estos cálculos produjeron mínimos de energía que exhibieron las principales características conformacionales ya predichas en los cálculos anteriores de dDMPs.

Se han obtenido dos o más *mínimos energéticos locales* para cada cdDMPs al iniciar la optimización directamente del fragmento correspondiente de la estructura experimental o después de la optimización preliminar realizada con cálculos de la Mecánica Molecular empleando el software *AMBER 9*. Los parámetros de todas las estructuras iniciales y las optimizadas, con las diferentes secuencias tomadas de los dos fragmentos de dúplex experimentales, se encuentran en las Tablas 4.4, 4.5, 4.6, 4.7, 4.8 y 4.9.

Podemos observar en la Tabla 4.4 que algunos de los ángulos de torsión del esqueleto azúcar-fosfato, para los cdDMPs dApdA:dTpdT, se encuentran fuera de las regiones en que se encuentran estos para estructuras de WC ¹, algunos ángulos de torsión de las dos cadenas de la conformación optimizada dApdA:dTpdT difieren más que las correspondientes torsiones de las estructuras iniciales; sin embargo, nótese que se tienen buenas posiciones mutuales de las bases y de los pares de bases, Tabla 4.6. Además, a pesar de tener diferentes cdDMPs iniciales, los mínimos obtenidos con DFT tienen una mínima diferencia de 2-3 *Kcal/mol*, siendo más favorable la energía de la estructura que presenta dos de sus ángulos fuera de las regiones ya mencionadas, comparando con la estructura que no presenta ningún ángulo fuera de región.

Tabla 4.4: *Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8), para los mínimos de dDMP-Na+ complementarios con pares de Hoogsteen [22].*

cdDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2	E
Pirimidina-Pirimidina/Purina-Purina										
dTpdT	-145	95	187	224	36	107	12	114	-118	
dApdA	72	107	162	278	319	171	55	122	47	
dTpdT	-121	139	198	175	29	78	49	106	-130	-519.57
dApdA	57	153	182	279	293	170	47	117	54	
dTpdT	-122	145	192	188	22	94	49	115	-122	-3029747.23
dApdA	67	151	183	272	298	173	52	128	77	
dTpdT	-112	128	196	232	288	158	72	127	-127	
dApdA	69	132	163	274	292	199	51	134	66	
dTpdT	-125	138	192	272	286	173	51	78	-123	-3029745.04
dApdA	71	149	182	275	295	175	51	119	75	
dTpdT	-129	131	181	263	286	173	52	87	-136	-527.42
dApdA	50	152	184	282	292	170	46	116	54	
dTpdT	-146	126	198	225	41	102	357	115	-107	
dApdA	96	130	161	273	300	183	59	115	54	
dTpdT	-154	106	196	198	34	82	46	109	-125	-511.71
dApdA	57	153	182	279	293	171	47	117	55	
dTpdT	-162	94	196	220	27	98	43	121	-116	-3029742.09
dApdA	68	152	184	274	295	174	51	126	75	

Ángulos en grados (°).

Energía en *Kcal/mol*.

¹Se puede ver en el *Apéndice A* las regiones de los ángulos para cada familia estructural.

Tabla 4.5: *Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios con pares de Hoogsteen [22].*

cdDMPs	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
	Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
Pirimidina-Pirimidina/Purina-Purina						
dTpdT	O4'-endo	95	40	C1'-exo	121	38
dApdA	C1'-exo	114	39	C1'-exo	131	44
dTpdT	C2'-endo	147	43	C1'-exo	112	43
dApdA	C2'-endo	172	40	C1'-exo	123	45
dTpdT	C2'-endo	164	37	C1'-exo	123	43
dApdA	C3'-endo	180	35	C1'-exo	137	42
dTpdT	C1'-exo	135	41	C3'-exo	137	36
dApdA	C1'-exo	144	42	C1'-exo	144	41
dTpdT	C2'-endo	167	29	C4'-exo	58	37
dApdA	C3'-exo	183	35	C1'-exo	132	41
dTpdT	C2'-endo	153	29	C3'-endo	29	32
dApdA	C2'-endo	169	41	C1'-exo	121	46
dTpdT	C1'-exo	136	37	C1'-exo	123	37
dApdA	C2'-endo	146	40	C1'-exo	120	40
dTpdT	O4'-endo	105	43	C1'-exo	115	43
dApdA	C2'-endo	172	40	C1'-exo	123	45
dTpdT	O4'-endo	90	41	C1'-exo	129	43
dApdA	C2'-endo	179	36	C1'-exo	134	43

P y v_{max} en grados ($^{\circ}$).

Por otro lado, en la estructura cristalina con las secuencias $d(ATATAT)_2$ [21], las dos cadenas de las conformaciones iniciales de los cdDMPs $(dApdT)_2$ y $(dTpdA)_2$ tienen diferentes ángulos de torsión. Después de la optimización, los parámetros conformacionales de ambas cadenas son iguales, excepto para una única estructura, Tabla 4.7. Nótese que una de las estructuras de $(dTpdA)_2$ tiene sus ángulos α y γ fuera de las regiones en que se encuentran estos para las estructuras del modelo de WC, aún así muestra una buena posición mutua de sus bases (Tabla 4.9); sin embargo, su energía es 2-3 $Kcal/mol$ menos favorable comparada con la energía de la estructura cuyos ángulos de torsión se encuentran dentro de las regiones esperadas.

Es importante notar que en éstas estructuras se obtuvo poca superposición de las bases apiladas para $Pyr-Pyr$ y considerable para el caso de $Pur-Pur$, Fig. 4.3 (ARRIBA), tal como se tiene en estructuras del modelo de WC. Por otro lado, se obtuvo algo de superposición para $Pur-Pyr$, mientras que para $Pyr-Pur$ se presentó una superposición notoria, ver Fig. 4.3 (CENTRO y ABAJO), esto último distinto a lo que se tiene para conformaciones del modelo de WC.

Tabla 4.6: *Parámetros de las bases y pares de bases de los dDMPs complementarios con pares de Hoogsteen [22] (véase designación en la fig. 2.10).*

cdDMPs	Traslape	<i>Buckle</i> _{1,2}	<i>Propeller</i> _{1,2}	<i>Tilt</i>	<i>Roll</i>		
Pirimidina-Pirimidina/Purina-Purina							
dTpdT dApdA	0.02 4.08	-5.6	-10.3	-8.5	-2.8	1.7	-1.2
dTpdT dApdA	0 4.09	4.6	1.6	-4.6	-22.4	-0.7	-9.6
dTpdT dApdA	0.03 3.74	14.0	7.3	-13.8	13.9	-10.8	15.8
dTpdT dApdA	0 3.48	6.4	0.2	9.4	10.4	1.7	0.1
dTpdT dApdA	0.23 3.34	-19.2	-7.0	0.8	-12.6	-0.9	-4.0
dTpdT dApdA	0 4.12	2.0	5.0	-25.0	-4.3	-1.0	-10.1
dTpdT dApdA	0.12 3.99	-1.6	-23.2	-15.9	-12.9	3.9	2.2
dTpdT dApdA	0 4.09	0.6	4.3	-4.6	-24.2	2.5	-10.1
dTpdT dApdA	0.09 3.44	-19.2	-5.1	0.8	-15.9	-0.3	-6.5

Área de traslape de las bases en Å², *Buckle*, *Propeller*, *Tilt* y *Roll* en grados (°).

Tabla 4.7: *Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8), para los mínimos de dDMP-Na+ complementarios con pares de Hoogsteen [21].*

cdDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2	E
Purina-Pirimidina										
dApdT	65	140	171	269	295	189	51	135	-108	
dApdT	99	134	192	233	311	149	68	130	-120	
dApdT	51	144	180	278	296	171	55	120	-150	-543.77
dApdT	51	144	180	278	296	171	55	120	-150	
dApdT	75	151	164	267	303	182	61	135	-126	-3029744.54
dApdT	75	151	165	267	303	182	61	135	-127	
dApdT	83	143	191	261	311	164	50	139	-104	
dApdT	65	145	184	263	311	176	41	131	-109	
dApdT	57	152	180	276	295	173	56	121	-150	-527.52
dApdT	57	152	180	276	295	173	56	121	-150	
dApdT	73	153	168	265	303	181	61	134	-128	-3029748.43
dApdT	73	153	168	265	303	181	61	134	-128	
dApdT**	77	153	166	268	302	185	61	132	-129	
dApdT	77	153	166	268	302	185	61	131	-129	
dApdT	66	143	180	265	302	183	42	134	-99	
dApdT	87	149	174	274	292	187	51	135	-93	
dApdT	51	144	180	278	296	171	56	120	-151	-535.65
dApdT	56	152	180	276	295	173	56	121	-150	
dApdT	57	147	180	270	297	179	58	129	-124	-3029756.98
dApdT	71	153	173	266	299	180	60	131	-125	
dApdT	52	150	182	278	294	172	55	119	-151	-530.82
dApdT	57	152	180	276	295	173	56	121	-150	
dApdT	76	151	164	267	304	182	61	136	-126	-3029746.49
dApdT	74	153	167	264	304	181	61	136	-127	
dApdT**	61	147	180	272	296	183	58	126	-126	
dApdT	76	153	173	269	298	184	59	127	-127	
Pirimidina-Purina										
dTpdA	-108	135	195	237	316	164	41	132	78	
dTpdA	-97	137	199	241	305	157	46	134	99	
dTpdA	-122	141	201	279	286	170	48	109	78	-3029745.72
dTpdA	-122	141	201	279	286	170	48	109	78	
dTpdA	-117	135	187	278	295	165	47	120	70	-518.27
dTpdA	-117	135	187	278	295	165	47	120	70	
dTpdA	-123	140	200	278	288	171	47	124	80	-3029747.24
dTpdA	-123	140	200	278	288	171	47	124	80	
dTpdA	-99	134	201	253	35	207	284	144	73	
dTpdA	-108	138	197	252	22	204	296	149	87	
dTpdA	-132	145	192	233	66	161	284	156	83	-3029743.77
dTpdA	-132	145	192	233	66	161	284	156	83	
dTpdA	-150	135	188	242	63	166	289	153	57	-514.50
dTpdA	-150	135	188	242	63	166	289	153	57	

Ángulos en grados ($^{\circ}$).

Energía en *Kcal/mol*.

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

Tabla 4.8: *Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios con pares de Hoogsteen [21].*

cdDMPs	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
	Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
Purina-Pirimidina						
dApdT	C2'-endo	159	39	C2'-endo	149	41
dApdT	C2'-endo	146	41	C1'-exo	143	39
dApdT	C2'-endo	161	38	C1'-exo	125	43
dApdT	C2'-endo	161	38	C1'-exo	125	43
dApdT	C3'-exo	188	37	C2'-endo	150	37
dApdT	C3'-exo	188	37	C2'-endo	149	37
dApdT	C2'-endo	176	29	C2'-endo	152	38
dApdT	C2'-endo	165	37	C2'-endo	150	35
dApdT	C2'-endo	169	40	C1'-exo	125	43
dApdT	C2'-endo	169	40	C1'-exo	125	43
dApdT	C3'-exo	183	37	C2'-endo	148	37
dApdT	C3'-exo	183	37	C2'-endo	148	37
dApdT**	C3'-exo	186	36	C2'-endo	147	35
dApdT	C3'-exo	187	36	C2'-endo	146	35
dApdT	C2'-endo	164	37	C2'-endo	154	36
dApdT	C3'-exo	183	32	C2'-endo	151	35
dApdT	C2'-endo	161	38	C1'-exo	125	43
dApdT	C2'-endo	169	40	C1'-exo	125	43
dApdT	C2'-endo	170	36	C1'-exo	139	39
dApdT	C2'-endo	180	37	C1'-exo	143	38
dApdT	C2'-endo	166	40	C1'-exo	123	44
dApdT	C2'-endo	169	40	C1'-exo	126	43
dApdT	C3'-exo	189	37	C2'-endo	151	37
dApdT	C3'-exo	183	37	C2'-endo	151	37
dApdT**	C2'-endo	173	37	C1'-exo	137	35
dApdT	C3'-exo	185	36	C1'-exo	140	36
Pirimidina-Purina						
dTpdA	C2'-endo	149	41	C2'-endo	147	42
dTpdA	C2'-endo	150	35	C2'-endo	147	41
dTpdA	C2'-endo	169	31	C1'-exo	119	42
dTpdA	C2'-endo	169	31	C1'-exo	119	42
dTpdA	C2'-endo	165	29	C1'-exo	124	45
dTpdA	C2'-endo	165	29	C1'-exo	124	45
dTpdA	C2'-endo	169	31	C1'-exo	130	43
dTpdA	C2'-endo	169	31	C1'-exo	130	43
dTpdA	C2'-endo	154	36	C2'-endo	175	32
dTpdA	C2'-endo	160	34	C3'-exo	183	32
dTpdA	C2'-endo	159	40	C3'-exo	205	37
dTpdA	C2'-endo	159	40	C3'-exo	205	37
dTpdA	C2'-endo	169	40	C1'-exo	125	43
dTpdA	C2'-endo	169	40	C1'-exo	125	43

P y v_{max} en grados ($^{\circ}$).

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

Tabla 4.9: *Parámetros de las bases y pares de bases de los dDMPs complementarios con pares de Hoogsteen [21] (véase designación en la fig. 2.10).*

cdDMPs	Traslape	$Buckle_{1,2}$	$Propeller_{1,2}$		
Purina-Pirimidina					
dApdT	0.86				
dApdT	1.22	-9.9	-30.1	26.9	3.3
dApdT	0.02	19.8	-19.8	-3.5	3.5
dApdT	0.02				
dApdT	0.83	2.5	-2.8	-3.6	3.3
dApdT	0.87				
dApdT	2.02	21.8	6.3	-2.4	-19.6
dApdT	1.55				
dApdT	0.03	22.2	-22.2	-4.1	4.1
dApdT	0.03				
dApdT	0.74	2.2	-2.3	-2.8	2.6
dApdT	0.74				
dApdT**	0.65	6.0	-6.2	-4.6	4.4
dApdT	0.65				
dApdT	0.97	1.1	-15.4	18.6	-8.7
dApdT	1.44				
dApdT	0.03	20.7	-21.3	-3.6	4.1
dApdT	0.03				
dApdT	0.71	1.3	-3.7	2.8	-1.5
dApdT	0.81				
dApdT	0.02	21.6	-22.4	-3.8	4.8
dApdT	0.02				
dApdT	0.81	2.6	-2.6	-3.1	3.0
dApdT	0.86				
dApdT**	0.58	5.0	-8.7	2.1	-0.5
dApdT	0.65				
Pirimidina-Purina					
dTpdA	1.59				
dTpdA	2.25	-30.1	7.9	3.3	14.2
dTpdA	2.32	-13.4	13.3	6.7	-6.7
dTpdA	2.33				
dTpdA	3.84	-7.6	7.6	-0.3	0.3
dTpdA	3.84				
dTpdA	2.23	-12.7	12.9	6.1	-6.2
dTpdA	2.23				
dTpdA	3.45	-15.4	10.6	-8.7	0.6
dTpdA	2.49				
dTpdA	2.11	-11.5	11.6	9.8	-9.8
dTpdA	2.11				
dTpdA	0.03	22.3	-22.3	-4.1	4.1
dTpdA	0.03				

Área de traslape de las bases en Å², *Buckle* y *Propeller* en grados (°).

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

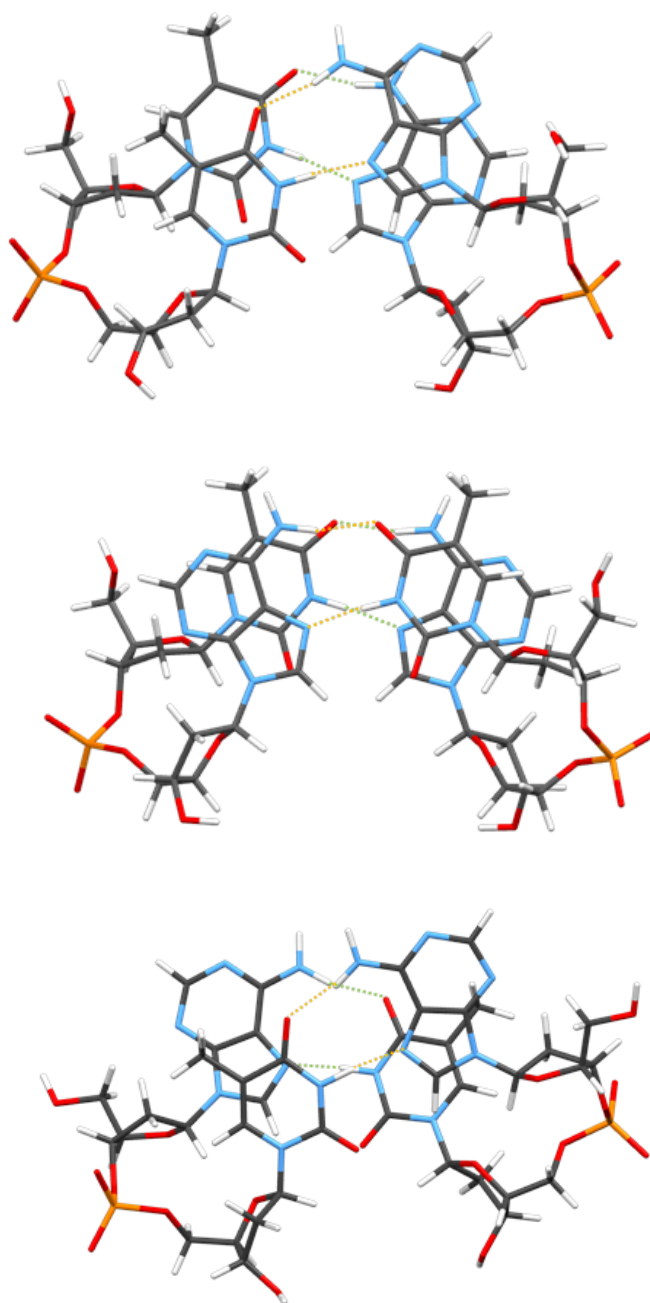


Figura 4.3: Conformaciones de dúplex con pares de Hoogsteen Ade(*syn*):Thy mostrando la superposición de sus bases apiladas. (ARRIBA) *dTpdT:dApdA*. (CENTRO) $(dTpdA)_2$. (ABAJO) $(dApdT)_2$.

4.3. Conformaciones de cadenas paralelas con el par reverso de Watson-Crick Ade:Thy

La formación de fragmentos de ADN de cadenas paralelas puede que desempeñe algún rol durante los procesos genéticos como la recombinación genética (apareamiento de cromosomas homólogos durante la meiosis), el plegamiento de los cromosomas y la regulación de la replicación y la transcripción. La existencia de un par de bases «Watson-Crick reverso» Ade:Thy (figura 2.6) sugiere la posibilidad de formación de dúplex con cadenas paralelas. Parvathy *et al* diseñaron un dodecámero de ADN con cadenas paralelas con secuencia intercalada de las bases Adenina y Timina [23]. Las estructuras tridimensionales de dúplex de cadenas paralelas fueron obtenidas usando simulaciones de dinámica molecular bajo restricciones de Resonancia Nucleomagnética (NMR), existe un total de 20 modelos [23]. Dichos dúplex contienen sólo dos tipos de fragmentos mínimos, dApdA:dTpdT y dApdT:dTpdA.

Para este trabajo hemos construido varias conformaciones iniciales de tales cdDMPs utilizando fragmentos de los modelos 1, 7, 10 y 16 con la orientación de los pares de bases menos distorsionada y más paralela. Los ángulos de torsión del esqueleto azúcar-fosfato, el empaquetamiento del anillo de azúcar, los ángulos de torsión glucosídicos (χ) y las posiciones mutuales de las bases de las estructuras iniciales y las optimizadas del dúplex con dApdT:dTpdA se encuentran en las Tablas 4.10, 4.11 y 4.12; mientras que los parámetros para los dúplex dApdA:dTpdT se muestran en las Tablas 4.13, 4.14 y 4.15.

La optimización con DFT de estos cdDMPs reveló regularidades de las posiciones mutuales de las bases en dúplex paralelos con acoplamiento Watson-Crick reverso de las bases. Podemos notar de la Tabla 4.10 que, para el dúplex con dApdT:dTpdA, para distintos modelos tenemos mínimos con energía prácticamente igual pero con pequeñas diferencias entre los valores de sus parámetros, esto nos indica que tenemos mínimos locales con barrera entre ellos. Cabe mencionar que, a pesar de que estos cdDMPs tienen una energía de 2 Kcal/mol más favorable presentan un ángulo entre sus pares de bases un poco más pronunciado, comparando con la estructura de energía más grande.

Tabla 4.10: *Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8), para los mínimos de dApdT:dTpdA-Na+ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy.*

cdDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2	E
Modelo 10										
dTpdA	-110	121	182	257	297	172	61	122	-105	
dApdT	-99	121	180	270	299	168	50	126	-103	
dTpdA	-128	136	200	280	283	179	46	132	-107	-3029743.96
dApdT	-123	141	191	270	302	170	59	136	-107	
dTpdA	-128	138	202	284	281	179	45	131	-109	-3029745.02
dApdT	-120	143	184	246	316	162	62	145	-112	
Modelo 7										
dTpdA	-116	126	188	268	297	174	56	123	-94	
dApdT	-101	145	178	261	296	172	65	126	-107	
dTpdA	-128	138	202	284	280	179	45	131	-108	-3029745.23
dApdT	-120	143	184	244	318	159	62	139	-115	

Ángulos en grados (°).

Energía en Kcal/mol.

Tabla 4.11: *Empaquetamiento del anillo de azúcar para los mínimos de dApdT:dTpdA-Na+ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy.*

cdDMPs	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
	Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
Modelo 10						
dTpdA	C1'-exo	126	36	C1'-exo	133	36
dApdT	C1'-exo	128	43	C1'-exo	137	53
dTpdA	C2'-endo	161	28	C1'-exo	136	43
dApdT	C3'-exo	182	28	C1'-exo	142	38
dTpdA	C2'-endo	160	30	C1'-exo	135	30
dApdT	C3'-exo	183	30	C2'-endo	160	30
Modelo 7						
dTpdA	C1'-exo	141	24	C1'-exo	134	38
dApdT	C2'-endo	174	33	C1'-exo	134	37
dTpdA	C2'-endo	161	30	C1'-exo	135	43
dApdT	C3'-exo	182	30	C2'-endo	156	35

P y v_{max} en grados (°).

Tabla 4.12: *Parámetros de las bases y pares de bases de los dApdT:dTpdA-Na+ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy* (véase designación en la fig. 2.10).

cdDMPs	Traslape	<i>Buckle</i> _{1,2}		<i>Propeller</i> _{1,2}	
Modelo 10					
dTpdA	0	-26.0	-28.0	-0.9	7.8
dApdT	3.16				
dTpdA	0.04	5.0	14.6	13.7	-17.5
dApdT	0.36				
dTpdA	0.2	-2.5	-1.5	11.7	-16.4
dApdT	0				
Modelo 7					
dTpdA	0	-14.1	18.9	2.2	15.0
dApdT	1.29				
dTpdA	0.22	1.7	2.1	17.2	-12.1
dApdT	0				

Área de traslape de las bases en Å², *Buckle* y *Propeller* en grados (°).

En las Tablas 4.13 y 4.14 se puede ver que para los dúplex con dApdA:dTpdT, de los distintos modelos, se tienen resultados semejantes (con ligeras desviaciones) para todas las optimizaciones realizadas con los diferentes métodos y funcionales; además de que, las energías de las estructuras optimizadas con DFT son prácticamente iguales para todos los modelos, excepto para el modelo 7 que presenta una energía 3 *Kcal/mol* más favorable. Cabe mencionar que, a pesar de que el dúplex mínimo tomado del modelo 7 presenta buenos parámetros conformaciones no fue posible iniciar cálculos con Mecánica Cuántica directamente del fragmento de banco de datos, por lo que se realizaron cálculos preliminares con Mec. Molecular utilizando los campos de fuerza AMBER y CHARM, posteriormente se sometieron las estructuras a cálculos con PBE. Dichas optimizaciones presentaron buena posición mutual de sus bases, sin embargo, se obtuvieron dos ángulos de torsión, α y β , fuera de las regiones en que se encuentran estos para estructuras de WC ².

Por otro lado, también se puede notar en la Tabla 4.13 que de los modelos 1, 10 y 16 se obtuvieron estructuras cuyos ángulos se encuentran todos dentro de las regiones esperadas para estructuras de WC, lo cual nos muestra la viabilidad de dichas estructuras estructuras de cadenas paralelas, es decir, que son físicamente realizables.

²Se puede ver en el *Apéndice A* las regiones de los ángulos para cada familia estructural.

Tabla 4.13: *Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8), para los mínimos de dApdA:dTpdT-Na⁺ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy.*

cdDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2	E
Modelo 7										
dTpdT	-107	126	182	270	295	171	55	123	-102	
dApdA	-94	123	177	248	293	182	59	141	-96	
dTpdT*	-119	136	190	273	307	166	44	129	-112	
dApdA	-113	118	182	241	307	169	50	138	-119	
dTpdT	-124	140	201	281	291	171	52	126	-102	-3029745.48
dApdA	-106	142	197	174	15	96	47	125	-119	
dTpdT	-122	138	187	278	296	165	54	128	-113	
dApdA	-83	130	192	223	296	167	52	143	-120	
dTpdT	-124	141	198	280	292	171	52	125	-103	-3029745.52
dApdA	-106	142	196	176	18	95	46	125	-117	
dTpdT	-121	135	186	275	297	164	56	125	-117	
dApdA	-91	137	198	165	26	74	49	109	-126	
dTpdT	-126	139	203	283	290	173	52	127	-104	-3029743.8
dApdA	-120	148	195	187	21	96	48	123	-121	
Modelo 10										
dTpdT	-108	127	181	271	295	172	54	125	-102	
dApdA	-95	124	177	249	294	182	59	139	-98	
dTpdT	-125	140	199	281	293	173	52	132	-101	-3029742.17
dApdA	-121	145	187	243	305	179	56	149	-107	
Modelo 16										
dTpdT	-108	127	181	271	295	172	54	125	-102	
dApdA	-95	123	177	249	293	182	58	139	-97	
dTpdT**	-119	138	201	277	291	172	53	104	-112	
dApdA	-120	111	181	267	293	193	51	142	-102	
dTpdT	-125	140	199	281	292	173	52	132	-101	-3029742.16
dApdA	-122	145	187	243	305	179	57	149	-101	
Modelo 1										
dTpdT	-108	126	182	270	295	172	54	124	-101	
dApdA	-94	124	177	249	294	181	59	140	-97	
dTpdT	-122	138	187	278	296	165	54	128	-114	
dApdA	-83	130	192	223	296	167	52	143	-120	
dTpdT**	-120	138	200	277	291	170	52	99	-118	
dApdA	-118	105	181	272	295	190	52	142	-101	
dTpdT	-122	137	187	277	296	165	55	127	-113	
dApdA	-116	91	183	272	298	173	57	141	-108	
dTpdT	-121	141	200	281	287	167	45	82	-118	-3029742.61
dApdA	-121	145	189	242	303	180	55	150	-103	
dTpdT	-122	138	187	278	296	165	54	128	-113	
dApdA	-83	130	192	224	296	167	52	143	-120	

Ángulos en grados (°).

Energía en Kcal/mol.

* Estructura optimizada con el campo de fuerzas CHARMM.

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

Tabla 4.14: *Empaquetamiento del anillo de azúcar para los mínimos de dApdA:dTpdT-Na+ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy.*

cdDMPs	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
	Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
Modelo 7						
dTpdT	C1'-exo	134	37	C1'-exo	129	41
dApdA	C1'-exo	134	38	C2'-endo	169	33
dTpdT*	C2'-endo	159	32	C1'-exo	135	43
dApdA	C1'-exo	121	52	C2'-endo	164	31
dTpdT	C2'-endo	164	31	C1'-exo	133	42
dApdA	C2'-endo	149	40	C1'-exo	134	42
dTpdT	C2'-endo	160	32	C1'-exo	131	40
dApdA	C1'-exo	135	45	C2'-endo	173	28
dTpdT	C2'-endo	165	31	C1'-exo	132	42
dApdA	C2'-endo	149	40	C1'-exo	133	42
dTpdT	C2'-endo	161	29	C1'-exo	128	40
dApdA	C2'-endo	144	45	C1'-exo	113	40
dTpdT	C2'-endo	162	31	C1'-exo	135	42
dApdA	C2'-endo	175	36	C1'-exo	131	43
Modelo 10						
dTpdT	C1'-exo	135	38	C1'-exo	132	42
dApdA	C1'-exo	135	38	C2'-endo	167	32
dTpdT	C2'-endo	164	32	C1'-exo	136	42
dApdA	C2'-endo	177	33	C2'-endo	175	33
Modelo 16						
dTpdT	C1'-exo	136	38	C1'-exo	132	42
dApdA	C1'-exo	134	38	C2'-endo	167	32
dTpdT**	C2'-endo	161	31	O4'-endo	100	40
dApdA	C1'-exo	113	41	C2'-endo	161	33
dTpdT	C2'-endo	163	32	C1'-exo	136	42
dApdA	C2'-endo	177	33	C2'-endo	175	33
Modelo 1						
dTpdT	C1'-exo	134	38	C1'-exo	130	41
dApdA	C1'-exo	135	37	C2'-endo	169	33
dTpdT	C2'-endo	160	32	C1'-exo	131	40
dApdA	C1'-exo	135	45	C2'-endo	173	28
dTpdT**	C2'-endo	161	31	O4'-endo	94	41
dApdA	O4'-endo	104	40	C2'-endo	159	34
dTpdT	C2'-endo	161	31	C1'-exo	131	40
dApdA	O4'-endo	82	39	C2'-endo	161	31
dTpdT	C2'-endo	167	31	C4'-exo	60	41
dApdA	C2'-endo	175	33	C2'-endo	178	33
dTpdT	C2'-endo	160	32	C1'-exo	131	40
dApdA	C1'-exo	135	45	C2'-endo	173	28

P y v_{max} en grados ($^{\circ}$).

* Estructura optimizada con el campo de fuerzas CHARMM.

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

Tabla 4.15: *Parámetros de las bases y pares de bases de los dApdA:dTpdT-Na+ de cadenas paralelas con el par Watson-Crick reverso Ade:Thy* (véase designación en la fig. 2.10).

cdDMPs	Traslape	<i>Buckle</i> _{1,2}		<i>Propeller</i> _{1,2}	
Modelo 7					
dTpdT	1.67	-18.9	-11.1	-15.0	-13.9
dApdA	0.24				
dTpdT*	0.65	17.0	39.0	-13.4	8.9
dApdA	0.54				
dTpdT	1.94	8.4	5.6	-3.1	-1.2
dApdA	0.58				
dTpdT	0.93	8.7	48.3	-11.4	20.1
dApdA	0.04				
dTpdT	1.73	11.4	8.4	-3.3	-1.3
dApdA	0.60				
dTpdT	0.81	17.7	24.5	-8.3	0
dApdA	1.15				
dTpdT	1.95	15.3	0.5	-8.7	-0.4
dApdA	0.39				
Modelo 10					
dTpdT	1.79	-20.0	-12.1	-11.9	-9.2
dApdA	0.17				
dTpdT	2.02	10.1	6.3	-4.2	-0.2
dApdA	0.09				
Modelo 16					
dTpdT	1.73	-20.1	-11.5	-12.3	-10.2
dApdA	0.20				
dTpdT**	1.24	-20.7	-0.2	5.6	8.9
dApdA	0.99				
dTpdT	2.03	-6.1	-9.9	0	4.3
dApdA	0.09				
Modelo 1					
dTpdT	1.65	-18.5	-11.9	-16.1	-14.7
dApdA	0.22				
dTpdT	0.93	8.7	48.6	-11.5	20.3
dApdA	0.03				
dTpdT**	1.38	1.2	22.4	-9.6	-7.8
dApdA	1.32				
dTpdT	0.59	16.1	45.9	-1.3	16.6
dApdA	0.68				
dTpdT	1.81	14.0	7.5	2.3	-0.6
dApdA	0.09				
dTpdT	0.91	-48.5	-9.0	-20.0	11.7
dApdA	0.04				

Área de traslape de las bases en Å², *Buckle* y *Propeller* en grados (°).

* Estructura optimizada con el campo de fuerzas CHARMM.

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

En los modelos optimizados con DFT, las secuencias dApdT:dTp dA se caracterizan por una superposición sustancial de los anillos de las bases en la cadena dApdT y por una superposición insignificante en la cadena dTp dA (como en los dúplex de WC). La optimización de la geometría de las secuencias dAp dA:dTp dT conduce a una superposición sustancial de los anillos de las bases en la cadena dTp dT y a un solapamiento menor en la cadena dAp dA (en contraste con los WCDs), véase las Tablas 4.12 y 4.15. Basándonos en estos datos computacionales, concluimos que los cdDMPs como fragmentos mínimos de los dúplex de cadenas paralelas tienen mínimos de energía bien definidos, y sus patrones conformacionales difieren de los de los WCDs (ver Fig. 4.4).

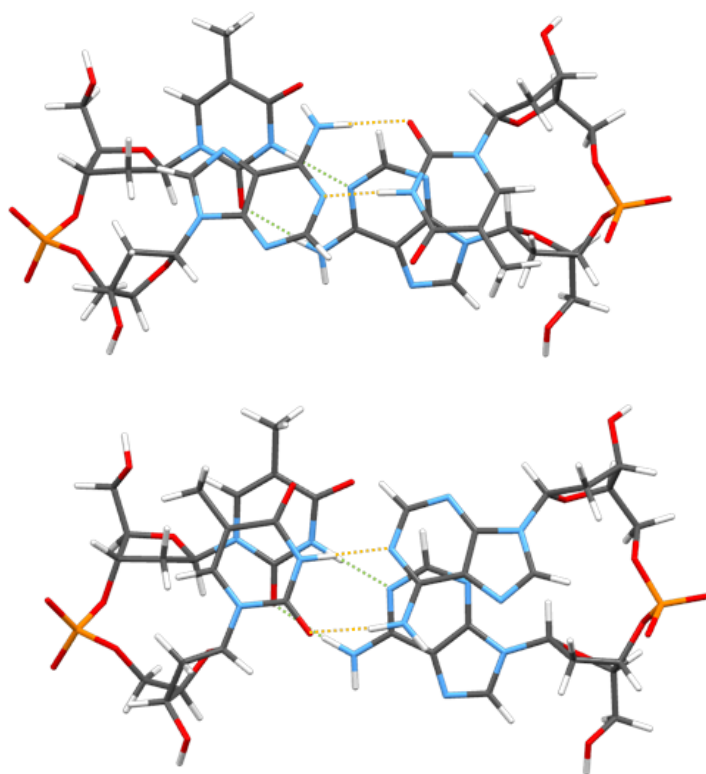


Figura 4.4: Conformaciones de dúplex de cadenas paralelas con el par reverso de Watson-Crick Ade:Thy mostrando la superposición de sus bases apiladas. (ARRIBA) dTp dA:dAp dT. (ABAJO) dTp dT:dAp dA.

4.4. Dúplex de Watson-Crick con pares no correctos

La formación de pares de bases distintos de Ade:Thy y Gua:Cyt en los procesos de biosíntesis de ácidos nucleicos fue sugerida por Watson y Crick en su trabajo pionero [48]. La evolución biológica y, por ende, la vida misma, no es posible sin la presencia de mutaciones puntuales o pares no correctos. Dada la importancia de la formación de estos pares, han sido estudiados cristales de varios dúplex con pares de bases «erróneos» usando difracción de rayos X. El Nucleic Acid Data Bank [1] contiene modelos de nivel atómico de varios dúplex de este tipo (por ejemplo, [25, 26, 27]).

En este trabajo consideramos dúplex de ADN con tres pares erróneos: con Gua:Thy [25], Gua:Ade(*syn*) [26], y el par de Ade(*syn*) con el análogo de guanosina modificada, inosina (Ino), que carece de grupo amino [27] (Ino:Ade(*syn*)). Los parámetros conformacionales de las estructuras iniciales y optimizadas con el par Gua:Thy se encuentran en las Tablas 4.16, 4.17 y 4.18; para el par Gua:Ade(*syn*) se tienen sus parámetros en las Tablas 4.19, 4.20 y 4.21; finalmente, en las Tablas 4.22, 4.23 y 4.24 se presentan los parámetros estructurales para las conformaciones con el par erróneo Ino:Ade(*syn*).

Las regiones de torsión de estos dúplex que contienen pares erróneos son en su mayoría similares a las de los dúplex de WC, pero algunos de los ángulos de torsión parecen estar fuera de las regiones estándar, ver Tablas 4.16, 4.19 y 4.22. Un análisis exhaustivo de las características conformacionales de las estructuras reveló una desviación de las regularidades en los arreglos de las bases vecinas de aquellas que son características para los WCDs, es decir, la superposición sustancial de los anillos de las bases para las secuencias Pyr-Pyr y Pyr-Pur.

Para los dúplex con el par no correcto Gua:Thy, el dCpdG:dTpdG presenta algunos de sus ángulos, ver Tabla 4.16, fuera de las regiones de torsión de los WCDs, el dDMP que presenta estas características formó enlaces de H entre OH de 5'-fin con el O4' y sus bases no son totalmente paralelas entre sí, ya que, además, el grupo *amino* de la Citosina de éste quiere formar enlace de H con el grupo *oxo* de la Timina del otro dDMP. Por otro lado, el dTpdT:dGpdA presenta una gran superposición entre los anillos de las bases apiladas del dTpdT (igual que en el caso del dTpdT:dApdA con cadenas paralelas) y menor solapamiento entre las bases del dGpdA, comparado con el resultado experimental. Para este último cdDMPs se obtuvieron buenos resultados para los cálculos con AMBER y DFT, ya que se tiene una buena posición mutua de sus bases y pares de bases, ver Tabla 4.18. Podemos notar que estas estructuras presentan los parámetros *h-Twist* y *h-Rise* semejantes a los que tenemos en los WCDs; en los otros casos no presentamos estos resultados ya que al tener cadenas paralelas o las bases en conformación *syn*, el programa *3DNA* no nos arroja resultados correctos.

Tabla 4.16: *Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8), para los mínimos de dDMP-Na+ complementarios con el par erróneo Gua:Thy [25].*

cdDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2
dCpdG:dTpdG									
dCpdG	-147	76	203	254	328	135	64	114	-115
dTpdG	-114	82	196	275	305	182	50	150	-81
dCpdG	-164	85	257	216	22	88	43	120	-109
dTpdG	-128	144	174	273	277	234	49	148	-70
dTpdT:dGpdA									
dTpdT	-115	115	186	276	288	173	58	85	-127
dGpdA	-99	142	151	282	281	220	56	132	-95
dTpdT	-124	143	194	283	287	170	53	81	-125
dGpdA	-144	145	169	277	283	235	48	148	-83
dTpdT	-126	139	186	280	295	166	54	120	-128
dGpdA	-169	106	191	275	293	186	48	143	-114

Ángulos en grados (°).

Tabla 4.17: *Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios con el par erróneo Gua:Thy [25].*

cdDMPs	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
	Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
dCpdG:dTpdG						
dCpdG	C3'-endo	35	39	C1'-exo	131	43
dTpdG	C1'-exo	115	30	C2'-endo	170	39
dCpdG	C4'-exo	36	37	C1'-exo	133	41
dTpdG	C2'-endo	168	36	C3'-exo	213	34
dTpdT:dGpdA						
dTpdT	C1'-exo	140	34	C4'-exo	69	32
dGpdA	C2'-endo	178	35	C3'-exo	184	30
dTpdT	C2'-endo	164	35	O4'-endo	73	39
dGpdA	C2'-endo	164	39	C3'-exo	208	34
dTpdT	C2'-endo	157	35	C1'-exo	122	42
dGpdA	O4'-endo	108	41	C2'-endo	164	34

P y v_{max} en grados (°).

Tabla 4.18: *Parámetros de las bases y pares de bases de los dDMP-Na+ complementarios con el par erróneo Gua:Thy [25] (véase designación en la fig. 2.10).*

cdDMPs	Traslape	<i>h-Twist</i>	<i>h-Rise</i>	<i>Buckle</i> _{1,2}		<i>Propeller</i> _{1,2}		<i>Tilt</i>	<i>Roll</i>
dCpdG:dTpdG									
dCpdG	0.07	20.7	2.7	-14.0	13.4	-6.9	-8.3	3.8	6.4
dTpdG	0								
dCpdG	0	41.6	3.6	32.2	-13.7	-33.7	-15.7	2.2	20.5
dTpdG	0								
dTpdT:dGpdA									
dTpdT	1.65	47.7	3.5	-5.7	-9.9	-17.0	-10.2	0.4	-2.7
dGpdA	2.62								
dTpdT	3.52	48.3	3.8	1.8	9.1	-11.7	-0.3	1.7	-5.4
dGpdA	0.78								
dTpdT	2.67	46.6	3.5	3.5	4.4	-12.3	-4.5	-1.3	-1.8
dGpdA	1.51								

Área de traslape de las bases en Å², *h-Rise* en Å, *h-Twist*, *Buckle*, *Propeller*, *Tilt* y *Roll* en grados (°).

Tal como se mencionó antes, podemos corroborar con los valores de los ángulos glucosídicos que, las bases Adeninas de los pares erróneos se encuentran en conformación *syn*, ver Tablas 4.19 y 4.22.

En el caso de los dúplex con el par Gua:Ade(*syn*) tenemos buenos resultados para ambos cdDMPs tomados del banco de datos, aunque uno de ellos presenta algunos de sus ángulos de torsión fuera de las regiones esperadas, ver Tabla 4.19, la posición mutual de sus bases es buena, se tienen pares de bases casi paralelos, ver Tabla 4.21. Después de optimizar con DFT al dTpdA:dGpdA se obtuvo una gran superposición entre las bases apiladas del dTpdA (con Adenina en conformación *syn*) y poca para las bases del dGpdA, regularidad que no se tiene para los WCDs. Por otro lado, se puede notar que después de optimizar con DFT al dApdG:dCpdG, por distintas rutas, se obtuvieron resultados semejantes, sus ángulos de torsión, el empaquetamiento del azúcar (ver Tabla 4.20) y los parámetros de posición mutual de sus bases varían poco, estas optimizaciones presentan una diferencia de 1 *Kcal/mol*; dicho dúplex muestra una superposición de sus bases apiladas semejante a la que se tiene en los WCDs.

Cabe mencionar que, después de los cálculos con DFT, los dos dúplex con el par Gua:Ade(*syn*) presentan el grupo Amino de la base Adenina(*syn*) fuera de plano, este quiere formar enlace de H con la base Purina apilada en cada caso.

Tabla 4.19: Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8), para los mínimos de dDMP-Na+ complementarios con el par erróneo Gua:Ade(syn) [26].

cdDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2	E
dTpdA:dGpdA										
dTpdA	-114	93	151	311	264	196	60	109	67	
dGpdA	-105	114	134	310	271	203	67	122	-111	
dTpdA	-123	141	201	278	285	177	45	115	74	-3085178.83
dGpdA	-142	145	164	278	281	241	49	148	-83	
dApdG:dCpdG										
dApdG	67	109	179	282	292	186	50	146	-82	
dCpdG	-143	70	194	277	293	187	65	114	-105	
dApdG	56	147	188	275	296	175	52	130	-103	-3095254.88
dCpdG	-165	91	211	269	297	184	54	146	-91	
dApdG	52	143	183	279	294	169	48	135	-109	
dCpdG	-131	84	180	273	299	177	57	144	-107	
dApdG	56	147	187	275	296	173	52	135	-105	-3095255.64
dCpdG	-161	89	208	267	299	176	57	146	-98	

Ángulos en grados ($^{\circ}$).

Energía en Kcal/mol.

Tabla 4.20: Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios con el par erróneo Gua:Ade(syn) [26].

cdDMPs	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
	Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
dTpdA:dGpdA						
dTpdA	C1'-exo	133	20	C1'-exo	124	42
dGpdA	C4'-endo	237	14	C2'-endo	154	36
dTpdA	C2'-endo	167	31	C1'-exo	126	41
dGpdA	C2'-endo	167	38	C3'-exo	202	34
dApdG:dCpdG						
dApdG	C1'-exo	124	42	C2'-endo	179	37
dCpdG	C4'-exo	47	41	C4'-endo	237	14
dApdG	C2'-endo	169	35	C2'-endo	146	40
dCpdG	O4'-endo	78	36	C3'-exo	193	34
dApdG	C2'-endo	162	37	C1'-exo	142	42
dCpdG	C4'-exo	71	41	C2'-endo	172	30
dApdG	C2'-endo	169	35	C2'-endo	145	41
dCpdG	O4'-endo	76	37	C2'-endo	176	33

P y v_{max} en grados ($^{\circ}$).

Tabla 4.21: *Parámetros de las bases y pares de bases de los dDMP-Na+ complementarios con el par erróneo Gua:Ade(syn)[26]* (véase designación en la fig. 2.10).

cdDMPs	Traslape	<i>Buckle</i> _{1,2}	<i>Propeller</i> _{1,2}	<i>Tilt</i>	<i>Roll</i>		
dTpdA:dGpdA							
dTpdA	0.75	-2.2	4.2	-12.8	13.1	-174.7	17.3
dGpdA	0.95						
dTpdA	3.8	1.5	12.4	-3.6	-21.6	-175.5	7.8
dGpdA	0.35						
dApdG:dCpdG							
dApdG	2.36	4.2	8.8	13.1	-8.5	134.7	-107.8
dCpdG	0						
dApdG	1.3	-7.8	-9.2	19.9	-15.9	-137.6	115.7
dCpdG	0						
dApdG	1.54	-10.3	-15.7	18.5	-19.8	138.8	-104.2
dCpdG	0						
dApdG	1.24	-7.5	-9.1	19.6	-16.2	-139.0	114.0
dCpdG	0						

Área de traslape de las bases en Å², *Buckle*, *Propeller*, *Tilt* y *Roll* en grados (°).

Finalmente, para el estudio de los dúplex con el par erróneo Ino:Ade(*syn*) sometimos a cálculos y analizamos 3 dúplex de cdDMPs con dicho par, dos de ellos con las mismas secuencias. El dApdG:dCpdI después de ser optimizado (tanto con el programa ADF y funcional PW91, como con el programa *Gaussian 09D.1* y el funcional PBE) muestra pares de bases no totalmente paralelos entre sí, aunque la estructura optimizada con *Gaussian 09D.1* tiene una menor inclinación entre sus bases apiladas, este resultado tal vez se deba a la formación de un enlace de H entre OH en 5'-fin con el O4' del dApdG; los ángulos para ambas optimizaciones se encuentran dentro de las regiones de torsión de los WCDs (ver Tablas 4.22 y 4.23), sin embargo, presentan una superposición nula entre las bases del dApdG, ver Tabla 4.24, distinto a lo que obtuvimos para el dúplex semejante con el par Gua:Ade(*syn*), ver Tabla 4.21.

Para la secuencia dTpdA:dIpdA analizamos dos dúplex tomados del banco de datos. El primer dúplex presenta todos sus ángulos en las regiones para los WCDs, pares de bases casi planos entre sí y una superposición de sus bases apiladas semejante a lo obtenido para la estructura con el par Gua:Ade(*syn*) con secuencia semejante. Por otro lado, el segundo dúplex con dicha secuencia, fue optimizado con el programa *Gaussian 09D.1* empleando los funcionales PW91 y PBE, en ambos casos presenta varios de sus ángulos de torsión fuera de las regiones de los WCDs, esto tal vez debido a que, en principio, las estructuras presentan dichos ángulos muy lejanos a las regiones y es difícil hacer un salto tan grande de una región a otra.

Las posiciones mutuales de sus bases son buenas, aunque ambos cálculos arrojaron una mayor superposición entre las bases del dIpdA, comparada con la que se tuvo en el primer dúplex; la superposición entre las bases del dTpdA es sustancial en todos los casos. Los empaquetamientos del azúcar de ambos dúplex, para todas las optimizaciones, es semejante. Cabe mencionar que la optimización del primer dúplex nos dió una energía de 4 *Kcal/mol* más favorable.

Las optimizaciones con DFT de cdDMPs con pares erróneos demuestran que las estructuras optimizadas tienen cambios aún más pronunciados en las regularidades de superposición de bases apiladas que aquellos que presenta la estructura experimental, aunque las torsiones se acercaron a las de los WCDs. Las conformaciones optimizadas con DFT de cdDMPs con los pares Gua:Thy y Gua:Ade(*syn*) se muestran en la Figura 4.5 y la Figura 4.6 respectivamente. La superposición sustancial de dos bases Thy en fragmentos mínimos con el par no correcto Gua:Thy (Fig. 4.5) difieren drásticamente de la superposición menor de estas bases que tienen lugar en WCDs sin pares erróneos (tanto en estructuras experimentales, como en fragmentos mínimos de ADN optimizados con DFT), se obtuvo también esta característica para las estructuras de cadenas paralelas (Fig. 4.4). El mismo resultado se aplica a la secuencia Thy-Ade (Fig. 4.6) en fragmentos mínimos con pares no correctos Pur:Pur(*syn*); esta regularidad se asemeja a la superposición de las bases apiladas que ocurre en el dúplex con pares de bases de Hoogsteen (Fig. 4.3), cuando se observó la superposición del anillo base de Pyr-Pur(*syn*). La posición mutua de los dos anillos de Pyr en el fragmento con el par erróneo Gua:Thy es el resultado del desplazamiento de bases en este par debido a la formación de dos enlaces de H del grupo Amino de la base Guanina.

Tal como se mencionó antes, se puede observar en la Fig. 4.7 (ARRIBA y CENTRO) que el dCpdG:dTpdG (con el par Gua:Thy) y el dApdG:dCpdG (con el par Gua:Ade(*syn*)), respectivamente, muestran una superposición de sus bases apiladas tal como los WCDs; mientras que el dApdG:dCpdI (con el par Ino:Ade(*syn*)) presenta nula superposición entre Pur(*syn*):Pur y muy poca entre Pyr:Pur, Fig. 4.7 (ABAJO).

Optimizaciones con DFT de dTpdT y dTpdA, fragmentos de cadena única, partiendo de las conformaciones mostradas en las Fig. 4.5 y 4.6 producen estructuras con torsiones próximas a las de las conformaciones iniciales pero sin superposición sustancial de los anillos de las bases (ver la Sección 4.5). En general, se puede predecir que las posiciones mutuales de las bases vecinas en la cadena para dúplex con pares erróneos es el resultado de la conformación preferencial del esqueleto de azúcar-fosfato combinada con los desplazamientos de las bases debido a los enlaces de H no estándares entre los pares de bases.

Tabla 4.22: Ángulos de torsión del esqueleto azúcar-fosfato y ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8), para los mínimos de *dDMP-Na+* complementarios con el par erróneo *Ino:Ade(syn)* [27].

cdDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2	E
dTpdA:dIpdA										
dTpdA	-103	133	170	282	271	199	57	107	72	
dIpdA	-116	129	162	286	291	232	32	153	-92	
dTpdA	-122	141	198	277	291	168	49	105	76	-3050470.88
dIpdA	-134	144	173	270	282	227	50	147	-82	
dTpdA	-124	82	71	43	240	186	52	116	70	
dIpdA	-127	102	166	291	276	202	59	122	-98	
dTpdA	-128	121	71	82	190	166	52	97	81	-3050466.8
dIpdA	-125	142	196	275	290	177	50	125	-104	
dTpdA*	-127	122	70	82	190	166	52	98	81	
dIpdA	-125	142	196	276	290	177	50	126	-104	
dApdG:dCpdI										
dApdG	70	116	185	272	297	180	50	151	-89	
dCpdI	-143	85	205	238	328	129	66	102	-127	
dApdG**	59	146	173	264	295	197	55	140	-101	
dCpdI	-135	88	193	274	296	189	56	142	-100	
dApdG	58	148	179	248	304	182	56	145	-99	-3060551.72
dCpdI	-153	88	220	279	308	166	59	146	-103	

Ángulos en grados ($^{\circ}$).

Energía en *Kcal/mol*.

* Estructura optimizada empleando el programa *Gaussian 09D.1* con el funcional PW91 y la base 6-31g*.

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

Tabla 4.23: *Empaquetamiento del anillo de azúcar para los mínimos de dDMP-Na+ complementarios con el par erróneo Ino:Ade(syn) [27].*

cdDMPs	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
	Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
dTpdA:dIpdA						
dTpdA	C2'-endo	177	29	C1'-exo	116	42
dIpdA	C1'-exo	143	46	C3'-exo	195	39
dTpdA	C2'-endo	169	32	C1'-exo	115	42
dIpdA	C2'-endo	173	35	C3'-exo	207	34
dTpdA	C4'-exo	47	27	C1'-exo	135	36
dIpdA	O4'-endo	96	28	C2'-endo	164	23
dTpdA	C2'-endo	147	24	O4'-endo	104	43
dIpdA	C2'-endo	175	29	C1'-exo	138	40
dTpdA*	C2'-endo	149	24	O4'-endo	105	42
dIpdA	C2'-endo	176	29	C1'-exo	138	40
dApdG:dCpdI						
dApdG	C1'-exo	135	36	C2'-endo	162	47
dCpdI	O4'-endo	73	42	O4'-endo	96	28
dApdG**	C2'-endo	176	36	C2'-endo	177	33
dCpdI	O4'-endo	82	42	C3'-exo	184	32
dApdG	C2'-endo	173	37	C3'-exo	185	34
dCpdI	C3'-endo	15	32	C3'-exo	185	34

P y v_{max} en grados ($^{\circ}$).

* Estructura optimizada empleando el programa *Gaussian 09D.1* con el funcional PW91 y la base 6-31g*.

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

Tabla 4.24: *Parámetros de las bases y pares de bases de los dDMP-Na+ complementarios con el par erróneo Ino:Ade(syn)[27]* (véase designación en la fig. 2.10).

cdDMPs	Traslape	<i>Buckle</i> _{1,2}		<i>Propeller</i> _{1,2}		<i>Tilt</i>	<i>Roll</i>
dTpdA:dIpdA							
dTpdA	1.66	-7.3	4.0	-13.4	-12.1	-5.2	7.1
dIpdA	2.58						
dTpdA	2.8	-19.1	1.4	1.7	-5.9	-6.7	-2.4
dIpdA	0.12						
dTpdA	0.25	-5.9	10.4	-17.6	17.0	177.7	-13.7
dIpdA	1.34						
dTpdA	2.68	-5.8	14.1	-6.8	5.7	-174.0	15.2
dIpdA	2.35						
dTpdA*	2.56	-5.7	14.3	-6.4	6.5	-174.0	14.7
dIpdA	2.4						
dApdG:dCpdI							
dApdG	1.35	10.4	9.7	17.0	-4.9	140.3	-103.2
dCpdI	0						
dApdG**	0	-21.7	-19.6	1.2	-16.8	-138.3	111.2
dCpdI	0.3						
dApdG	0	-13.1	-12.5	9.7	-26.6	134.1	-114.5
dCpdI	0.63						

Área de traslape de las bases en Å², *Buckle*, *Propeller*, *Tilt* y *Roll* en grados (°).

* Estructura optimizada empleando el programa *Gaussian 09D.1* con el funcional PW91 y la base 6-31g*.

** Estructura optimizada empleando el programa ADF con el funcional PW91 y la base TZP.

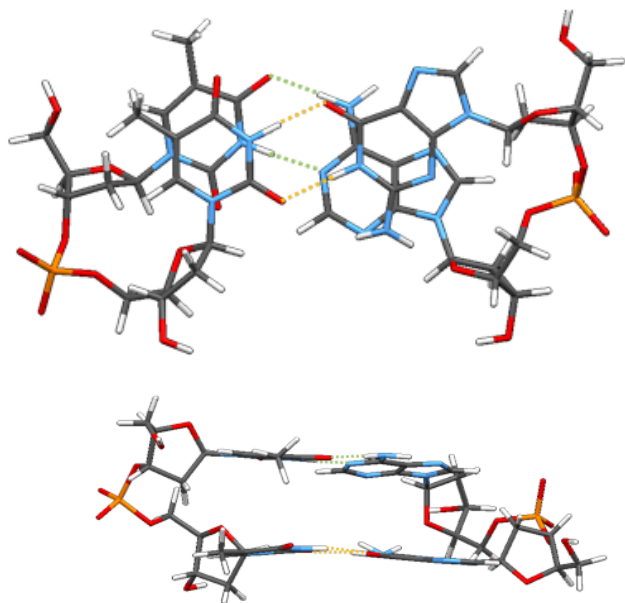


Figura 4.5: *Dúplex optimizado con DFT del dTpdT:dGpdA con el par erróneo Gua:Thy, mostrando la superposición de sus bases apiladas y la posición mutua de sus pares de bases.*

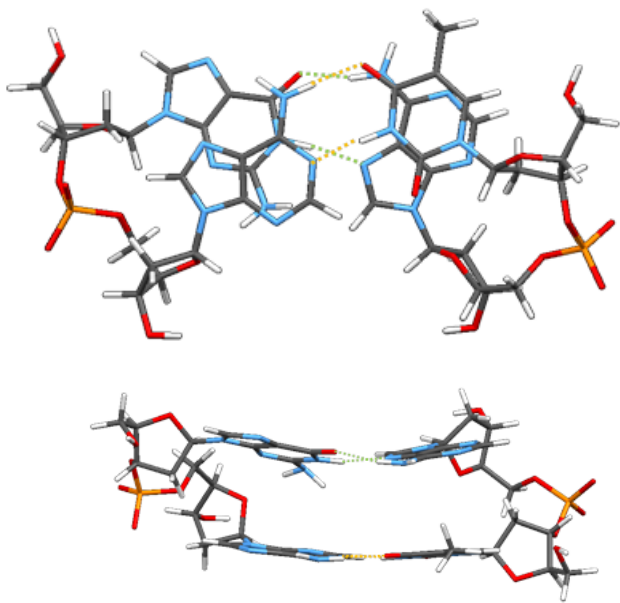


Figura 4.6: *Dúplex optimizado con DFT del dGpdA:dTpdA con el par erróneo Gua:Ade(syn), mostrando la superposición de sus bases apiladas y la posición mutua de sus pares de bases.*

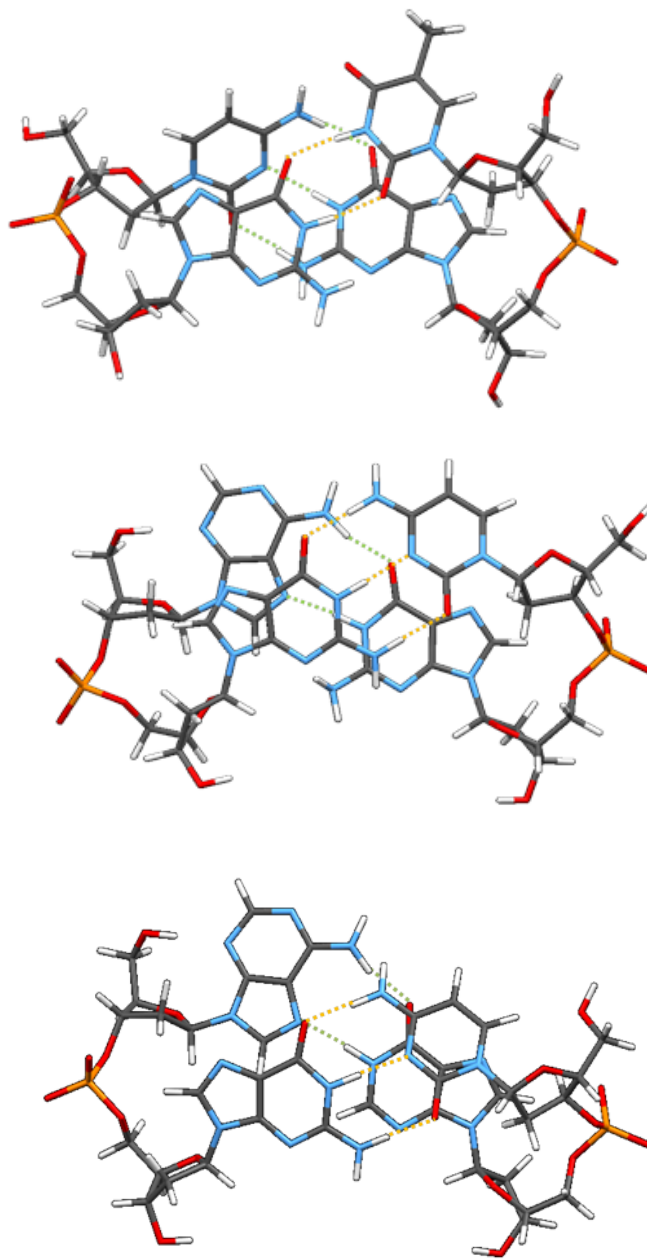


Figura 4.7: Conformaciones de dúplex con pares erróneos mostrando la superposición de sus bases apiladas. (ARRIBA) *dCpdG:dTpdG* (*Gua:Thy*). (CENTRO) *dApdG:dCpdG* (*Gua:Ade(syn)*). (ABAJO) *dApdG:dCpdI* (*Ino:Ade(syn)*).

4.5. Estructuras de cadena única

En trabajos anteriores [3, 6, 7, 8], con WCDs, se mostró que desoxidinucléosidos monofosfatos (dDMPs) con secuencias *purina-pirimidina* y *purina-purina* muestran una amplia superposición en los anillos de las bases, mientras que dDMPs con secuencias *pirimidina-purina* y *pirimidina-pirimidina* tienen superposición de los anillos de las bases despreciable o nula. A fin de ver si las regularidades obtenidas, es decir, si la gran superposición entre Pyr-Pyr (mostrada en las conformaciones de cadenas paralelas con el par de Watson-Crick reverso Ade:Thy y en los cdDMPs con el par erróneo Gua:Cyt) y entre Pyr-Pur(*syn*) (para las conformaciones con pares Ade(*syn*):Thy de Hoogsteen y en los cdDMPs con el par erróneo Gua:Ade(*syn*)) se cumplen también para estructuras de cadenas únicas (lo cual conduciría a que son posibles para WCDs, distinto a lo que se había encontrado antes), se sometieron a optimización estructuras de dTpdT y dTpdA cortados de los dúplex ya optimizados que presentaron dichas características.

En las Tablas 4.25 y 4.26 se encuentran los parámetros obtenidos para las estructuras iniciales y optimizadas de dTpdT, obtenidas de los dúplex con el par no correcto Gua:Thy y de los de cadenas paralelas que presentaron una gran superposición entre Pyr-Pyr después de ser optimizados con DFT. Como se puede observar en estas tablas, partiendo de distintos puntos y con distintos funcionales se obtuvieron parámetros estructurales casi iguales para todos los dDMPs optimizados, las torsiones se encuentran dentro de las regiones para los WCDs, además de que presentan un buen traslape entre los anillos de las bases apiladas; sin embargo, se tienen estructuras físicamente no realizables, pues presentan un gran ángulo entre los planos de las bases apiladas (Fig. 4.8), ver ángulo ϕ en la Tabla 4.25, lo cual no se tiene en el ADN ya que de ser así no podría enrollarse al tener otras interacciones.

Para las optimizaciones con DFT de la estructura de dTpdA, tomada de los dúplex con pares de Hoogsteen Ade(*syn*):Thy y del cdDMPs con el par no correcto Gua:Ade(*syn*), se obtuvieron resultados semejantes a los ya mencionados, ángulos de torsión en las regiones para los WCDs, superposición sustancial entre Pyr-Pur(*syn*) pero un gran ángulo entre las bases apiladas (Fig. 4.9).

De estos resultados, de las optimizaciones realizadas a conformaciones de cadena única dTpdT y dTpdA con gran superposición de las bases apiladas, se pudo corroborar que se siguen cumpliendo las consideraciones obtenidas anteriormente: no es posible tener sustancial traslape de bases Pyr-Pyr y Pyr-Pur(*syn*) en conformaciones de WC; sin embargo, sí se pueden tener estas regularidades en dúplex no de WC, ya que se tienen otras interacciones que estabilizan su estructura y dan pie a estas regularidades.

Tabla 4.25: *Ángulos de torsión del esqueleto azúcar-fosfato, ángulo glucosídico de torsión, χ , (véase designación en la fig. 2.8) y valores promedio del ángulo entre los planos de las bases apiladas, ϕ , para los mínimos de dTpdT-Na+ con gran superposición entre Pyr-Pyr.*

dDMPs	χ_1	δ_1	ϵ	ζ	α	β	γ	δ_2	χ_2	ϕ
dTpdT:dApdA de cadenas paralelas										
dTpdT ¹	-121	141	200	281	287	167	45	82	-118	12
dTpdT	-133	141	196	282	291	169	50	82	-108	35
dTpdT:dGpdA con par erróneo Gua:Thy										
dTpdT ¹	-124	143	194	283	287	170	53	81	-125	8
dTpdT	-133	141	195	282	288	170	54	76	-110	34
dTpdT ²	-115	115	186	277	288	173	58	86	-127	8
dTpdT*	-134	141	195	282	288	170	54	76	-110	35

Ángulos en grados (°).

1 dDMP de dúplex optimizado con DFT empleando el funcional PBE y la base 6-31g*.

2 dDMP de dúplex de banco de datos.

* Estructura optimizada empleando el programa *Gaussian 09D.1* con el funcional PW91 y la base 6-31g*.

Tabla 4.26: *Empaquetamiento del anillo de azúcar y área de superposición entre los anillos de las bases apiladas para los mínimos de dTpdT-Na+ con gran superposición entre Pyr-Pyr.*

dDMPs	Traslape	Nucleósido 5'-extremo			Nucleósido 3'-extremo		
		Empaquetamiento	P	v_{max}	Empaquetamiento	P	v_{max}
dTpdT:dApdA de cadenas paralelas							
dTpdT ¹	1.81	C2'-endo	167	31	C4'-exo	60	41
dTpdT	74 %	C2'-endo	159	35	C4'-exo	52	40
dTpdT:dGpdA con par erróneo Gua:Thy							
dTpdT ¹	3.52	C2'-endo	164	35	O4'-endo	73	39
dTpdT	76 %	C2'-endo	160	34	C4'-exo	54	38
dTpdT ²	1.65	C1'-exo	140	34	C4'-exo	69	32
dTpdT*	71 %	C2'-endo	160	34	C4'-exo	54	38

Área de traslape de las bases en Å², P y v_{max} en grados (°).

1 dDMP de dúplex optimizado con DFT empleando el funcional PBE y la base 6-31g*.

2 dDMP de dúplex de banco de datos.

* Estructura optimizada empleando el programa *Gaussian 09D.1* con el funcional PW91 y la base 6-31g*.

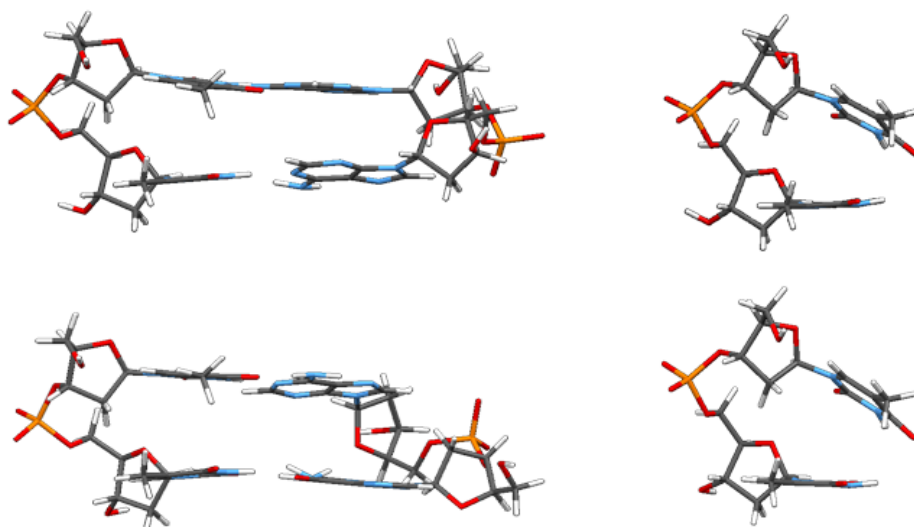


Figura 4.8: Conformaciones de cdDMPs optimizados con DFT que presentaron sustancial traslape entre Pyr-Pyr, mostrando la posición mutual de sus bases. (ARRIBA-Izq.) Dúplex de cadenas paralelas, dTpdT:dApdA, con superposición entre las bases Thy. (ARRIBA-Der.) Estructura de cadena única optimizada con DFT presentando bases no paralelas, dTpdT, tomada del dúplex con cadenas paralelas mencionado. (ABAJO-Izq.) Dúplex con par no correcto Gua:Thy, dTpdT:dGpdA, con superposición entre las bases Thy. (ABAJO-Der.) Estructura de cadena única optimizada con DFT presentando bases no paralelas, dTpdT, tomada del dúplex con par erróneo Gua:Thy mencionado.

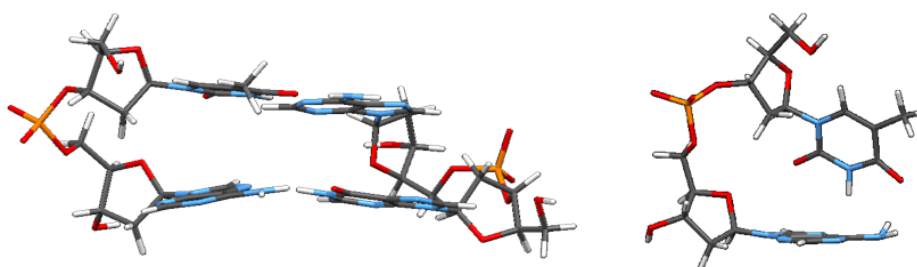


Figura 4.9: Conformación de cdDMPs optimizados con DFT que presentaron sustancial traslape entre Pyr-Pur(syn), mostrando la posición mutual de sus bases. (IZQ.) Dúplex con el par no correcto Gua:Ade(syn), dTpdA:dGpdA, con superposición entre las bases Thy-Ade. (DER.) Estructura de cadena única optimizada con DFT presentando gran ángulo entre sus bases, dTpdA, tomada del dúplex con par erróneo Gua:Ade(syn) mencionado.

Cabe mencionar que, al optimizar los dDMPs con Mecánica Molecular, dTpdT y dTpdA que presentaron una gran superposición en los dúplex optimizados con DFT, éstas estructuras de cadena única presentan parámetros geométricos semejantes a los WCDs, además de que sus bases apiladas se encuentran paralelas entre sí, ver Fig. 4.10; sin embargo, muestran una gran superposición entre Pyr-Pyr y Pyr-Pur. Esto nos indica que AMBER no reproduce las regularidades que DFT sí da para dDMPs (DFT realiza cálculos más exactos), con lo cual es claro que el campo de fuerzas AMBER necesita un mejoramiento en las interacciones de apilamiento pues considera un poco más de lo normal.

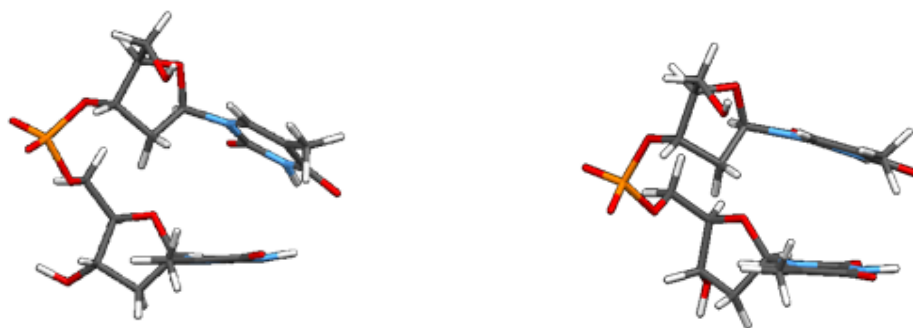


Figura 4.10: Estructuras de cadena única, dDMPs, tomadas de dúplex que presentaron una gran superposición entre Pyr-Pyr. (IZQUIERDA) dTpdT con gran traslape entre sus bases apiladas, optimizado con DFT. (DERECHA) dTpdT con gran traslape entre sus bases apiladas, optimizado con AMBER.

Capítulo 5

Conclusiones

Las leyes físicas subyacentes necesarias para la teoría matemática de gran parte de la física y la química son completamente conocidas, la dificultad es sólo que la aplicación exacta de estas leyes conduce a ecuaciones mucho más complicadas de resolver.

P. A. M. DIRAC, 1929

Los cálculos no empíricos de la Mecánica Cuántica realizados a fragmentos mínimos de dúplex de desoxidinucleósidos monofosfatos complementarios con iones de sodio, cdDMPs-Na⁺, mostraron que la posición mutua de las subunidades en conformaciones de la familia estructural Z no corresponden a los mínimos de energía de cdDMPs, las disposiciones mutuas de las bases están muy lejos de las del dúplex experimental, tal como se había encontrado antes para dDMPs [3, 8]. Estos resultados refuerzan la conclusión de que la formación de las estructuras de la familia Z se origina a partir de las interacciones del ADN con su entorno, como se sugiere en la literatura [20].

La posición mutua de las subunidades en otras conformaciones de hélices dobles, tales como dúplex de Hoogsteen y estructuras de cadenas paralelas que tienen apareamiento de nucleósidos diferentes de los de Watson-Crick, así como estructuras con pares erróneos, corresponden a mínimos de energía de cdDMPs, pero, la dependencia de las secuencias en tales dúplex difiere drásticamente de la de los WCDs.

Es notoria la importancia del análisis de estas conformaciones diferentes del modelo de Watson-Crick pues pueden ser partícipes de distintos procesos y funciones del ADN en las células. Las optimizaciones realizadas a estos dúplex con los diferentes métodos y funcionales permitieron constatar la viabilidad de su estructura.

Del análisis realizado a dúplex con pares de Hoogsteen, con las secuencias dTpdT:dApdA, (dTpdA)₂ y (dApdT)₂, es claro que dichas conformaciones presentan buenos parámetros estructurales, tal como los WCDs, pero, muestran distinta superposición entre los anillos de las bases apiladas al que se tiene en WCDs. Estas conformaciones con pares Ade(*syn*):Thy de Hoogsteen no tienen secuencias con traslape base-base casi nulo en los pares apilados, y, en consecuencia, tendrían dificultades para doblarse.

Por otro lado, los cálculos realizados a las conformaciones de cadenas paralelas con el par reverso de Watson-Crick Ade:Thy permitieron obtener interesantes regularidades, ya que, a pesar de presentar sus torsiones dentro de las regiones para los WCDs, estas estructuras muestran un notorio traslape entre Pyr-Pyr, mientras que es casi nulo en las secuencias Pur-Pur, Pur-Pyr y Pyr-Pur.

Las regularidades de la disposición mutua de las bases en fragmentos de WCDs con pares erróneos difieren de las de los dúplex con pares *canónicos*. Estos arreglos corresponden a mínimos energéticos locales para cdDMPs con pares «no correctos», y pueden tener considerables superposiciones de las bases apiladas Pyr-Pyr y Pyr-Pur(*syn*). Los enlaces de H «no-estándares» entre los pares de bases que existe en estos dúplex es responsable del desplazamiento de las bases en los patrones de la posición mutua de nucleósidos vecinos.

Además, de los resultados obtenidos al optimizar conformaciones de cadena única, dTpdT y dTpdA con gran superposición de las bases apiladas, se pudo corroborar que se siguen cumpliendo las consideraciones obtenidas anteriormente: no es posible tener sustancial traslape de bases Pyr-Pyr y Pyr-Pur(*syn*) en conformaciones de WC; sin embargo, sí se pueden tener estas regularidades en dúplex no de WC, ya que se tienen otras interacciones que estabilizan su estructura y dan pie a estas regularidades.

Los cálculos computacionales realizados con DFT empleando los funcionales PBE (base 6-31g*) y PW91 (base *TZP*) muestran que ambos métodos nos arrojan resultados semejantes, siendo buenos para reproducir las características estructurales de conformaciones diferentes de los WCDs. En cuanto a los cálculos con Mecánica Molecular, parece ser que éstos pueden reproducir la mayoría de regularidades reveladas en los cálculos de DFT para los cdDMPs. Sin embargo, no es así para los dDMPs, no todas las características dependientes de la secuencia pueden ser reproducidas por tales cálculos. Obtuvimos evidencia de que AMBER, y posiblemente otros campos de fuerza populares, sobrestiman algunas interacciones de apilamiento.

Apéndice A

Distribución de ángulos de torsión y de pseudorotación para las familias estructurales AI, AII, BI y BII

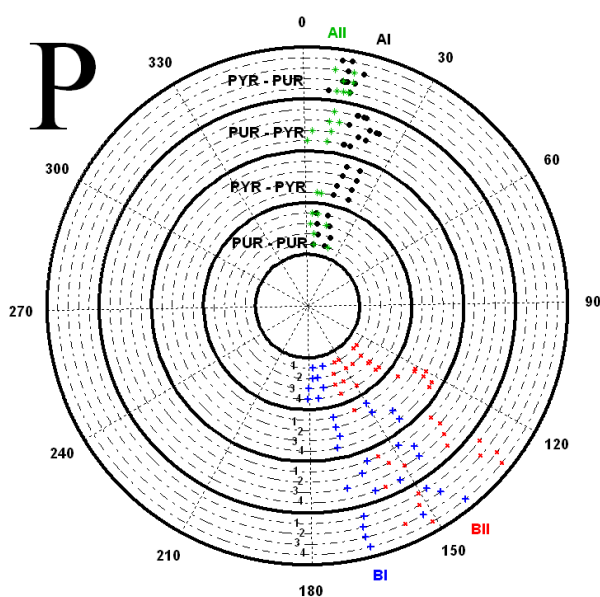


Figura A.1: Distribución del ángulo de pseudorotación, P , del anillo de azúcar en las estructuras optimizadas de dDMPs, correspondientes a dúplex diestros. Cada secuencia de bases se localiza en una circunferencia punteada. Cuatro grupos de secuencias están separados por círculos sólidos. Estos son Pur-Pur: 1-dApdA, 2-dGpdG, 3-dAdpG, 4-dGdpA; Pyr-Pyr: 1-dCpdC, 2-dTpdC, 3-dCpdT, 4-dTpdT; Pur-Pyr: 1-dApdC, 2-dApdT, 3-dGpdT, 4-dGpdC; Pyr-Pur: 1-dCpdA, 2-dTpdG, 3-dCpdG, 4-dTpdA.

¹Gráficas tomadas de la Ref.[28]

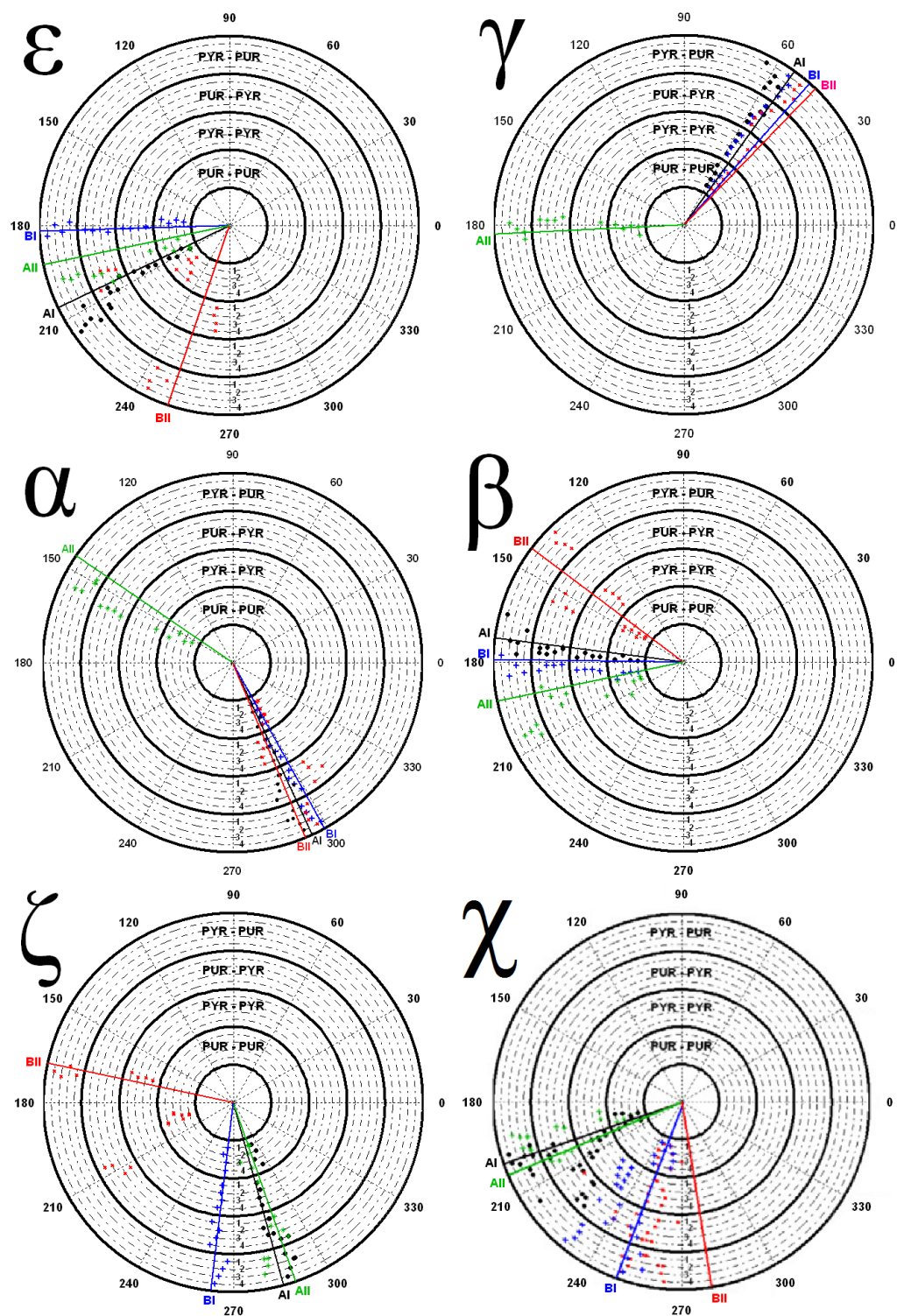


Figura A.2: Distribución de los ángulos de torsión del esqueleto azúcar-fosfato para estructuras optimizadas de dDMPs. Los valores promedio de los ángulos de torsión para cada una de cuatro familias de dúplex en cristales se indican mediante líneas rectas. Otras designaciones son explicadas en la leyenda de la Fig. A.1.

Bibliografía

*Los libros son faros en el gran
mar de tiempo.*

E. P. WHIPPLE

- [1] H. M. Berman, W. K. Olson, D. I. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S. H. Hsieh, A. R. Srinivasan, and B. Schneider. The nucleic acid database. a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophysical Journal*, 63, No. 3, pp. 751-759, 1992.
- [2] J. D. Watson and F. H. Crick. Molecular structure of nucleic acids. *Nature*, 171, No. 4356, pp. 737-738, 1953.
- [3] Vareli Poltev, Victor M. Anisimov, Victor I. Danilov, Dolores García, Carolina Sánchez, Alexandra Deriabina, Eduardo González, Francisco Rivas, and Nina Polteva. The role of molecular structure of sugar-phosphate backbone and nucleic acid bases in the formation of single-stranded and double-stranded dna structures. *Biopolymers*, 101, no. 6, pp. 640-650, 2014.
- [4] Bill Bryson. *Una Breve Historia de Casi Todo*. RBA bolsillo, séptima edición, 2008.
- [5] D. Svozil, J. Kalina, M. Omelka, and B. Schneide. Dna conformations and their sequence preferences. *Nucleic Acids Research*, 36, pp. 3690-3706, 2008.
- [6] V. I. Poltev, V. M. Anisimov, V. I. Danilov, A. Deriabina, E. Gonzalez, A. Jurkiewicz, A. Leś, and N. Polteva. Dft study of b-like conformations of deoxydinucleoside monophosphates containing gua and/or cyt and their complexes with na+ cation. *Journal of Biomolecular Structure and Dynamics*, 25, pp. 563-572, 2008.
- [7] V. I. Poltev, V. M. Anisimov, V. I. Danilov, A. Deriabina, E. Gonzalez, D. Garcia, F. Rivas, A. Jurkiewicz, and A. Leś and N. Polteva. Dft study of minimal fragments of nucleic acid single chain for explication of sequence dependence of

- dna duplex conformation. *Journal of Molecular Structure THEOCHEM*, 912, pp. 53-59, 2009.
- [8] Vareli I. Poltev, Victor M. Anisimov, Victor I. Danilov, Tanja Van Mourik, Alexandra Deriabina, Eduardo González, Maria Padua, Dolores García, Francisco Rivas, and Nina Polteva. Dft study of polymorphism of the dna double helix at the level of dinucleoside monophosphates. *International Journal of Quantum Chemistry*, 110, pp. 2548-2559, 2010.
- [9] K. Hoogsteen. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta crystallographica*, 12, no. 10, pp. 822-823, 1959.
- [10] G. A. Patikoglou, J. L. Kim, L. Sun, S. H. Yang, T. Kodadek, and S. Burley. Tata element recognition by the tata box-binding protein has been conserved throughout evolution. *Genes and Development*, 13(24), pp. 3217-3230, 1999.
- [11] N. A. Tchurikov, B. K. Chernov, Y. B. Golova, and Y. D. Nechipurenko. Parallel dna: generation of a duplex between two drosophila sequences in vitro. *FEBS letters*, 257, pp. 415-418, 1989.
- [12] F. H. Crick, L. Barnett, S. Brenner, and R. J. Watts-Tobin. General nature of the genetic code for proteins. *Nature*, 192, no. 4809, pp. 1227-1232., 1961.
- [13] W. Traub and J. L. Sussman. Adenine-guanine base pairing in ribosomal rna. *Nucleic acids research*, 10, no. 8, pp. 2701-2708, 1982.
- [14] Y. Takahashi, K. Kato, Y. Hayashizaki, T. Wakabayashi, E. Ohtsuka, S. Matsuki, and K. Matsubara. Molecular cloning of the human cholecystokinin gene by use of a synthetic probe containing deoxyinosine. *Proceedings of the National Academy of Sciences*, 82, no. 7, pp. 1931-1935, 1985.
- [15] F. H. Martin, M. M. Castro, F. Aboul-ela, and I. Tinoco. Base pairing involving deoxyinosine: implications for probe design. *Nucleic acids research*, 13, no. 24, pp. 8927-8938, 1985.
- [16] F. H. Crick. Codon-anticodon pairing: the wobble hypothesis. *Journal of molecular biology*, 19, no. 2, pp. 548-555, 1966.
- [17] Verónica Domínguez Benítez. Estudios de conformaciones de fragmentos mínimos de la hélice doble de adn. Tesis de Licenciatura, FCFM, BUAP, Noviembre 2014.
- [18] Xiang-Jun Lu and Wilma K. Olson. 3dna: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research*, 31, no. 17, 2003.

- [19] Richard E. Dickerson, Horace R. Drew, Benjamin N. Conner, Richard M. Wing, Albert V. Franiti, and Mary L. Kopka. The anatomy of a-, b-, and z-dna. *Science*, 216, pp. 475-485, 1982.
- [20] K. Brzezinski, A. Brzuszkiewicz, M. Dauter, M. Kubicki, M. Jaskolski, and Z. Dauter. High regularity of z-dna revealed by ultra high-resolution crystal structure at 0.55Å. *Nucleic Acids Research*, gkr202, 2011.
- [21] N.G. Abrescia, C. González, C. Gouyette, and J.A. Subirana. X-ray and nmr studies of the dna oligomer d(atatat)₂: Hoogsteen base pairing in duplex dna. *Biochemistry*, 43, no. 14, pp. 4092-4100, 2004.
- [22] Francisco J. Acosta-Reyes, Elida Alechaga, Juan A. Subirana, and J. Lourdes Campos. Structure of the dna duplex d(attaat)₂ with hoogsteen hydrogen bonds. *PLoS ONE*, 2015.
- [23] V. Rani Parvathy, Sukesh R. Bhaumik, Kandala V. R. Chary, Girjesh Govil, Keliang Lui, Frank B. Howard, and H. Todd Miles. Nmr structure of a parallel-stranded dna duplex at atomic resolution. *Nucleic Acids Research*, 30, no. 7, 2002.
- [24] C. M. Radding. Genetic recombination: strand transfer and mismatch repair. *Annual review of biochemistry*, 47, no. 1, pp. 847-880, 1978.
- [25] William N. Hunter, Tom Brown, G. Kneale, Naveen N. Anand, Dov Rabinovich, and Olga Kennard. The structure of guanosine-thymidine mismatches in b-dna at 2.5-Å resolution. *Journal of Biological Chemistry*, 262(21):9962-9970, 1987.
- [26] T. Brown, W. N. Hunter, G. Kneale, and O. Kennard. Molecular structure of the ga base pair in dna and its implications for the mechanism of transversion mutations. *Proceedings of the National Academy of Sciences*, 83, no. 8, pp. 2402-2406, 1986.
- [27] Peter W. R. Corfield, William N. Hunter, Tom Brown, Paul Robinson, and Olga Kennard. Inosine.adenine base pairs in a b-dna duplex. *Nucleic Acids Research*, 15, no. 19, 1987.
- [28] V. I. Poltev, V. M. Anisimov, V. I. Danilov, D. García, A. Deriabina, E. González, R. Salazar, F. Rivas, and N. Polteva. Dft study of dna sequence dependence at the level of dinucleoside monophosphates. *Computational and Theoretical Chemistry*, 975, pp. 69-75, 2011.
- [29] J. C. Slater. The theory of complex spectra. *Physical Review*, 34, no. 10, pp. 1293-1322, 1929.

- [30] J. C. Slater. Molecular energy levels and valence bonds. *Physical Review*, 38, no. 6, pp. 1109-1144, 1931.
- [31] J. C. Slater. Atomic shielding constants. *Physical Review*, 36, no. 1, pp. 57-64, 1930.
- [32] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka. Prescription for the design and selection of density functional approximations: More constraint satisfaction with fewer fits. *The Journal of chemical physics*, 123, no. 6, pp. 062201-1(9), 2005.
- [33] Valeri Poltev. *Molecular Mechanics: Principles, History, and Current Status*, pages 1–48. Springer Netherlands, Dordrecht, 2016.
- [34] J.P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77, no. 18, pp. 3865, 1996.
- [35] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical review B*, 46, no. 11, pp. 6671, 1992.
- [36] T. Tsuzuki and H. P. Luthi. Interaction energies of van der waals and hydrogen bonded systems calculated using density functional theory: Assessing the pw91 model. *Journal of Chemical Physics*, 114, pp. 3949-3957, 2001.
- [37] Van der Wijst, C. F. Guerra, M. Swart, and F. M. Bickelhaupt. Performance of various density functionals for the hydrogen bonds in dna base pairs. *Chemical Physics Letters*, 426, pp. 415-421, 2006.
- [38] J. Sponer, P. Jurecka, and P. Hobza. Interaction energies of hydrogen-bonded nucleic acid base pairs. *Journal of the American Chemical Society*, 13, pp. 1245-1257, 2004.
- [39] M. J. Frisch, et al G. W. Trucks, and D. J. Fox. *Gaussian 09, Revision D.01*. Gaussian, Inc., Wallingford CT, 2013.
- [40] G.T. Te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J. van Gisbergen, and J.G. Snijders and T. Ziegler. *Chemistry with ADF*, 2001.
- [41] D.A. Case, T.A. Darden, III T.E. Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C. Schafmeister, W.S. Ross, and P.A. Kollman. *AMBER 9*. University of California, San Francisco.

- [42] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, and D. M. Ferguson and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117, no. 19, pp. 5179-5197, 1995.
- [43] T. E. Cheatham III, P. Cieplak, and P. A. Kollman. A modified version of the cornell et al. force field with improved sugar pucker phases and helical repeat. *Journal of Biomolecular Structure and Dynamics*, 16, no. 4, pp. 845-862, 1999.
- [44] J. Wang, P. Cieplak, and P. A. Kollman. How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules? *Journal of computational chemistry*, 21, no. 12, pp. 1049-1074, 2000.
- [45] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Loughton, and M. Orozco. Refinement of the amber force field for nucleic acids: improving the description of α/γ conformers. *Biophysical journal*, 92, no. 11, pp. 3817-3829, 2007.
- [46] A.H. Wang, G.J. Quigley, F.J. Kolpak and J.L Crawford, J.H. Van Boom, G. van der Marel, and A. Rich. Molecular structure of a left-handed double helical dna fragment at atomic resolution. *Nature*, 282, no. 5740, pp. 680, 1979.
- [47] M.D Frank-Kamenetskii and S.M. Mirkin. Triplex dna structures. *Annual review of biochemistry*, 64, no. 1, pp. 65-95, 1995.
- [48] J.D. Watson and F.H. Crick. Genetical implications of the structure of deoxyribonucleic acid. *JAMA*, 269, no. 15, pp. 1967-1969, 1993.