



BUAP

BENEMÉRITA UNIVERSIDAD
AUTÓNOMA DE PUEBLA

Facultad de Ciencias Fisico Matemáticas

Estudios de Conformaciones de fragmentos mínimos de la hélice doble de ADN

Tesis presentada al

Colegio de Física

como requisito parcial para la obtención del grado de

Licenciada en Física

por

Verónica Domínguez Benítez

asesorada por

Dr. Valeri Poltev (FCFM-BUAP)
Dra. Dolores García Toral (FIQ-BUAP)

NOVIEMBRE 2014

Este documento está preparado para ser impreso a doble cara.



BUAP

BENEMÉRITA UNIVERSIDAD
AUTÓNOMA DE PUEBLA

Facultad de Ciencias Fisico Matemáticas

Estudios de Conformaciones de fragmentos mínimos de la hélice doble de ADN

Tesis presentada al

Colegio de Física

como requisito parcial para la obtención del grado de

Licenciada en Física

por

Verónica Domínguez Benítez

asesorada por

Dr. Valeri Poltev (FCFM-BUAP)

Dra. Dolores García Toral (FIQ-BUAP)

NOVIEMBRE 2014

Copyright © Verónica Domínguez Benítez

Título: Estudios de Conformaciones de fragmentos mínimos de la hélice doble de ADN.

Estudiante: Verónica Domínguez Benítez

COMITÉ

Dr. Eduardo González Jiménez
PRESIDENTE

Dra. Alexandra Deriavina
SECRETARIO

Dra. Ma. Luisa Roxana Licona Ibarra
VOCAL

M.C. Pablo Martínez Garcilazo
SUPLENTE

Dr. Valeri Poltev
ASESOR

Dra. Dolores García Toral
ASESORA

*Con todo mi amor para mis papás:
Nelly y Santiago, y a mis hermanas,
quienes me han apoyado en toda mi travesía,
me brindaron su mano cuando sentía que
el camino se terminaba, me motivaron a seguir,
me dieron su amor incondicional y
no me dejaron sola cuando más los necesité.*

*A tí que recorriste a mi lado
todo mi trayecto, fuiste mi apoyo,
me brindaste tu amor y
sacrificaste tu tiempo para que yo
pudiese cumplir con el mío.*

*A tí mi brillante lucerito,
tú que llegaste a llenar de alegría
y de luz mi vida, haciendo que cada día
tuviese un motivo por el cual sonreír,
por el cual luchar y no desistir,
alguien por quién dar lo mejor de mí...
A mis dos angelitos, mis dos amores:
Sergio y Mateo.*

Agradecimientos

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

Doy gracias a Dios por la familia que me dió. Unos papás amorosos que me han brindado su apoyo incondicional, han velado por mí, no me dieron la espalda cuando pensé que todo estaba perdido, y, en su lugar, me apoyaron con amor para poder seguir adelante, corrigiendo mis faltas y celebrando mis triunfos. Unas hermanas con las cuales crecer, con quienes contar en los buenos y malos momentos, con quienes reír de la vida y formar un lazo para siempre. Unos abuelitos cariñosos dispuestos a dar su amor en todo momento, y ese abuelito juguetón con quién reírse, ver lo divertido de la vida y volver a ser niño, quien vió que nunca me faltara nada y me apoyó en mi estancia fuera de casa.

Doy gracias por tener alguien con quien compartir mi vida, con quien conocer el amor de pareja, quien, a pesar de los senderos oscuros que se presentaron, estuvo a mi lado en toda esta travesía, de quien me pude apoyar y salir adelante juntos, quien hizo de los momentos juntos tiempo de respirar, de recordarme quién soy y quién quiero ser. Bendecida me siento por el fruto de ese amor, ese angelito que llena de luz mi vida, de alegría mis tristezas, de ilusiones mis sueños, de fortaleza mis derrotas; quien ha tenido que sacrificar nuestro tiempo juntos y sin embargo tener una sonrisa, chistosadas y amor para darme, a tí todos mis triunfos, por tí lucharé por ser mejor.

Agradezco la oportunidad de conocer y trabajar con el Dr. Valeri Poltev, un excelente científico y persona, quien compartió sus anécdotas y conocimientos conmigo, recordándome lo fascinante que es la ciencia y que uno debe luchar para lograr sus sueños y estar orgulloso de lo que logró en la vida; y a su esposa Nina quien me mostró calidez y empatía. Agradezco a la Dra. Dolores García Toral, por su paciencia, por su apoyo cuando lo necesité y por alentarme a mejorar.

A todos aquéllos profesores que tuve la oportunidad de conocer, pues me brindaron sus conocimientos y amor por la física, e hicieron que me enamorara de ésta aún más y me adentrara en su encanto, haciéndome tener la necesidad de compartir esa fascinación. Doy gracias por todas aquéllas personas que conocí en ésta etapa de mi vida, aquéllos amigos con los que comencé este camino y otros con los que lo terminé, pues tuve su apoyo e hicieron que fuese un recorrido más grato y ameno. Todo lo que viví a su lado son recuerdos que llevaré por siempre conmigo, al igual que su amistad.

Finalmente a la VIEP por financiar el proyecto “*Estudio comparativo de la organización espacial de fragmentos de ácidos nucleicos cuyos modelos se construyen según resultados experimentales y computacionales*”, del cual forma parte ésta tesis.

Resumen

«Soy de las que piensan que la ciencia tiene una gran belleza. Un sabio en su laboratorio no es solamente un teórico, es también un niño colocado ante los fenómenos naturales que le impresionan como un cuento de hadas.»

MARIE CURIE,
ganadora de los Premios Nobel de Física y Química.

El ADN es la molécula biológica más importante ya que es capaz de almacenar toda la información genética necesaria para formar un ser vivo completo y además duplicarse. Dicho papel se cumple gracias a su gran riqueza estructural. Ésta macromolécula es una hélice doble, la cual consta de dos cadenas complementarias de polinucleótidos, las cuales están formadas por un conjunto de nucleósidos conectados unos con otros mediante un grupo fosfato. En éste trabajo, en primera instancia, se estudiaron Trinucleósidos Difosfatos Complementarios para analizar si son o no buenos modelos para realizar nuestros estudios. Posteriormente se hace un estudio de diez posibles fragmentos mínimos del ADN, los cuales están conformados por dos Desoxidinnucleósidos Monofosfatos (dDMPs) complementarios, de la familia estructural B1. Éstos fueron sometidos a minimización utilizando teoría de primeros principios, en este caso se recurrió a la Teoría de Funcional de Densidad (DFT, por sus siglas en inglés), para analizar las características comunes y las diferencias, así como las dependencias de secuencias de nucleósidos, ya que es sabido que el arreglo mutal de las bases en los dDMPs es dependiente de éstas, es decir, dos dDMPs que contienen los mismos nucleósidos pero en distinta secuencia presentan una diferencia drástica en las posiciones mutuales de las bases. Además, se estudiaron tres fragmentos mínimos de ADN, con las secuencias A-T (Adenina-Timina), T-A (Timina-Adenina) y G-C (Guanina-Citosina), sometidos a minimización por los métodos MP2 y el Funcional PWB6K para el análisis de diferencias en las características conformacionales entre dichas estructuras y las minimizadas por la Funcional Estándar PW91, ésto con el fin de comparar los diferentes métodos.

Objetivos

No hay ningún área del mundo que no deba ser investigada por los científicos. Siempre existirá algunas preguntas que no han sido contestadas. En general, éstas son las preguntas que aún no han sido planteadas.

LINUS PAULING,
cristalógrafo importante ganador de los Premios
Nobel de Química y de la Paz.

1. Objetivo General

Comparación de características estructurales de diez fragmentos mínimos del ADN, de la familia conformacional B1, optimizados por el código ADF basado en la Teoría de Funcional de Densidad. Estudio de la dependencia de las conformaciones de los fragmentos mínimos obtenidos mediante diferentes métodos de mecánica cuántica.

2. Objetivos Particulares

- Preparación de diez archivos posibles de fragmentos mínimos de ADN, de la familia BI, para su minimización por el método DFT.
- Medición de características estructurales y su organización en tablas.
- Construcción de gráficos para la comparación de parámetros y características estructurales.
- Conclusiones de características comunes para las diferentes secuencias.
- Diferencias y dependencias debidas a las secuencias de nucleósidos.
- Prueba del Funcional PWB6K.
- Análisis de los cálculos para las características conformacionales de fragmentos mínimos del ADN de la conformación de dDMPs complementarios, optimizados por los métodos MP2 y DFT.

Índice

Agradecimientos	XI
Resumen	XIII
Objetivos	XV
1. Introducción	1
1.1. Breve historia sobre el ADN	1
1.2. Importancia biológica y estudios del ADN	5
2. Estructura del ADN	9
2.1. Estructura Secundaria del ADN	13
2.1.1. Parámetros geométricos para el estudio del ADN	15
3. Metodología para el estudio Teórico-Computacional del ADN	19
3.1. Mecánica Cuántica	20
3.1.1. Teoría del Funcional de la Densidad (DFT)	23
3.1.2. Teoría de perturbaciones Møller-Plesset de segundo orden (MP2)	28
3.2. Metodología	31
3.2.1. Construcción de diez estructuras iniciales	34
3.2.2. Medición de parámetros	35
4. Resultados	37
4.1. Selección de fragmentos de la hélice doble del ADN y detalles de método de DFT para estudio	37
4.2. Estudio DFT de Fragmentos Mínimos de la hélice doble del ADN . .	42
4.2.1. Estudio de diez fragmentos mínimos de la hélice doble del ADN con el funcional <i>PW91</i>	42

4.2.2. Comparación de resultados obtenidos con los Funcionales <i>PWB6K</i> y <i>PW91</i> para dDMPs-Na ⁺ complementarios con la secuencia Adenina-Timina	53
4.3. Comparación de resultados obtenidos por los métodos MP2 y DFT . . .	57
4.3.1. dDMPs complementarios con secuencia Timina-Adenina . . .	57
4.3.2. dDMPs complementarios con secuencia Guanina-Citosina . . .	61
5. Conclusiones	65
A. Distribución de ángulos de torsión y de pseudorotación para las familias estructurales AI, AII, BI y BII	67
Bibliografía	69

Índice de figuras

1.1.	<i>“Foto 51”, tomada por Franklin, que muestra un patrón de difracción de rayos X de ADN de la forma B</i>	4
1.2.	<i>Estructura de la hélice doble del ADN. (ARRIBA) Fragmento de ADN de la familia B1, vista aérea (imagen a la izq.) y frontalmente con diferentes efectos. (ABAJO) Fragmento mínimo de ADN, ejemplo de dDMPs complementarios, con los pares de bases Adenina-Timina y Guanina-Citosina.</i>	5
1.3.	<i>Desoxydinucleósido Monofosfato neutralizado con Na+, fragmento mínimo de la cadena única del ADN.</i>	7
2.1.	<i>Estructura de un nucleótido</i>	10
2.2.	<i>Bases nitrogenadas y pares que se forman.</i>	10
2.3.	<i>Estructura Secundaria, Terciaria y Cuaternaria del ADN</i>	12
2.4.	<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 [9].</i>	14
2.5.	<i>Designación de los ángulos de torsión de un dDMP.</i>	15
2.6.	<i>Empaquetamiento del azúcar.</i>	16
2.7.	<i>Descripciones gráficas de los parámetros Buckle, Propeller, h-Rise y h-Twist, empleados para describir la geometría de pares de bases complementarios y secuencias de pares de bases.</i>	17
3.1.	<i>Unidad mínima de la hélice doble del ADN. Dinucleósidos monofosfatos complementarios.</i>	32
3.2.	<i>Bases de los dDMPs con su enlace glucosídico.</i>	34
4.1.	<i>Designación de las bases nitrogenadas en el programa 3DNA para determinar la superposición entre éstas.</i>	42

4.2.	<i>Distribución de ángulos para las estructuras minimizadas de los dDMPs complementarios (con diferentes secuencias) de la familia estructural BI. a) (ARRIBA) Ángulos glucosídicos (χ) y de torsión. b) (ABAJO) Ángulos de pseudorotación y ángulos ϵ. NOTA: En la designación de secuencia de nucleótidos sólo se menciona la secuencia de un dDMP, no del complementario.</i>	47
4.3.	<i>Parámetros de los pares de bases de los dDMPs complementarios minimizados para la familia estructural BI, manteniendo el mismo esqueleto, con respecto al eje de la hélice.</i>	48
4.4.	<i>Valores de los parámetros Buckle y Propeller para los pares de bases en los dDMPs complementarios previamente minimizados, de la familia estructural BI.</i>	49
4.5.	<i>Superposición de las bases, para los dDMPs complementarios minimizados de la familia estructural BI, con el mismo esqueleto inicial, debida al Stacking entre ellas. a) (ARRIBA) Área de superposición de los anillos de las bases. b) (ABAJO) Área de superposición de los anillos de las bases y sus radicales.</i>	50
4.6.	<i>Estructuras optimizadas de dDMPs complementarios de la familia estructural BI, ejemplos característicos: dApdT b) dGpdG y dCpdC c) dTpdA.</i>	51
4.7.	<i>Estructura optimizada de dDMPs complementarios con las secuencias Timina-Citosina (dDMP a la izquierda) y Guanina-Adenina (dDMP a la derecha), vista de distintos ángulos.</i>	52
4.8.	<i>Estructuras optimizadas de dDMPs complementarios de la familia estructural BI, con la secuencia Timina-Adenina. a) (ARRIBA) dDMPs complementarios optimizados mediante el método MP2. b) (ABAJO) dDMPs complementarios optimizados mediante el método DFT-PW91.</i>	58
A.1.	<i>Distribución del ángulo pseudorotation 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>	67
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.</i>	68

Índice de Tablas

2.1.	<i>Terminología empleada para referirse a los nucleósidos y nucleótidos.</i>	10
4.1.	<i>Ángulos de torsión del esqueleto azúcar-fosfato (véase designación en la figura 2.5) y valores promedio del ángulo entre los planos de las bases (ϕ), para los $(dA)_3:(dT)_3$ optimizados, en distintas condiciones.</i>	38
4.2.	<i>Empaquetamiento del anillo de azúcar y ángulo glucosídico de torsión, χ, para los $(dA)_3:(dT)_3$ optimizados, en distintas condiciones.</i>	39
4.3.	<i>Parámetros de las bases y pares de bases de los $(dA)_3:(dT)_3$ optimizados. (véase designación en la figura 2.7).</i>	40
4.4.	<i>Distancias más cortas entre las bases del $dApdApdA-Na^+$, modelo con agua.</i>	41
4.5.	<i>Ángulos de torsión del esqueleto azúcar-fosfato (véase designación en la figura 2.5) y valores promedio del ángulo entre los planos de las bases (ϕ), para los mínimos de $dDMP-Na^+$ complementarios correspondientes a la familia estructural BI.</i>	43
4.6.	<i>Empaquetamiento del anillo de azúcar y ángulo glucosídico de torsión, χ, para los mínimos de $dDMP-Na^+$ complementarios correspondientes a la familia estructural BI.</i>	44
4.7.	<i>Parámetros de las bases y pares de bases de los $dDMPs$ complementarios, con el mismo esqueleto (véase designación en la figura 2.7).</i>	45
4.8.	<i>Ángulos de torsión del esqueleto azúcar-fosfato (véase designación en la figura 2.5) y valores promedio del ángulo entre los planos de las bases (ϕ), para los mínimos de $dApdT-Na^+$ complementarios, obtenidos con distintas funcionales.</i>	54
4.9.	<i>Empaquetamiento del anillo de azúcar y ángulo glucosídico de torsión, χ, entre los mínimos $dApdT-Na^+$ complementarios, obtenidos con distintas funcionales.</i>	55

4.10. <i>Parámetros de las bases y pares de bases de los dApdT complementarios, minimizados con disintas funcionales (véase designación en la figura 2.7).</i>	56
4.11. <i>Ángulos de torsión del esqueleto azúcar-fosfato y valores promedio del ángulo entre los planos de las bases (ϕ), para los mínimos de dTpdA-Na+ complementarios, obtenidos con MP2 y DFT.</i>	57
4.12. <i>Empaquetamiento del anillo de azúcar y ángulo glucosídico de torsión, χ, para los mínimos de dTpdA-Na+ complementarios, obtenidos con MP2 y DFT.</i>	59
4.13. <i>Parámetros de las bases y pares de bases de los dTpdA complementarios, minimizados con MP2 y DFT (véase designación en la figura 2.7).</i>	60
4.14. <i>Ángulos de torsión del esqueleto azúcar-fosfato (véase designación en la figura 2.5) y valores promedio del ángulo entre los planos de las bases (ϕ), para los mínimos de dGpdC-Na+ complementarios, obtenidos con distintos métodos.</i>	61
4.15. <i>Empaquetamiento del anillo de azúcar y el ángulo glucosídico de torsión, χ, para los mínimos de dGpdC-Na+ complementarios, obtenidos con distintos métodos.</i>	62
4.16. <i>Parámetros de las bases y pares de bases de los dGpdC complementarios, minimizados con distintos métodos (véase designación en la figura 2.7).</i>	63

Capítulo 1

Introducción

Un hombre honrado, armado con todo el conocimiento a nuestra disposición ahora, solo podría declarar que, en algún sentido, por el momento el origen de la vida parece casi milagroso

FRANCIS CRICK,

ganador del Premio Nobel, codescubridor de la estructura de hélice doble del ADN.

1.1. Breve historia sobre el ADN

El ADN es una macromolécula descubierta hace más tiempo del que podríamos pensar. Lo descubrió, en 1869 ¹, Johann Friedrich Miescher, un científico suizo que trabajaba en la Universidad de Tubinga, en Alemania. Cuando examinaba pus de vendajes quirúrgicos, al microscopio, encontró una sustancia que no identificó y a la que llamó nucleína (porque residía en los núcleos de las células). Miescher hizo por entonces poco más que reseñar su existencia, pero veintitrés años después, en una carta suya a un tío, planteó la posibilidad de que aquéllas moléculas pudieran ser los agentes básicos de la herencia[1].

Durante la mayor parte del medio siglo siguiente, el criterio general era que aquel material (llamado ácido desoxirribonucleico o ADN) tenía como máximo un papel secundario en cuestiones de herencia. Era demasiado simple. No tenía más que cuatro componentes básicos llamados nucleótidos, lo que era como tener un alfabeto de sólo cuatro letras. ¿Cómo iba a poder escribir la historia de la vida con un alfabeto tan rudimentario? La respuesta es que lo hace de un modo muy parecido al que utilizas para crear mensajes complejos con simples puntos y líneas del

¹Ridley,Matt, *Genome: The Autobiography of a Species*.London: Fourth Estate, 1999, p. 48.

código Morse: combinándolos. El ADN no hacía nada en absoluto, que se supiera². Estaba simplemente instalado allí en el núcleo, tal vez ligando al cromosoma de algún modo, añadiendo una pizca de acidez cuando se lo mandaban o realizando alguna otra tarea trivial que aún no se le había ocurrido a nadie. Se creía que la complejidad necesaria tenía que estar en las proteínas del núcleo³.

Sin embargo, dos problemas planteaban el hecho de desdeñar el ADN. Primero, había tanto (casi dos metros en cada núcleo) que era evidente que las células lo necesitaban y que tenía que ser importante para ellas. Además, en dos estudios en particular, uno sobre la bacteria *Pneumococcus* y otro sobre bacteriógrafos (virus que infectan bacterias), el ADN demostró tener una importancia que sólo se podía explicar si su papel era más trascendental. Las pruebas parecían indicar que el ADN participaba de algún modo en la formación de proteínas, un proceso decisivo para la vida, pero estaba también claro que las proteínas se hacían *fuera* del núcleo, bien lejos del ADN que dirigía supuestamente su ensamblaje.

Nadie podía entender cómo era posible que el ADN enviase mensajes a las proteínas. Hoy sabemos que la respuesta es el ARN, o ácido ribonucleico, que actúa como intérprete entre los dos. Es una notable rareza de la biología que el ADN y las proteínas no hablen el mismo idioma, para comunicarse necesitan un mediador en la forma de ARN. El ARN, trabajando con una especie de empleado químico llamado ribosoma, les traduce a las proteínas la información del ADN de una célula a un lenguaje comprensible para que ellas puedan utilizarlas, es decir, que se sinteticen las proteínas.

Reanudando nuestra historia, en 1904, Thomas Hunt Morgan, justo cuatro años antes del oportuno redescubrimiento de los experimentos de Mendel con plantas de guisantes y todavía casi una década antes de que “*gen*” se convirtiera siquiera en una palabra, empezó a hacer cosas notablemente nuevas con los cromosomas. Los cromosomas se habían descubierto por casualidad en 1888 y se llamaron así porque absorbían enseguida la tinción y eran fáciles de ver al microscopio.

Morgan eligió como objeto de estudio una mosca pequeña y delicada, llamada oficialmente *Drosophila melanogaster*, conocida como la mosca de la fruta. Morgan y su equipo trabajaban en un pequeño laboratorio de Schermerhorn⁴, en la universidad de Columbia, Nueva York. Intentaron, durante seis años, producir mutaciones por todos los medios que se les ocurrieron (exponiendo a las moscas a radiación, criándolas en un medio de luz brillante y de oscuridad, haciéndolas girar en

²Wallace *et al.*, *Biology: The Science of Life*, 2nd edn. Glenview, 1986, p. 48.

³De Duve, Christian, *A Guided Tour of the Living Cell*, vol. 2, 1984, p. 295.

⁴Clark, Ronald W., *The Survival of Charles Darwin: A Biography of a Man and an Idea*, London: Daedalus Book, 1985, p. 259

centrifugadoras), pero nada resultó. Cuando Morgan estaba ya a punto de renunciar se produjo una mutación súbita y repetible: una mosca tenía los ojos de color blanco en vez del rojo habitual. Después de este éxito, Morgan y sus ayudantes consiguieron generar provechosas deformidades que les permitieron rastrear un rasgo a lo largo de sucesivas generaciones. Pudieron determinar así las correlaciones entre caracteres particulares y cromosomas individuales, llegando a demostrar, por fin, que los cromosomas eran una de las claves de la herencia.

Pero seguía en pie el problema en el nivel siguiente de la herencia: los enigmáticos genes y el ADN del que se componían.

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.

Que Wilkins y Franklin no congeniasen fue un hecho que Watson y Crick parece ser que explotaron en beneficio propio. Tras un “problema” que tuvieron⁵, en enero de 1953, Wilkins mostró a Watson las imágenes de Franklin, al parecer sin que ella lo supiese ni lo consintiese⁶ (figura 1.1). 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. A partir de ahí construyeron un modelo tipo Mecano, que consistía en placas metálicas atornilladas en espiral.

⁵White, *Rivals*, p.257; Maddox, Brenda, *Rosalind Franklin: The Dark Lady of DNA*, London: HarperCollins, 2002, p. 185.

⁶Portal internet PBS, «A Science Odyssey», n.d.

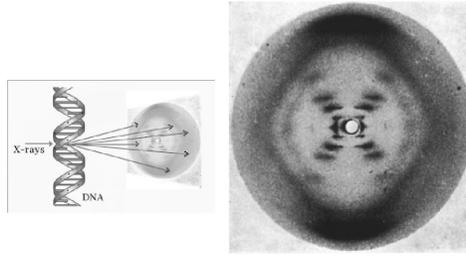


Figura 1.1: “Foto 51”, tomada por Franklin, que muestra un patrón de difracción de rayos X de ADN de la forma B

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 desoxirribonucleico»⁷. 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»⁸. En realidad casi no se 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 más⁹, 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.

⁷De Duve, *A Guided Tour of the Living Cell*, vol. 2, p.290

⁸Crick, Francis, *What Mad Pursuit: A Personal View of Scientific Discovery*, London: Penguin Press, p. 73-74

⁹*National Geographic*, «Secrets of the Gene», octubre de 1995, p. 55.

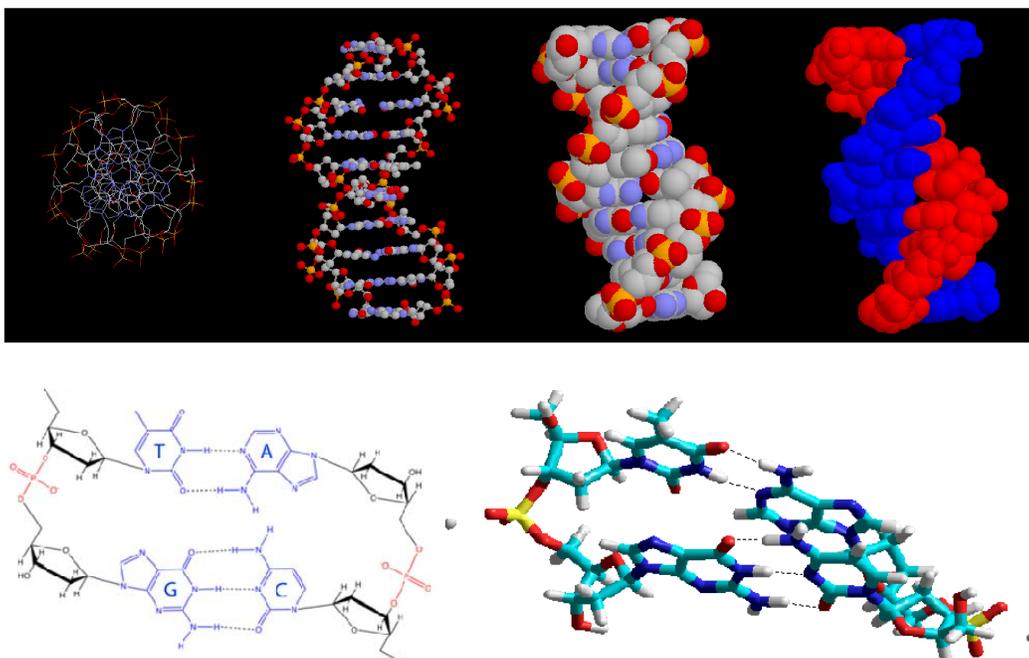


Figura 1.2: Estructura de la hélice doble del ADN. (ARRIBA) Fragmento de ADN de la familia B1, vista aérea (imagen a la izq.) y frontalmente con diferentes efectos. (ABAJO) Fragmento mínimo de ADN, ejemplo de dDMPs complementarios, con los pares de bases Adenina-Timina y Guanina-Citosina.

1.2. Importancia biológica y estudios del ADN

El ADN es la molécula biológica más importante, ya que es la portadora de la información genética. Toda la información necesaria para que se lleve a cabo la vida y reproducción en los sistemas biológicos está contenida en ésta macromolécula. El ADN es una hélice doble, consta de dos cadenas complementarias de polinucleótidos, cada una de las cuales está compuesta por un conjunto de nucleósidos conectados por un grupo fosfato (figura 1.2). Cada nucleósido consiste de una base nitrogenada (Adenina, Timina, Guanina o Citosina) y un azúcar desoxirribosa. A pesar del importantísimo papel biológico del ADN, la diversidad química de ésta es sorprendentemente simple y está limitada a la combinación de las cuatro posibles bases nitrogenadas a lo largo de la cadena de polinucleótidos. Sin embargo, esto es suficiente para codificar la gran diversidad de la vida. Pero, ¿por qué sólo cuatro bases nitrogenadas para escribir la gran diversidad de vida?, pues bien, porque de no ser así no podríamos tener la maravillosa cadena alimenticia que tenemos, donde podemos sintetizar los compuestos de otros organismos para poder sobrevivir.

A pesar de su estructura 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.

Han sido revelados, en experimentos estructurales, varios patrones de la estructura tridimensional de los fragmentos de ADN, incluyendo las bien conocidas familias de dobles hélices (A y B) y conformaciones con secuencias específicas (familia Z). Ésta importante información estructural obtenida ha sido acumulada en el *Nucleic Acid Data Bank* [2] y ha sido analizada por Svoil y col. [3] Las familias y los fragmentos específicos de las secuencias de polinucleósidos dentro de cada familia se distinguen por los ángulos de torsión del esqueleto azúcar-fosfato y por el arreglo mutuo de las bases.

Sin embargo, determinar el papel de los fragmentos individuales en la conformación dúplex va más allá del alcance de las técnicas experimentales disponibles, por ello es indispensable el uso de herramientas teórico-computacionales, como cálculos de mecánica cuántica, para entender las complejas propiedades de los bloques de construcción y la naturaleza física de sus interacciones. Debido al alto costo computacional de éstos métodos, es necesario limitar los sistemas de estudio hasta el mínimo posible.

Para evaluar la contribución de las conformaciones preferentes presentes en una sola cadena de ADN, respecto a la variación de la estructura de la doble hélice, se han realizado con anterioridad cálculos de mecánica cuántica de primeros principios de fragmentos mínimos de una sola cadena de ADN, usando el método de la Teoría de Funcional de Densidad (DFT, por sus siglas en inglés) [4, 5]. Este nivel de la teoría provee un compromiso razonable entre la exactitud de los cálculos y la demanda de recursos de cómputo, en comparación con los métodos computacionales, no empíricos, disponibles. Los cálculos específicos con DFT han revelado que la conformación característica BI de la cadena de ADN está inicialmente predispuesta en una sola hebra y concretamente en los perfiles estructurales de su dDMP en la forma de mínimo local de energía [6].

Siendo la unidad estructural más pequeña de una sola cadena de ADN, los desoxydinucleósidos monofosfatos (dDMPs), (figura 1.3), llevan cierta información sobre determinada conformación local de ADN. Éstos consisten de dos nucleósidos situados a lo largo de la hebra de ADN conectados por un grupo fosfato. Los dDMPs pueden ser descritos totalmente por los ángulos de torsión del esqueleto

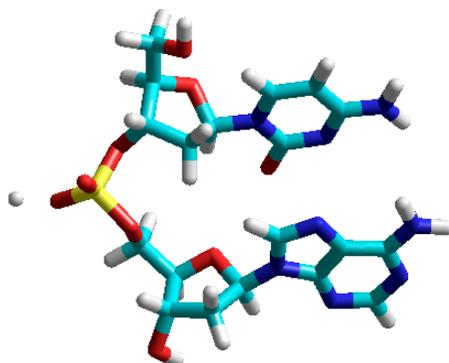


Figura 1.3: *Desoxydinucleósido Monofosfato neutralizado con Na⁺, fragmento mínimo de la cadena única del ADN.*

azúcar-fosfato, el empaquetamiento de los anillos de la desoxirribosa, y los ángulos de torsión glucosídicos, uno por base.

Trabajos anteriores [4, 5, 6, 7] mostraron que la contribución del esqueleto azúcar-fosfato al paralelismo de las bases es dependiente de la secuencia de las bases, 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.

Las interacciones debido al apilamiento o también llamado *Stacking* y los enlaces por puente de hidrógeno de las bases del ADN representan una importante fuente de variabilidad conformacional de este mismo, pues éstas interacciones son las encargadas de la estabilización de la hélice doble. Datos experimentales del ADN cristalino se han utilizado ampliamente para examinar cómo las interacciones de *stacking* y la orientación de la cadena principal de azúcar-fosfato se correlacionan con parámetros estructurales promedio, de secuencias de nucleótidos, tales como *Twist*, *Rise*, *Propeller*, *Roll*, y *Slide*. Éstas investigaciones han revelado que desviaciones en éstos parámetros son dependientes de la secuencia de bases[s]. Algunos estudios indican que las variaciones en el parámetro *Twist* del ADN pueden ser atribuidas a las interacciones de *stacking* que involucran repulsiones estéricas entre los radicales de los anillos de las bases en surcos mayores y menores de ADN y las interacciones $\pi - \pi$ entre pares de bases apiladas de ácidos nucleicos, mientras que un mayor análisis de datos de rayos X, sugiere que el esqueleto también juega un papel significativo en la determinación del ángulo de torsión helicoidal (*h-Twist*) del ADN.

A fin de continuar los estudios sobre el polimorfismo del ADN, en el presente trabajo estudiamos diez fragmentos mínimos de la hélice doble del ADN con diferentes secuencias de bases, los cuales están conformados por dDMPs complementarios, de la familia estructural B1. Éstos fueron sometidos a minimización utilizando el método DFT, para analizar las características comunes, diferencias y dependencias de secuencias de nucleósidos, pues se sabe muy bien que el arreglo mutual de las bases en los dDMPs es dependiente de la secuencia, es decir, dos dDMPs que contienen los mismos nucleósidos pero en distinta secuencia presentan una diferencia drástica en las posiciones mutuales de las bases. Además, como un estudio extra, se analizaron tres fragmentos mínimos, con las secuencias A-T (Adenina-Timina), T-A (Timina-Adenina) y G-C (Guanina-Citosina), sometidos a minimización por los métodos MP2 y el Funcional PWB6K para el análisis de diferencias en las características conformacionales entre las estructuras con la misma secuencia de nucleósidos, ésto con el fin de comparar los diferentes métodos.

Cabe resaltar que éste trabajo es producto de la colaboración entre diferentes investigadores y estudiantes. El papel principal de la tesista en tal colaboración fué la de crear archivos iniciales a partir de estructuras experimentales (previamente optimizadas por distintos métodos), medir las características estructurales de los dDMPs complementarios ya optimizados y organizar estos resultados en gráficos y tablas. En la presente tesis se describe el método de Teoría del Funcional de Densidad para comprender la estructura electrónica.

Capítulo 2

Estructura del ADN

«Dime y lo olvido, enséñame y lo recuerdo,
involúcrame y lo aprendo.»

BENJAMIN FRANKLIN,
político, científico e inventor.

Los ácidos nucleicos constituyen los últimos peldaños en la jerarquía molecular de los seres vivos, por ser depositarios de la información genética, cuyas instrucciones vienen codificadas por el ordenamiento 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. Por otra parte, su capacidad de replicarse permite la obtención de copias exactas del código genético con las que dota a las células descendientes, lo que supone la base para la perpetuación de la vida.

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 un polinucleótido de doble cadena cuya misión es conservar la información genética, especificando la secuencia de aminoácidos de todas y cada una de las proteínas celulares. Cada nucleótido está formado por tres unidades (figura 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: dos púricas (Adenina y Guanina) y dos pirimidínicas (Timina y Citosina). Cabe mencionar que un nucleósido sólo está formado por la desoxirribosa y una de las cuatro bases nitrogenadas, así, *Nucleósido + Grupo Fosfato = Nucleótido*.

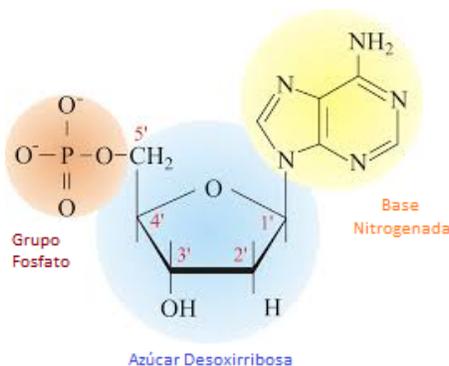


Figura 2.1: Estructura de un nucleótido

BASE NITROGENADA	NUCLEÓSIDO	NUCLEÓTIDO
Adenina	Adenosina	Ácido Adenílico
Guanina	Guanidina	Ácido Guanílico
Citosina	Citidina	Ácido Citidílico
Timina	Timidina	Ácido Timidílico

Tabla 2.1: Terminología empleada para referirse a los nucleósidos y nucleótidos.

Los nucleótidos de cada una de las dos cadenas que forman el ADN establecen una asociación específica con los correspondientes de la otra cadena. Debido a la afinidad química entre las bases, los nucleótidos que contienen Adenina forman puentes de hidrógeno (enlaces químicos débiles) siempre con los que contienen Timina, y los que contienen Citosina con los que contienen Guanina (figura 2.2).

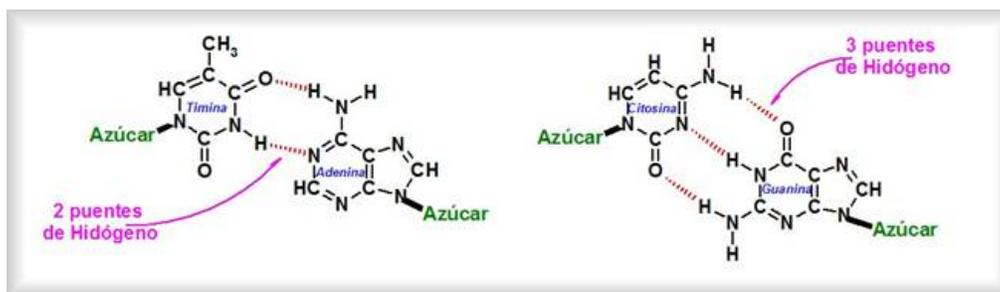


Figura 2.2: Bases nitrogenadas y pares que se forman.

Una de las tareas más importantes de la célula es la síntesis de proteínas, moléculas que intervienen en la mayoría de las funciones celulares, estas 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.

De las dos cadenas de polinucleótidos que forman una molécula de ADN, sólo una, llamada *cadena adelantada*, contiene la información necesaria para la producción de una determinada secuencia de aminoácidos. La otra, llamada *cadena retrasada*, ayuda a la replicación.

Cabe mencionar que un gen es una secuencia de nucleótidos de ADN que especifica el orden de aminoácidos de una proteína por medio de una molécula intermedia 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.

El ADN se puede presentar en cuatro niveles de organización (figura 2.3):

- **Estructura Primaria:** Cadena única del ADN, tiene cada uno de sus nucleósidos unidos por un *enlace fosfodiéster* (enlace covalente que se produce entre un grupo hidroxilo ($-OH$) en el carbono $3'$ y un grupo fosfato (PO_4) en el carbono $5'$ del nucleótido entrante). Aquí se encuentra la información genética, radicando en la secuencia de las bases nitrogenadas.
- **Estructura Secundaria:** Postulada por Watson y Crick, es una hélice doble, dextrógira o levógira, según la conformación de ADN. Permite explicar el almacenamiento de la información genética y el mecanismo de duplicación del ADN. Ambas cadenas son complementarias y antiparalelas, debido a la complementariedad de las bases para formar enlaces de hidrógeno y debido a que el extremo $3'$ de una se enfrenta al extremo $5'$ de la otra.
- **Estructura Terciaria:** Se refiere a cómo se almacena el ADN en un volumen reducido para formar los cromosomas. Varía según se trate de organismos procariontes o eucariontes:
 - a) En procariontes se pliega como una super-hélice en forma, generalmente, circular y asociada a una pequeña cantidad de proteínas. Lo mismo ocurre en las mitocondrias y en los plastos.

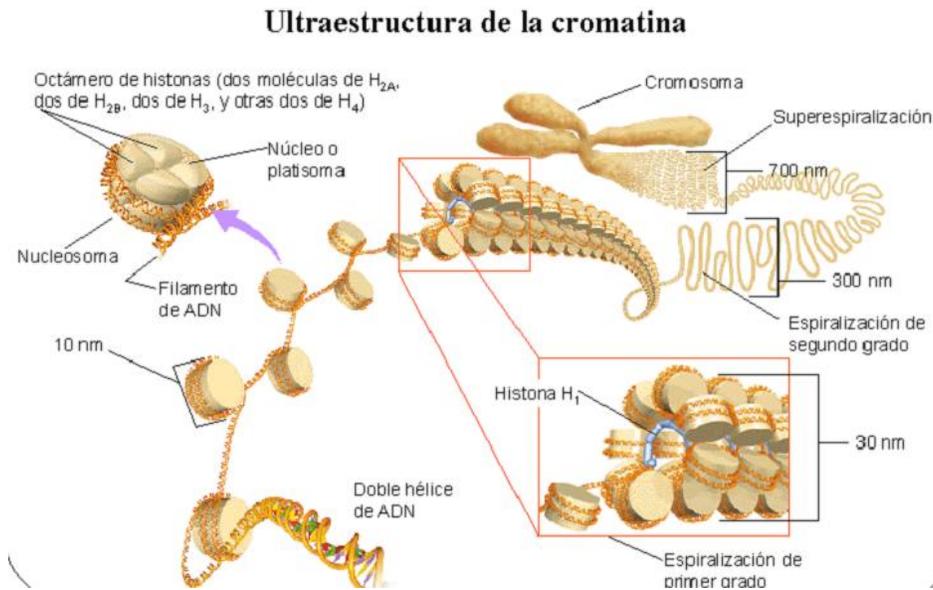


Figura 2.3: Estructura Secundaria, Terciaria y Cuaternaria del ADN

b) En eucariontes el empaquetamiento ha de ser más complejo y compacto y para esto necesita la presencia de proteínas. A esta unión de ADN y proteínas se conoce como cromatina, en la cual se distinguen diferentes niveles de organización:

- Nucleosoma
 - Collar de perlas
 - Fibra cromatínica
 - Bucles radiales
 - Cromosoma
- **Estructura Cuaternaria:** Sólo esta presente si hay más de una cadena polipeptídica. La cromatina en el núcleo tiene un grosor de 300Å. La fibra de cromatina de 100Å se empaqueta formando una fibra de cromatina de 300Å. El enrollamiento que sufre el conjunto de nucleosomas recibe el nombre de solenoide. Los solenoides se enrollan formando la cromatina del núcleo interfásico de la célula eucariota. Cuando la célula entra en división, el ADN se compacta aún más, formando los cromosomas.

Dado que en el presente trabajo nuestro objeto de estudio son fragmentos de la doble hélice del ADN (Estructura Secundaria), hablaremos un poco más a fondo de ésta.

2.1. Estructura Secundaria del ADN

Experimentalmente, gracias a la difracción de rayos X para fragmentos de ADN cristalizados, se ha podido clasificar a las estructuras (secundarias) de ADN en 3 grandes familias conformacionales (figura 2.4): familia A, familia B (ambas Dextrógiras) y la familia Z (Levógira). Cabe mencionar que el ADN de tipo B es el más abundante y es el descubierto por Watson y Crick.

Familias estructurales del ADN [10]:

- **ADN-A:** Tenemos ésta familia conformacional cuando el ADN, (fibra de ADN), se encuentra en un medio con escasa agua. Éste es una gran hélice, rechoncha con bases fuertemente inclinadas, dejando de ser perpendiculares al eje de la hélice, 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, con planos de las bases esencialmente 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 A-T de los T-A, 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...). 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 zigzag que es responsable de la designación Z.

Cabe mencionar que en el presente trabajo el estudio se limita a la familia estructural B, específicamente BI.

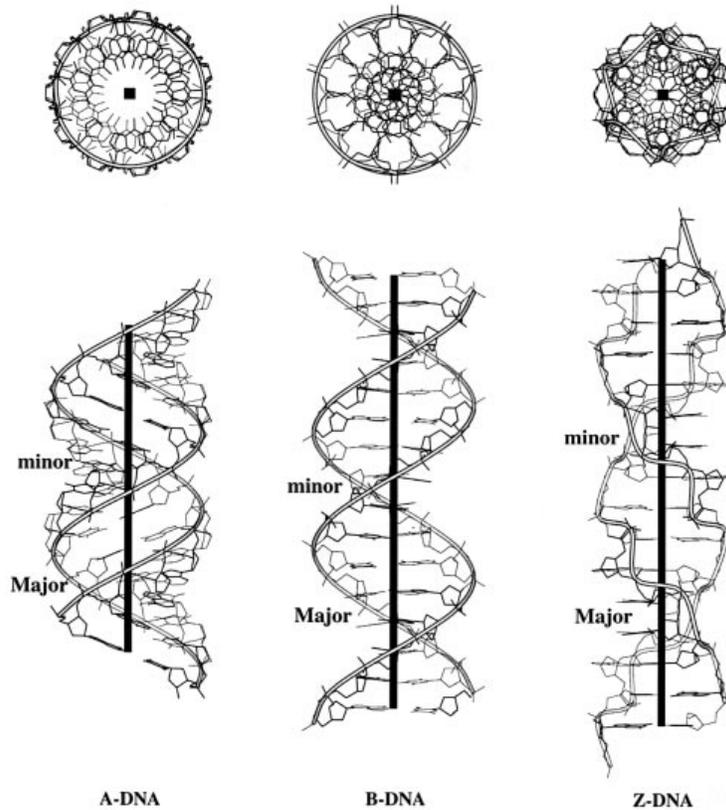


Figura 2.4: Vistas laterales y superiores que ilustran los rasgos característicos de las estructuras helicoidales regulares A, B y ADN-Z deducidas a partir de los modelos de difracción de rayos X de fibras representativas. Las cintas rastrean la progresión de la columna vertebral definida por los átomos de fósforo y las líneas negras gruesas (cajas) representan los ejes helicoidales [9].

Pero, ¿cuál es la importancia de estudiar dichas familias estructurales?. Si bien no es posible estudiar al ADN en pleno funcionamiento, dentro de las células, dado que sabemos de qué están conformadas y conocemos sus distintos procesos, podemos darnos una idea del papel que puede jugar cada familia de ADN. Por ejemplo, el ADN-A dado sus grupos hidrofóbicos podemos encontrarlo en los lípidos, por ejemplo en las membranas..., por otro lado el ADN-B que es el más abundante y se observa en solución con agua, podemos tenerlo en general en las células, del ADN-Z no se sabe mucho aún, pero hay indicaciones de que existen estructuras semejantes en los organismos, por ello la importancia de estudiarlo también.

2.1.1. Parámetros geométricos para el estudio del ADN

Para el estudio del ADN nos podemos valer de distintos parámetros que nos aportan información sobre éste, en este trabajo analizamos los ángulos de torsión (Fig. 2.5), el empaquetamiento del azúcar (Fig. 2.6), además de los parámetros *h-Twist*, *h-Rise*, *Buckle*, *Propeller* y superposición de las bases debido al *Stacking* (Fig. 2.7).

Para los ácidos nucleicos son especificados seis ángulos principales de torsión de la cadena azúcar-fosfato (figura 2.5). Uno de estos ángulos de torsión, δ , es particularmente informativo porque 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. 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 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 2.6). 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 [10].

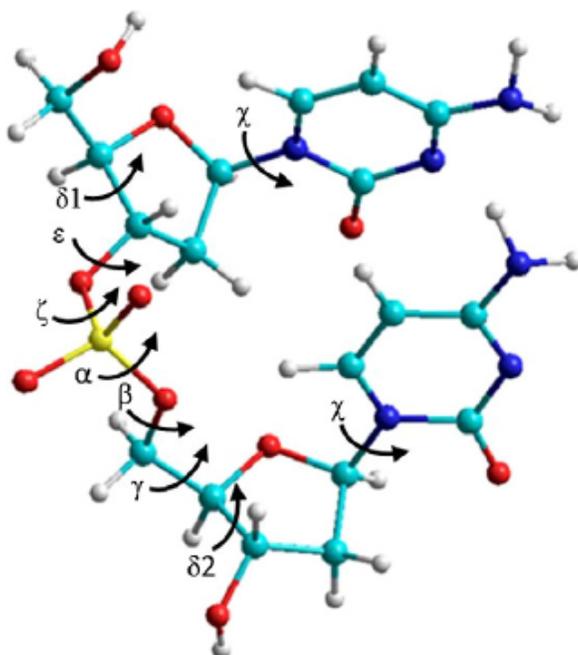


Figura 2.5: Designación de los ángulos de torsión de un dDMP.

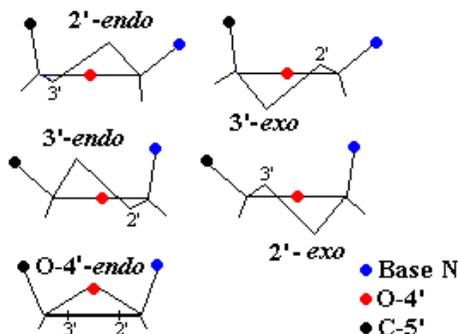


Figura 2.6: *Empaquetamiento del azúcar.*

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 2.5). 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 zigzag 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 sólo se encuentra la conformación *anti*.

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

Para cada conjunto de ángulos tenemos ciertos parámetros de posición de las bases, cada familia tiene dichos ángulos en regiones específicas, lo cual nos permite determinar a qué tipo de familia pertenece cierto fragmento de ADN, es por ello que en el presente trabajo se estudiaron los parámetros *Buckle*, *Propeller* y la superposición de las bases debida al *Stacking*, además de *h-Twist* y *h-Rise* para las diez posibles conformaciones de dinucleósidos monofosfatos, como unidad mínima de la hélice doble de ADN.

Los enlaces de hidrógeno y *Stacking* de las bases nitrogenadas son fundamentales para la organización del ADN y ARN, dando a ambos sus estructuras tridimensionales y modos de reconocimiento. Cabe mencionar que los pares de bases no son necesariamente planos, lo cual podemos comprobarlo con el ángulo φ' . El *Stacking*

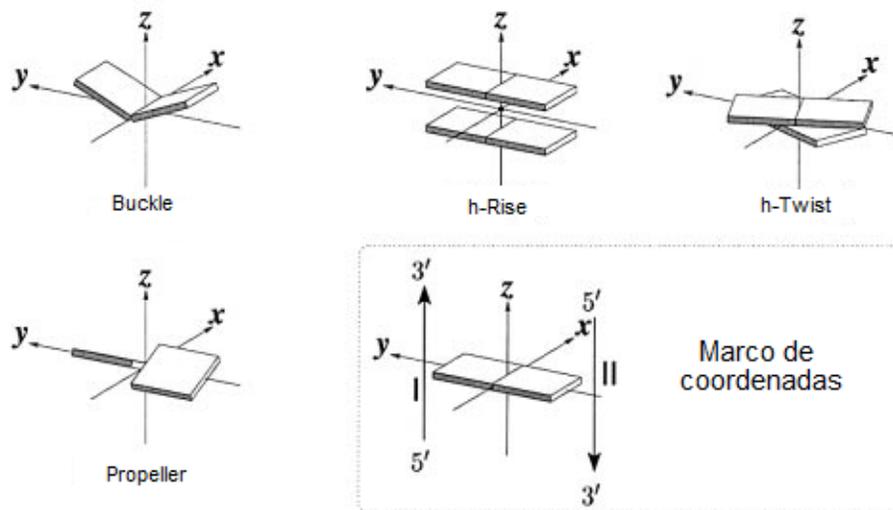


Figura 2.7: Descripciones gráficas de los parámetros *Buckle*, *Propeller*, *h-Rise* y *h-Twist*, empleados para describir la geometría de pares de bases complementarios y secuencias de pares de bases.

cara a cara de los anillos aromáticos depende sutilmente de la secuencia de nucleótidos. Comprobando que el grado de superposición entre purina-pirimidina es mayor que el que hay entre pirimidina-purina [11].

Cada par de bases se caracteriza de forma única por un conjunto de seis parámetros de cuerpo rígido: *Shear*, *Stretch*, *Stagger*, *Buckle*, *Propeller* y *Opening*. Ya que los seis parámetros de pares de bases definen unívocamente la posición y orientación relativa de dos bases, se pueden utilizar para reconstruir el par de bases. Por otra parte, los parámetros proporcionan un mecanismo simple para la clasificación de las estructuras y la búsqueda en bases de datos. Entre los seis parámetros del par bases, sólo *Shear*, *Stretch* y *Opening* son críticos en la caracterización clave de los enlaces de hidrógeno. *Buckle*, *Propeller* (ver Fig. 2.7, a la izquierda) y *Stagger*, en contraste, son parámetros secundarios, que simplemente describen las imperfecciones, es decir, la no planaridad de determinado par de bases [9]. 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° . Los otros cuatro muestran mucho menos variaciones.

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 y Twist, que describen la geometría del apilamiento de una secuencia de dinucleótido desde una perspectiva local; y el conjunto de parámetros helicoidales, X-desplazamiento, Y-desplazamiento, Inclination, Tip *helical Rise* y *helical Twist* (ver Fig. 2.7, a la derecha), que describen la regularidad de la hélice, por ejemplo, *h-Twist* es el ángulo de rotación alrededor del eje helicoidal que tienen pares de bases sucesivos. Generalmente se supone que la longitud del ADN es el mismo para cualquier familia de ADN del mismo tamaño y que el valor de *h-Rise* de cualquier secuencia de dinucleótido no cambia en gran medida del valor promedio 3.4\AA . Resultados más recientes, 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 [9].

Capítulo 3

Metodología para el estudio Teórico-Computacional del ADN

Nunca consideres el estudio como una obligación, sino como una oportunidad para penetrar en el bello y maravilloso mundo del saber.

ALBERT EINSTEIN,
ganador del Premio Nobel de Física en 1921.

Es indiscutible que el ADN juega un papel esencial en los procesos de la vida, gracias a su riqueza estructural. Pero, ¿cuáles son los factores que permiten dicha riqueza estructural?. Para poder responder ésta 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 y resonancia magnética nuclear, 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 trabajamos no son tan complejos, es posible hacer uso de ambos. En este trabajo se empleó el modelo mecánico cuántico para optimizar los dDMPs complementarios (que se detallan en

la Sección 3.2.1), como unidad mínima de la hélice doble de ADN.

3.1. Mecánica Cuántica

El Método de Mecánica Cuántica 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 son estrictamente matemáticas, pues buscan emplear una forma más simple de las funciones utilizadas para obtener soluciones aproximadas a las ecuaciones diferenciales complicadas.

Una característica más de los métodos *ab initio*, es resolver la estructura electrónica de la molécula, resolviendo la ecuación de Schrödinger independiente del tiempo para cada átomo o partícula:

$$\hat{H}\Psi = E\Psi \tag{3.1}$$

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.

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

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 (LCAO-Linear Combination of Atomic Orbitals) del sistema.

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, sino que 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, debe ser anti-simétricos.

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 de perturbaciones Moller-Plesset, conocido como *Método de Perturbaciones* (MP). El cual incluye el efecto de correlación electrónica al aumentar el Hamiltoniano no perturbado H_0 de HF una pequeña perturbación H' :

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}' \quad (3.2)$$

Donde λ es un parámetro arbitrario.

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:

$$\left[-\frac{\hbar^2}{8\pi^2m} \nabla^2 + V \right] \psi(r, t) = \frac{i\hbar}{2\pi} \frac{\partial \psi(r, t)}{\partial t} \quad (3.3)$$

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.

$$\psi(r, t) = \psi(r)\tau(t) \quad (3.4)$$

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:

$$H\psi(r) = E\psi(r) \quad (3.5)$$

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

$$H = -\frac{\hbar^2}{8\pi^2m}\nabla^2 + V \quad (3.6)$$

Las soluciones a la Ec. 3.5 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.

$$\psi_{total} = \psi_{ele}\psi_{nuc} \quad (3.7)$$

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:

$$H^{elec}\psi^{elec}(r, R) = E^{eff}\psi^{elec}(r, R) \quad (3.8)$$

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.

$$H^{nucl} = T^{nucl}(R) + E^{eff}(R) \quad (3.9)$$

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, 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 nuestro trabajo, dicho sistema son los dinucleósidos monofosfatos complementarios, 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.

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:

$$\hat{H}\Psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots) = E\Psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots) \quad (3.10)$$

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:

$$\hat{H} = \hat{T} + \hat{V} + \hat{V}_{ee} \quad (3.11)$$

Donde el operador de energía cinética es:

$$\hat{T} = \sum_{i=q}^N \frac{-\nabla_i^2}{\hat{V}_{ee}} \quad (3.12)$$

La energía potencial de la interacción núcleo-electrón es:

$$\hat{V} = \sum_i^N v\vec{r}_i = - \sum_{i=1}^N \sum_{\alpha} \frac{Z_{\alpha}}{r_{\alpha i}} \quad (3.13)$$

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:

$$\hat{V}_{ee} = \sum_{i < j}^N \left(\frac{1}{r_{ij}} \right) \quad (3.14)$$

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 $\rho(\mathbf{r})$:

$$\rho(\vec{r}_i) = \sum_{\sigma=\uparrow\downarrow} \int |\psi(\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots)| dx_1 dx_2 \dots dx_N \quad (3.15)$$

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.

$$\hat{V} \leftrightarrow \psi \leftrightarrow \rho \quad (3.16)$$

Comprobemos esta relación, en primer lugar notemos que dado el potencial V es posible, usando la ecuación de Schrödinger, encontrar la función de onda ψ ; así, con la función de onda y empleando la Ec. 3.15 podemos encontrar la densidad electrónica. Es decir:

$$V \rightarrow \psi \rightarrow \rho \quad (3.17)$$

Ahora analicemos si se satisface $V \leftarrow \psi$. Como ya hemos probado que $V \rightarrow \psi$ sólo resta verificar que para dos potenciales, V y V' , diferentes se tendrán dos funciones de onda, ψ y ψ' , respectivamente. *"Si V y 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:

$$(\hat{T} + \hat{V} + \hat{V}_{ee}) |\psi\rangle = E_{gs} |\psi\rangle \quad (3.18)$$

Y

$$(\hat{T} + \hat{V}' + \hat{V}_{ee}) |\psi\rangle = E'_{gs} |\psi\rangle \quad (3.19)$$

Restando las dos ecuaciones anteriores tenemos que:

$$(\hat{V} - \hat{V}') |\psi\rangle = (E_{gs} - E'_{gs}) |\psi\rangle \quad (3.20)$$

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:

$$E_{gs} = \langle \psi | \hat{H} | \psi \rangle < \langle \psi' | \hat{H} | \psi' \rangle = \langle \psi' | \hat{H}' + \hat{V} - \hat{V}' | \psi' \rangle = E'_{gs} + \int \rho'(\vec{r}) [v(\vec{r}) - v'(\vec{r})] \quad (3.21)$$

$$E'_{gs} = \langle \psi' | \hat{H} | \psi' \rangle < \langle \psi | \hat{H}' | \psi \rangle = \langle \psi | \hat{H} + \hat{V}' - \hat{V} | \psi \rangle = E_{gs} + \int \rho(\vec{r}) [v'(\vec{r}) - v(\vec{r})] \quad (3.22)$$

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

$$E_{gs} + E'_{gs} < E_{gs} + E'_{gs} \quad (3.23)$$

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

$$\psi \longleftrightarrow \rho(\vec{r}) \quad (3.24)$$

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:

$$\left(-\frac{\nabla^2}{2} + \hat{V}_{so}(\vec{r}) \right) \phi_i(\vec{r}) = \epsilon_i \phi_i(\vec{r}) \quad (3.25)$$

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

$$\rho_s(\vec{r}) = \sum_{i=1} |\phi_i(\vec{r})| \quad (3.26)$$

Y la energía cinética no interactiva del sistema es:

$$T_s[\rho] = \sum_{i=1}^N \left\langle \phi_i \left| -\frac{\nabla^2}{2} \right| \phi_i \right\rangle \quad (3.27)$$

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

$$E[\rho] = T_s[\rho] + \int V_0(\vec{r})\rho(\vec{r})d\vec{r} + \frac{1}{2} \int \int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}d\vec{r}' + E_{XC}(\rho) \quad (3.28)$$

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

$$E_{XC}[\rho] = F_{HK}[\rho] - T_s\rho - \frac{1}{2} \int \int \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}d\vec{r}' \quad (3.29)$$

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:

$$\frac{\delta E[\rho]}{\delta\rho(\vec{r})} \Big|_{\rho=\rho_0} = 0 \quad (3.30)$$

Así, empleando las Ecuaciones 3.28, 3.29 y 3.30 tenemos:

$$\frac{\delta F_{HK}[\rho]}{\delta\rho(\vec{r})} \Big|_{\rho=\rho_0} = -V_0(\vec{r}) \quad (3.31)$$

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

$$\frac{\delta F_{HK}[\rho]}{\delta\rho(\vec{r})} \Big|_{\rho=\rho_0} = -V_{s,o}(\vec{r}) \quad (3.32)$$

Al resolver la derivada parcial indicada obtenemos:

$$V_{s,o}(\vec{r}) = V_0(\vec{r}) + V_H(\vec{r}) + V_{X,C}(\vec{r}) \quad (3.33)$$

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

$$\left[-\frac{\nabla^2}{2} + V_0(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d(\vec{r}') + \frac{\delta E_{XC}}{\delta\rho(\vec{r})} \right] \phi_i(\vec{r}) = \epsilon_i \phi_i(\vec{r}) \quad (3.34)$$

Donde el potencial de intercambio correlación está dado por:

$$V_{X,C}(\vec{r}) = \frac{\delta E_{XC}}{\delta\rho(\vec{r})} \quad (3.35)$$

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

$$V_{S,O}(\vec{r}) = V_0(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \frac{\delta E_{XC}}{\delta \rho(\vec{r})} \quad (3.36)$$

$$V_{X,C}(\vec{r}) = \frac{\delta E_{XC}}{\delta \rho(\vec{r})} \quad (3.37)$$

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.

3.1.2. Teoría de perturbaciones Møller-Plesset de segundo orden (MP2)

La ecuación de Schrödiger para el estado perturbado Ψ_n es:

$$\hat{H}\Psi_n = (\hat{H}^{(0)} + \lambda\hat{V})\Psi_n = E_n\Psi_n \quad (3.38)$$

donde λ es un parámetro real arbitrario, \hat{V} es una perturbación del Hamiltonio imperturbable $\hat{H}^{(0)}$, $n = 1, 2, 3 \dots$ denota diferentes estados discretos.

Pero las expresiones producidas por la *teoría de perturbaciones* no son exactas. Los resultados exactos se pueden obtener siempre y cuando el parámetro de expansión λ sea muy pequeño.

Considerando eso, se hace una expansión en serie de Taylor, en potencias de λ , de Ψ_n y E_n . La ecuación de eigenvalores es entonces:

$$(\hat{H}^{(0)} + \lambda\hat{V})(\sum_{i=0} \lambda^i \Psi_n^{(i)}) = (\sum_{i=0} \lambda^i E_n^{(i)})(\sum_{i=0} \lambda^i \Psi_n^{(i)}) \quad (3.39)$$

Escribiendo sólo los primeros términos se obtiene:

$$(\hat{H}^{(0)} + \lambda\hat{V})(\Psi_n^{(0)} + \lambda\Psi_n^{(1)}) = (E_n^{(0)} + \lambda E_n^{(1)})(\Psi_n^{(0)} + \lambda\Psi_n^{(1)}) \quad (3.40)$$

La ecuación de orden cero es simplemente la ecuación de Schrödinger para el sistema no perturbado:

$$\hat{H}^{(0)}\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(0)} \quad (3.41)$$

Los términos de primer orden son los que tienen como coeficiente a λ :

$$\hat{H}^{(0)}\Psi_n^{(1)} + \hat{V}\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(1)} + E_n^{(1)}\Psi_n^{(0)} \quad (3.42)$$

Si multiplicamos esta ecuación por $\Psi_n^{(0)*}$ e integramos, el primer término en el lado izquierdo se cancela con el primer término del lado derecho ($\hat{H}^{(0)}$, que es hermitiano). Esto conduce a un cambio de primer orden en la energía:

$$E_n^{(1)} = \langle \Psi_n^{(0)} | V | \Psi_n^{(0)} \rangle \quad (3.43)$$

Podemos ver que la perturbación es aplicada, pero se mantiene el estado cuántico $|\Psi_n^{(0)}\rangle$, el cual es un estado cuántico válido aunque no es un estado propio de energía. La perturbación provoca que la energía promedio de este estado incrementa en $\langle \Psi_n^{(0)} | V | \Psi_n^{(0)} \rangle$. El verdadero cambio de energía es ligeramente diferente, ya que el eigenestado perturbado no es exactamente el mismo que $|\Psi_n^{(0)}\rangle$. Cambios adicionales se dan por correcciones de segundo orden en la energía.

Para obtener la corrección de primer orden en el estado propio de energía, recordemos la Ec. 3.40 y multipliquémosla por $\Psi_n^{(0)*}$, ($m \neq n$) a la izquierda e integremos. Se obtiene:

$$(E_m^{(0)} - E_n^{(0)}) \langle \Psi_m^{(0)} | \Psi_n^{(1)} \rangle = - \langle \Psi_m^{(0)} | V | \Psi_n^{(0)} \rangle \quad (3.44)$$

Expandamos $|\Psi_n^{(1)}\rangle$ como:

$$|\Psi_n^{(1)}\rangle = \sum_m a_m |\Psi_m^{(0)}\rangle \quad (3.45)$$

Así, combinando las dos ecuaciones anteriores se obtiene:

$$|\Psi_n^{(1)}\rangle = \sum_{m \neq n} \frac{\langle \Psi_m^{(0)} | V | \Psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |\Psi_m^{(0)}\rangle \quad (3.46)$$

El cambio de primer orden en la n -eigenfunción de energía tiene una contribución de cada uno de los eigenestados de energía ($m \neq n$).

La corrección de segundo orden a la energía es:

$$|\Psi_n^{(2)}\rangle = \sum_{m \neq n} \frac{|\langle \Psi_m^{(0)} | V | \Psi_n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} \quad (3.47)$$

Las correcciones MP en la energía se obtienen con la perturbación (potencial de correlación):

$$\hat{V} = \hat{H} - \hat{F} - \langle \Phi_0 | \hat{H} - \hat{F} | \Phi_0 \rangle \quad (3.48)$$

donde H es el hamiltoniano electrónico habitual y el determinante de Slater (Φ_0) es la eigenfunción del operador de Fock (\hat{F}).

$$\hat{F}\Phi_0 = 2\left(\sum_{i=1}^{N/2} \epsilon_i\right)\Phi_0 \quad (3.49)$$

donde N es el número de electrones de la molécula a considerar, ϵ_i la energía orbital que pertenece al orbital espacial doblemente ocupado. El operador de Fock cambiado es:

$$\hat{H}_0 = \hat{F} + \langle \Phi_0 | \hat{H} - \hat{F} | \Phi_0 \rangle \quad (3.50)$$

sirve como operador imperturbable (de orden cero). El determinante de Slater (Φ_0) siendo una eigenfunción de \hat{F} , se deduce como:

$$\hat{H}_0\Phi_0 = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle \Phi_0 \quad (3.51)$$

de manera que la energía de orden cero es el valor esperado de \hat{H} con respecto a Φ_0 , es decir, la energía de Hartree-Fock:

$$E_{MPO} = E_{HF} = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle \quad (3.52)$$

Dado que la energía de MP para el primer orden:

$$E_{MP1} = \langle \Phi_0 | \hat{V} | \Phi_0 \rangle = 0 \quad (3.53)$$

es cero, la energía de correlación de MP de orden más bajo aparece en el segundo orden. Este resultado es el **teorema de Møller-Plesset**: *El potencial de correlación no contribuye en primer orden a la energía electrónica exacta.*

A fin de obtener la fórmula **MP2** para una molécula closed-shell, la fórmula de segundo orden está escrita en base a los determinantes de Slater doblemente excitados (determinantes de Slater individualmente excitados no contribuyen, ver *teorema de Brillouin*). Después de integrar fuera del espín, se convierte en:

$$E_{MP2} = \sum_{i,j,a,b} \left\langle \psi_i(1)\psi_j(2) \left| \frac{1}{r_{12}} \right| \psi_a(1)\psi_b(2) \right\rangle \times$$

$$\frac{2 \left\langle \psi_a(1)\psi_b(2) \left| \frac{1}{r_{12}} \right| \psi_i(1)\psi_j(2) \right\rangle - \left\langle \psi_a(1)\psi_b(2) \left| \frac{1}{r_{12}} \right| \psi_j(1)\psi_i(2) \right\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \quad (3.54)$$

donde ψ_i y ψ_j son orbitales ocupados, mientras que ψ_a y ψ_b son orbitales virtuales (no ocupados). Las cantidades ϵ_i , ϵ_j , ϵ_a y ϵ_b son las energías de los orbitales correspondientes. Mediante el segundo orden en el potencial de correlación, la energía electrónica total viene dada por la energía de Hartree-Fock además de la corrección de segundo orden de MP:

$$E = E_{HF} + E_{MP2} \quad (3.55)$$

3.2. Metodología

Estamos interesados en estudiar los factores geométricos que permiten la gran riqueza estructural del ADN, es por ello que en el presente trabajo, como objeto principal de estudio, se analizaron las diez posibles conformaciones de dinucleósidos monofosfatos complementarios como unidad mínima de la doble hélice del ADN.

Si bien sabemos, ésto se encuentra fuera de las técnicas experimentales, hasta hoy en día, es necesario que recurramos a las técnicas teórico-computacionales. Dichas técnicas nos permiten en primera instancia optimizar nuestros objetos de estudio, pues minimizan su energía lo cual nos lleva a estructuras físicamente realizables (la naturaleza siempre busca estar en un estado de mínima energía). Ya que las dimensiones del problema son atómicas, podemos hacer uso de la mecánica cuántica para optimizar dichos objetos en estudio, pero recordemos que entre más grande el objeto, más costo computacional, es por ello que nuestro estudio se basa en los dDMPs complementarios como unidad mínima para estudiar las características conformacionales de la hélice doble, pues está comprobado que los dDMPs son las unidades mínimas de la cadena única del ADN que nos pueden aportar información estructural de ésta (ver 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

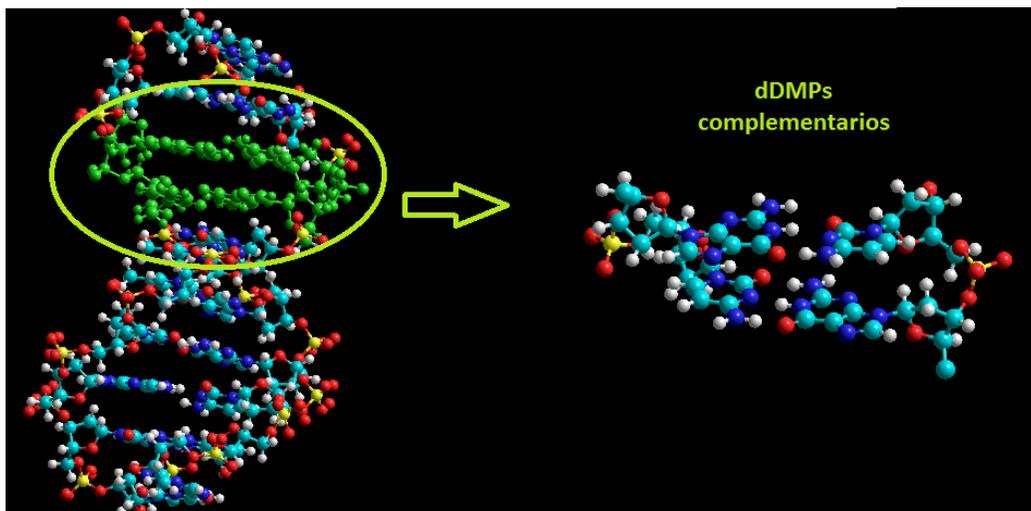


Figura 3.1: *Unidad mínima de la hélice doble del ADN. Dinucleósidos monofosfatos complementarios.*

manteniendo el plano oxígeno-fósforo-oxígeno (O-P-O), del esqueleto azúcar-fosfato, equidistantemente a 2.3\AA de ambos oxígenos. Una vez agregado el ión sodio los dDMP- Na^+ complementarios (ver Fig. 2.5) pueden ser descritos totalmente por los ángulos de torsión del esqueleto azúcar-fosfato (α , β , γ , ϵ , ζ), el empaquetamiento de los dos anillos de la desoxirribosa, los dos ángulos de torsión glucosídicos ($\delta 1$ y $\delta 2$), uno por base y el ángulo promedio formado por las bases nitrogenadas, además de los parámetros *Buckle*, *Propeller*, *h-Twist* y *h-Rise* y la superposición de las bases debida al *Stacking* (ver Fig. 2.7).

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

La Teoría del Funcional de Densidad (DFT) 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. El funcional *PW91* está entre los pocos 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 [12]. Es por ello que dicho funcional es ideal para realizar nuestros estudios, 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* [12]. Estudios más recientes con el funcional *PW91* demostraron que éste reproduce, razonablemente, buenos resultados *ab initio* para complejos con enlaces de hidrógeno [13] y para complejos con bases de ADN apiladas [14].

Para realizar los cálculos de minimización se dispuso del programa ADF (Amsterdam Density Functional). 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. En este trabajo empleamos el conjunto DZVP, el cual además está aumentado con funciones de polarización.

Las estructuras iniciales de dDMPs-Na⁺ complementarios fueron creadas empleando como base el esqueleto de una estructura minimizada anteriormente, con la secuencia de bases Adenina-Timina (dApdT), la cual cuenta con ángulos característicos de la familia estructural BI. Dicha estructura es el resultado de optimizar con el Funcional *PW91*, empleando el programa *Q-Chem*, una estructura experimental ¹, y posteriormente optimizarla nuevamente con éste Funcional, empleando el programa *Gaussian*.

Para algunos dDMPs-Na⁺ complementarios se realizó mas de dos optimizaciones con el objeto de obtener resultados más parecidos con los obtenidos experimentalmente. En tales optimizaciones, después de realizar la optimización con ADF y verificar sus parámetros, la estructura ya optimizada se sometía a un segundo cálculo, seguido de una nueva ronda de optimizaciones de ser necesario.

Cabe mencionar que en éste trabajo también se comparan los resultados obtenidos al optimizar con cálculos *ab initio* al nivel *MP2* estructuras con las secuencias Guanina-Citosina (dGpdC complementarios) y Timina-Adenina (dTpdA complementarios), con los que se obtienen al optimizarlas con *DFT*. Además de comparar los resultados obtenidos al optimizar estructuras con la secuencia Adenina-Timina (dApdT complementarios) con los Funcionales *PWB6K* y *PW91*. Todo esto con el objeto de comparar los funcionales y métodos y analizar cuáles resultados son más parecidos a los obtenidos experimentalmente.

¹Estructura tomada del banco de datos, <http://ndbserver.rutgers.edu/>

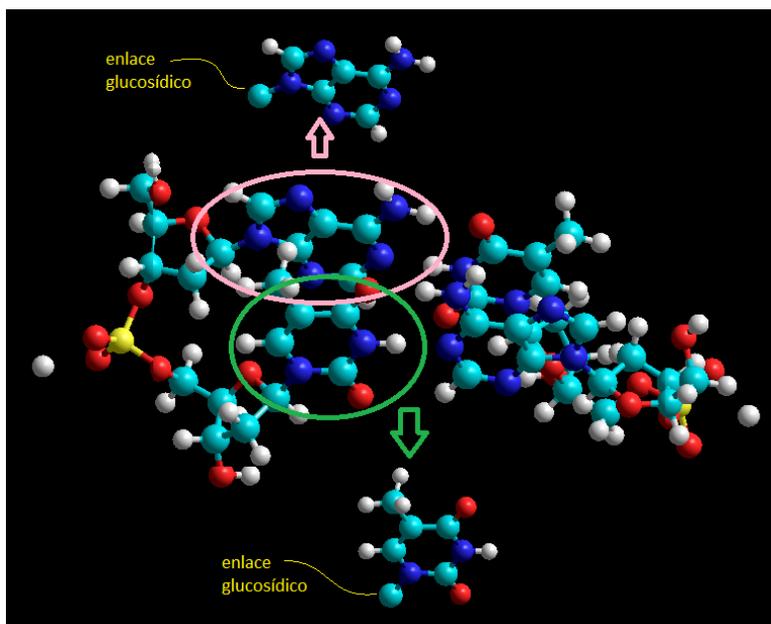


Figura 3.2: Bases de los dDMPs con su enlace glucosídico.

3.2.1. Construcción de diez estructuras iniciales

Como ya se ha mencionado anteriormente, las estructuras iniciales de dDMPs- Na^+ complementarios fueron creadas empleando como base el esqueleto de una estructura, ya optimizada, con la secuencia de bases Adenina-Timina (dApdT), de la familia BI. Fueron diez estructuras las que se crearon, dado que son las necesarias para tener todas las combinaciones posibles de diferentes secuencias de bases nitrogenadas.

Para crearlas es necesario un archivo con extensión *.pdb* que contenga el tipo de átomo, las coordenadas de cada uno, así como los enlaces que forman. Por medio del programa *Hyperchem* fue posible visualizar las estructuras en 3D, moverlas, rotarlas, eliminar átomos... De esta manera nos fue posible, a partir de la estructura con secuencia Adenina-Timina, obtener la estructura con secuencia Timina-Adenina. Para lograrlo es necesario mover las bases de lugar, manteniendo su enlace glucosídico (ver Fig. 3.2), es decir, donde estaba Adenina poner Timina y viceversa (en el mismo dDMP); es importante también colocarlas en el mismo plano que se encontraba la anterior. Esto es necesario hacerlo en ambos dDMPs.

Así, ya tenemos una estructura con purina-pirimidina (A-T) y otra con pirimidina-purina (T-A). Estas fueron la base para construir las otras estructuras con las

secuencias pertinentes. Para poder hacerlo, fue necesario eliminar los radicales y poner los correctos para tener las bases Guanina (purina) y/o Citosina (pirimidina), según sea el caso, teniendo en cuenta que se formen los pares correctos del ADN (ver figura 2.2). Para hacer estos cambios de radicales se empleó el programa *Oxford*, pero fue necesario un cambio de extensión, *.pdb* a *.mol*, en el archivo ya que este programa trabaja con archivos con la extensión *.mol*; una vez que se tuvo la estructura deseada, se realizó un cambio de extensión en el archivo, con la nueva estructura, para trabajar con extensión *.pdb*

Para obtener las estructuras con secuencia purina-purina (en un dDMP, en el otro es pirimidina-pirimidina, debido a los pares de bases que deben formarse), se hizo algo semejante a lo hecho para obtener la estructura con pirimidina-purina con la única diferencia que en lugar de intercambiar las bases de un mismo dDMP, se copió una misma base (Adenina) en ambos nucleósidos, recordar que es necesario que tenga su enlace glucosídico; para obtener el dDMP complementario se realizó lo mismo sólo que en este caso la base de interés fue Timina. Una vez que se tuvo la estructura base, siguiendo el mismo procedimiento con el programa *Oxford*, se crearon las demás estructuras iniciales pertinentes.

3.2.2. 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*, superposición de las bases debido al *Stacking*, *Buckle*, *Propeller*, *h-Rise* y *h-Twist*), se empleó el programa en línea *3DNA*² (ver sección 3.2.2.1).

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 en primera instancia se eliminaron los iones de sodio (Na⁺) y todos los H's, de esta manera, se guardaron por separado los dDMPs de una estructura. Finalmente, se verificaron ambos archivos con los dDMPs, se copiaron sólo las líneas que contenían los átomos y sus coordenadas y se unieron en un solo archivo (con extensión *.pdb*). Éste formato fue el que se empleó para poder medir todas las estructuras optimizadas.

Por otro lado, para obtener el empaquetamiento del azúcar de los dDMPs complementarios se empleó el programa en línea *PROSIT*³, el cual no necesita un formato específico para poder medir las estructuras, por lo que se emplearon las estructuras obtenidas en primera instancia.

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

³<http://cactus.nci.nih.gov/prosit/>

3.2.2.1. 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 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 [9].

Capítulo 4

Resultados

*El problema del físico es el problema de los
orígenes y leyes naturales últimas. El
problema del biólogo es el problema de la
complejidad.*

RICHARD DAWKINS,
etólogo, zoólogo, teórico evolutivo y divulgador
científico.

4.1. Selección de fragmentos de la hélice doble del ADN y detalles de método de DFT para estudio

Hasta ahora, los fragmentos dDMP-Na⁺ son buenos modelos para la cadena única del ADN y las contribuciones que conforman la estructura espacial a las interacciones en la cadena única del ADN.

Ahora la siguiente pregunta es contestar, ¿qué modelos debemos seleccionar para estudiar las características de la hélice doble y sus interacciones entre ambas cadenas?

Existe una sola publicación sobre el estudio de complejos de *Trinucleósidos Difosfatos Complementarios* (dTDPs complementarios) de la familia estructural BI, (dA)₃:(dT)₃ y (dG)₃:(dC)₃ [16]. Para poder hacer una selección de un buen modelo, para estudiar las características estructurales de la hélice doble del ADN, analizamos las estructuras optimizadas de (dA)₃:(dT)₃ en distintos medios, se analizaron tres casos: 1) sin iones de Sodio, 2) con iones de Sodio (Na⁺) y 3) con Na⁺ más inclusión de Agua (como medio continuo).¹

¹No se incluyen resultados en esta tesis para las estructuras de (dG)₃:(dC)₃, pero se cuenta con éstos cálculos. Cabe mencionar que las coordenadas de las estructuras analizadas fueron tomadas de dicho artículo.

Tabla 4.1: *Ángulos de torsión del esqueleto azúcar-fosfato* (véase designación en la figura 2.5) *y valores promedio del ángulo entre los planos de las bases (ϕ), para los $(dA)_3 : (dT)_3$ optimizados, en distintas condiciones.*

dDMPs	$\delta 1$	ϵ	ζ	α	β	γ	$\delta 2$	ϕ
dApdA*	143	219	199	330	136	31	143	5
dTpdT	143	219	199	330	136	31	143	5
dApdA	143	219	199	330	136	31	143	5
dTpdT	143	219	199	330	136	31	143	5
dApdA**	145	190.5	182	12	92	46.5	136.5	7
dTpdT	151	187	188	20	97	42	132	18
dApdA	136.5	192	170	38	73	42	101	4
dTpdT	132	194	171	37	72	43	95	11
dApdA***	150	162	267	294.5	193	55	148	1
dTpdT	145	173	261	298.5	181	58.5	140	12
dApdA	148	162	263	298	194	56	147	5
dTpdT	140	171	259	298.5	182	59	142	4

* estructuras del modelo en vacío sin Na+, ** estructuras del modelo en vacío con Na+, *** estructuras del modelo con Na+ y Agua. Ángulos en grados (°).

Para realizar las optimizaciones de dichas estructuras se empleó el funcional *M06-2X* usando el conjunto base 6-31G(d,p) con el paquete *Gaussian 09*. Este funcional, desarrollado por Zhao y Truhlar, ha demostrado describir las interacciones no covalentes, incluyendo las interacciones de dispersión, mejor que los funcionales de densidad que comúnmente se emplean. Este funcional es lo suficientemente universal como para predecir ambos parámetros geométricos y energéticos, incluyendo la termodinámica y la cinética [17].

En la Tabla 4.1 podemos observar que los ángulos obtenidos para las estructuras optimizadas (en distintos ambientes) no son muy buenos. Podemos observar que para los dTDPs complementarios con Na+ no se obtienen resultados muy buenos, pues los ángulos de torsión no corresponden a los que se tiene en la familia estructural BI, ésto tal vez por la repulsión generada al tener dos iones de Sodio. Por otro lado, podemos ver que para los dTDPs complementarios con Na+ e inclusión de Agua sí existen ángulos pertenecientes a la familia estructural BI. ²

Por otro lado, observando la Tabla 4.2 se tienen algunos empaquetamientos, del anillo de azúcar, fuera de lo que normalmente existe en la forma BI. Además, analizando la Tabla 4.3 podemos ver que tenemos superposición entre las bases nitrogenadas (Purina-Purina), aunque no es mucha como debíamos esperar.

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

Tabla 4.2: *Empaquetamiento del anillo de azúcar y ángulo glucosídico de torsión, χ , para los $(dA)_3:(dT)_3$ optimizados, en distintas condiciones.*

	Nucleósido 5'-extremo				Nucleósido 3'-extremo			
	Empaq.	P	v_{max}	χ	Empaq.	P	v_{max}	χ
dTDPs	Empaq.	P	v_{max}	χ	Empaq.	P	v_{max}	χ
dApdApdA*	C2'-endo	154	45	-98	C2'-endo	154	45	-98
dTpdTpdT	C2'-endo	154	45	-98	C2'-endo	154	45	-98
dApdApdA**	C2'-endo	153	42	-80	C1'-exo	140	48	-104
dTpdTpdT	C2'-endo	165	40	-112	C1'-exo	135	48	-112
dApdApdA***	C3'-exo	184	35	-101	C2'-endo	177	36	-106
dTpdTpdT	C2'-endo	172	33	-114	C2'-endo	160	34	-107
					C2'-endo	157	38	-112

* estructuras del modelo en vacío sin Na⁺, ** estructuras del modelo en vacío con Na⁺, *** estructuras del modelo con Na⁺ y Agua. P, v_{max} y χ en grados (°).

Tabla 4.3: *Parámetros de las bases y pares de bases de los (dA)₃(dT)₃ optimizados.* (véase designación en la figura 2.7).

Estructura	<i>Superposición de las bases</i>					<i>h-Twist</i>	<i>h-Rise</i>	<i>Buckle</i>			<i>Propeller</i>		
	i1-i2	i1-j2	j1-i2	j1-j2	Suma			par1	par2	par3	par1	par2	par3
Modelo*	0.20	0.00	0.00	2.03	2.23	36.0	3.37	0.04	0.04	0.04	-15.13	-15.13	-15.13
	0.20	0.00	0.00	2.03	2.23	36.0	3.37						
Modelo**	0.00	0.00	0.00	2.85	2.85	40.29	3.37	-8.62	-9.64	-19.87	-10.67	-24.25	-29.90
	0.00	0.00	0.00	2.82	2.82	39.14	3.37						
Modelo***	0.00	0.00	0.00	1.20	1.20	44.77	3.03	13.49	6.63	-12.66	-8.93	-22.04	-23.24
	0.00	0.00	0.00	2.20	2.20	44.14	3.33						

* estructuras del modelo en vacío sin Na⁺, ** estructuras del modelo en vacío con Na⁺, *** estructuras del modelo con Na⁺ y Agua.

Para designación de las bases de los dDMFs complementarios, empleada en la superposición, ver Fig. ??

Superposición en (Å²), *h-Twist* en grados (°), *h-Rise* en (Å), *Buckle* y *Propeller* en grados (°).

La Tabla 4.4 muestra algunas distancias entre las bases del dApdApdA-Na⁺, modelado con Agua. Se puede observar que existen contactos cortos entre las bases, por ejemplo, se tienen las distancias 3.08Å (entre los átomos C4-N7, de las dos primeras bases Adeninas), 3.18Å y 3.06Å (entre los átomos C2-C4, de las dos primeras y las dos últimas Adeninas). Cabe mencionar que en los complejos con y sin Na⁺ se tienen aún más contactos cortos entre las bases.

Tabla 4.4: *Distancias más cortas entre las bases del dApdApdA-Na⁺, modelo con agua.*

Átomos	Distancias(Å)	
	1er dApdA	2do dApdA
N1-N1	3.25	3.22
N1-C2	3.24	3.50
N1-C6	3.44	3.18
C2-N3	3.18	3.31
C2-C4	3.18	3.06
C2-C5	3.58	3.22
N3-C4	3.20	3.19
N3-C5	3.46	3.20
N3-C8	3.53	3.18
N3-N9	3.25	3.19
C4-C5	3.26	3.33
C4-N7	3.25	3.08
C6-C6	3.12	3.27
N9-N7	3.25	3.28

Es necesario hacer notar el hecho de que distancias largas, entre los átomos de las bases, no son correctas desde el punto de vista experimental, sin embargo, son posibles. Pero distancias cortas entre las bases no son posibles.

De nuestro análisis para las estructuras de dTDPS complementarios, es claro que este modelo, de dTDPs complementarios, no es bueno, pues se tiene dos Na⁺, por dTDP que hay que estabilizar, y sólo es posible reproducir las características de la hélice doble del ADN con dTDPs-Na⁺ complementarios si se tiene inclusión de Agua (como modelo). Pero, se obtiene como resultado de cálculos contactos cortos. Además de que ese modelo, casi no muestra superposición entre las bases. Por lo tanto, podemos decir que, este modelo no es bueno para estudiar la estructura de la hélice doble del ADN. La propuesta que se hace en ésta tesis es describir como modelo mínimo a: *Dinucleósidos Monofosfatos Complementarios*, en las siguientes secciones.

4.2. Estudio DFT de Fragmentos Mínimos de la hélice doble del ADN

4.2.1. Estudio de diez fragmentos mínimos de la hélice doble del ADN con el funcional *PW91*

Se minimizaron las estructuras iniciales mediante el método DFT (con la funcional *PW91* y base DZVP) haciendo uso del programa ADF, construidas como se menciona en la Sección 3.2.1. Cabe mencionar que las estructuras se sometieron a minimización más de una vez, es decir, algunas estructuras se minimizaron dos, tres, cuatro o cinco veces, en las tablas, sólo se muestran los resultados para una optimización de cada estructura como ejemplo. Las características de tales estructuras minimizadas (dDMPs complementarios) se presentan en tablas. En la Tabla 4.5 se muestran los ángulos de torsión del esqueleto azúcar-fosfato así como el ángulo promedio entre las bases de dichas estructuras, mientras que el empaquetamiento del anillo de azúcar de éstas y los ángulos de torsión glucosídicos (χ) se encuentra en la Tabla 4.6; por otro lado, en la Tabla 4.7 se encuentra la superposición de las bases, los parámetros helicoidales *h-Twist* y *h-Rise*, y los parámetros de los pares de bases *Propeller* y *Buckle*.

Podemos observar en la Tabla 4.6, que para el anillo de azúcar de los nucleósidos en 3'-fin tenemos un menor empaquetamiento a diferencia del que se tiene en el anillo de azúcar de los nucleósidos en 5'-fin, ésto es consecuencia de los cálculos, lo cual ya se ha mostrado anteriormente para los dDMPs en la cadena única[11].

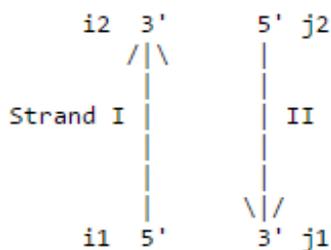


Figura 4.1: Designación de las bases nitrogenadas en el programa *3DNA* para determinar la superposición entre éstas.

Designación de las bases de los dDMPs complementarios en la Tabla 4.7 : Como ya se ha mencionado antes, algunos de los parámetros estudiados en éste trabajo han sido obtenidos mediante el programa en línea *3DNA*. Dicho programa designa a las bases de los dDMPs complementarios, para determinar la superposición entre éstas, como se muestra en la Figura 4.1

Tabla 4.5: Angulos de torsion del esqueleto azucar-fosfato (vease designacion en la figura 2.5) y valores promedio del angulo entre los planos de las bases (ϕ), para los mınimos de dDMP-Na+ complementarios correspondientes a la familia estructural BI.

dDMPs	$\delta 1$	ϵ	ζ	α	β	γ	$\delta 2$	ϕ
Purina - Pirimidina								
dApdT	140	198	280	288	175	51	101	4
dApdT	140	199	280	288.5	175	51	101	4
dApdC	139	198	280	289	174	51	103	7
dGpdT	141	198	278	287	176	50	103	2
dGpdC	141	197	279	286	178	49	103	20
dGpdC	140	196.5	280	286	178	49	104	18
Purina - Purina / Pirimidina - Pirimidina								
dApdA	141	198	279	286	177	51	99	2
dTpdT	137	200	280	287	177.5	51	98	7
dGpdG	141	200	278	288	177.5	50	104	5
dCpdC	137	199	280	287.5	177	50	103	1
dApdG	140	198	279	290	174	51	104	2
dCpdT	139	198	280	288	177	50	103	2
dGpdA	141	196.5	278	286	181	49	103	4
dTpdC	137	199	281	287	176	50	101	7
Pirimidina - Purina								
dTpdA	137	199	279	288	176.5	50	102	0.3
dTpdA	138	197	279	289	175	50	100	1
dCpdA	140	195.5	280	287	178	50	102	0.3
dTpdG	138	199	279	288	176	50	106	1
dCpdG	138	197	280	289	175.5	50	104	1
dCpdG	138	198	280	288	177	50	105	1
BI	132.8	181.7	263.2	299.0	179.3	48.4	132.8	

BI son los valores promedio de la Ref. [3].

Angulos en grados ($^{\circ}$).

Tabla 4.6: Empaquetamiento del anillo de azúcar y ángulo glucosídico de torsión, χ , para los mínimos de dDMP-Na+ complementarios correspondientes a la familia estructural BI.

dDMPs	Nucleósido 5'-extremo				Nucleósido 3'-extremo			
	Empaquetamiento	P	v_{max}	χ	Empaquetamiento	P	v_{max}	χ
dApdT	C2'-endo	172.5	30	-123	C1'-exo	110	43	-120
dApdT	C2'-endo	171	30	-123	C1'-exo	110	43	-120
dApdC	C2'-endo	171	29	-120	C1'-exo	112	44	-120
dGpdT	C2'-endo	173	30.5	-121	C1'-exo	112	42	-118.5
dGpdc	C2'-endo	170	30	-118	C1'-exo	113	42	-122
dGpdc	C2'-endo	172	30	-118	C1'-exo	111.5	42	-122
Purina - Purina / Pirimidina - Pirimidina								
dApdA	C2'-endo	175	30	-121	O4'-endo	106	41	-121
dTpdt	C2'-endo	161	31	-125	O4'-endo	105	42	-121
dGpddG	C2'-endo	172	31	-121	C1'-exo	113	42	-116
dCpdc	C2'-endo	161.5	31	-126.5	C1'-exo	111	43	-121
dApdG	C2'-endo	173	30	-122	C1'-exo	113	42.5	-118
dCpdt	C2'-endo	165	31.5	-125	C1'-exo	112	43	-119.5
dGpda	C2'-endo	170	31	-121	C1'-exo	111	41	-118
dTpdc	C2'-endo	161	30	-125	C1'-exo	110	43	-121
Pirimidina - Purina								
dTpda	C2'-endo	162	32	-122	O4'-endo	107.5	42	-122
dTpda	C2'-endo	162.5	31	-124	C1'-exo	109	42	-117
dCpda	C2'-endo	165	32	-122	C1'-exo	110	42	-119
dTpdc	C2'-endo	164	31	-123	C1'-exo	115	42	-115
dCpdc	C2'-endo	162	31	-122.5	C1'-exo	113	42.5	-118
dCpdc	C2'-endo	163	31	-124	C1'-exo	113	42	-113

P, v_{max} Y χ en grados ($^{\circ}$).

Tabla 4.7: Parametros de las bases y pares de bases de los dDMPs complementarios, con el mismo esqueleto (vease designacion en la figura 2.7).

dDMPs	Superposicion de las bases				<i>h-Twist</i>	<i>h-Rise</i>	<i>Buckle</i>		<i>Propeller</i>		
	i1-i2	i1-j2	j1-i2	j1-j2			Suma	par1	par2	par1	par2
Purina - Pirimidina											
dApdT - dApdT	3.59	0.00	0.00	3.64	7.24	24.52	3.85	0.10	0.06	-1.00	-0.76
dApdC - dGpdT	3.65	0.00	0.00	3.48	7.13	25.85	3.78	0.63	-1.35	0.51	-3.23
dGpdC - dGpdC	3.51	0.00	0.00	3.67	7.18	33.15	2.80	4.68	-4.81	-0.17	-0.47
Purina - Purina / Pirimidina - Pirimidina											
dApdA - dTpdT	3.20	0.00	0.00	0.71	3.91	25.46	4.17	1.21	-1.60	-0.52	-3.31
dGpdG - dCpdC	0.51	0.00	0.00	3.49	4.00	24.94	4.54	0.57	-2.53	-5.01	1.58
dApdG - dCpdT	3.56	0.00	0.00	0.70	4.27	23.53	4.15	-1.04	-0.94	-0.19	-3.48
dGpdA - dTpdC	0.56	0.00	0.00	3.11	3.67	26.96	3.81	0.33	-0.52	-1.50	0.34
Pirimidina - Purina											
dTpdA - dTpdA	0.26	0.00	0.19	0.37	0.82	25.52	4.07	1.90	2.62	-3.19	-5.56
dCpdA - dTpdG	0.35	0.00	0.09	0.33	0.77	24.43	4.13	-1.65	-1.70	-4.56	-1.55
dCpdG - dCpdG	0.53	0.00	0.01	0.29	0.83	23.29	4.12	-2.77	-2.04	-6.45	-3.78

Para designacion de las bases de los dDMPs complementarios, empleada en la superposicion, ver Fig. ??
 Superposicion en (^2), *h-Twist* en grados ($^\circ$), *h-Rise* en (), *Buckle* y *Propeller* en grados ($^\circ$).

En la Figura 4.2 se presenta la distribución de los ángulos de torsión característicos, el ángulo glucosídico (χ), el ángulo de pseudorotación (P) para las estructuras minimizadas, así como el valor de éstos ángulos obtenidos experimentalmente (tomados de la Ref. [3]).

Cada secuencia de dDMPs complementarios se localiza en la región entre dos circunferencias consecutivas. Las secuencias se separan en tres grupos por circunferencias sólidas. Para las estructuras con secuencias **Pur-Pur** tenemos: 1) dApdT·dApdT, 2) dApdC·dGpdT, 3) dGpdC·dGpdC; para **Pur-Pur/Pir-Pir** tenemos: 4) dApdA·dTpdT, 5) dGpdG·dCpdC, 6) dApdG·dCpdT, 7) dGpdA·dTpdC; y para **Pir-Pur** tenemos: 8) dTpdA·dTpdA, 9) dCpdA·dTpdG, 10) dCpdG·dCpdG.

En tales mínimos estructurales se observan pocos cambios entre los ángulos de torsión, de todos los dDMPs complementarios para las diferentes secuencias, esto confirma que no existe dependencia entre los ángulos para la secuencia de las bases. Pero, sí tenemos una clara dependencia en la superposición de las bases con la secuencia de éstas, de ello hablaremos más adelante.

Cabe mencionar que los ángulos obtenidos para todas las estructuras optimizadas están en la región perteneciente a la familia estructural BI, comparando con los resultados obtenidos experimentalmente para cadena única[3]; aunque para los ángulos α y χ se tiene desviación sistemática, y para el ángulo ζ la desviación es más que sistemática debido a la estructura con la que se contruyeron los diez dDMPs complementarios a estudiar, pues dicha estructura es resultado de previas optimizaciones y durante *gaussiana* dicho ángulo va de 264° a 280° , ver Fig. 4.2. Es importante mencionar que los ángulos δ_2 están en una región que no corresponde a la familia estructural BI, esto es una consecuencia de que el esqueleto empleado para construir las diez estructuras tenía éste ángulo fuera de dicha región desde la estructura experimental empleada para obtener este esqueleto, por lo que no se presenta dicho ángulo en la Fig. 4.2.

El empaquetamiento del anillo del azúcar se describe por el ángulo de fase de pseudorotación, \mathbf{P} (0° - 360°), mientras que el grado de desviación de la planaridad del anillo de furanosa es indicado por el máximo *pucker* fuera del plano v_{max} . Estos dos parámetros se definieron por primera vez por Altona y Sundaralingam cuando introdujeron el concepto de pseudorotación para describir el movimiento de empaquetamiento en el anillo de azúcar en nucleósidos y nucleótidos.

El valor de P junto con la amplitud del empaquetamiento v_{max} , pueden describir adecuadamente el *pucker* del azúcar, pues sólo dos ángulos de torsión son variables independientes a pesar de que hay cinco ángulos de torsión endocíclicos en un anillo de furanosa[18].

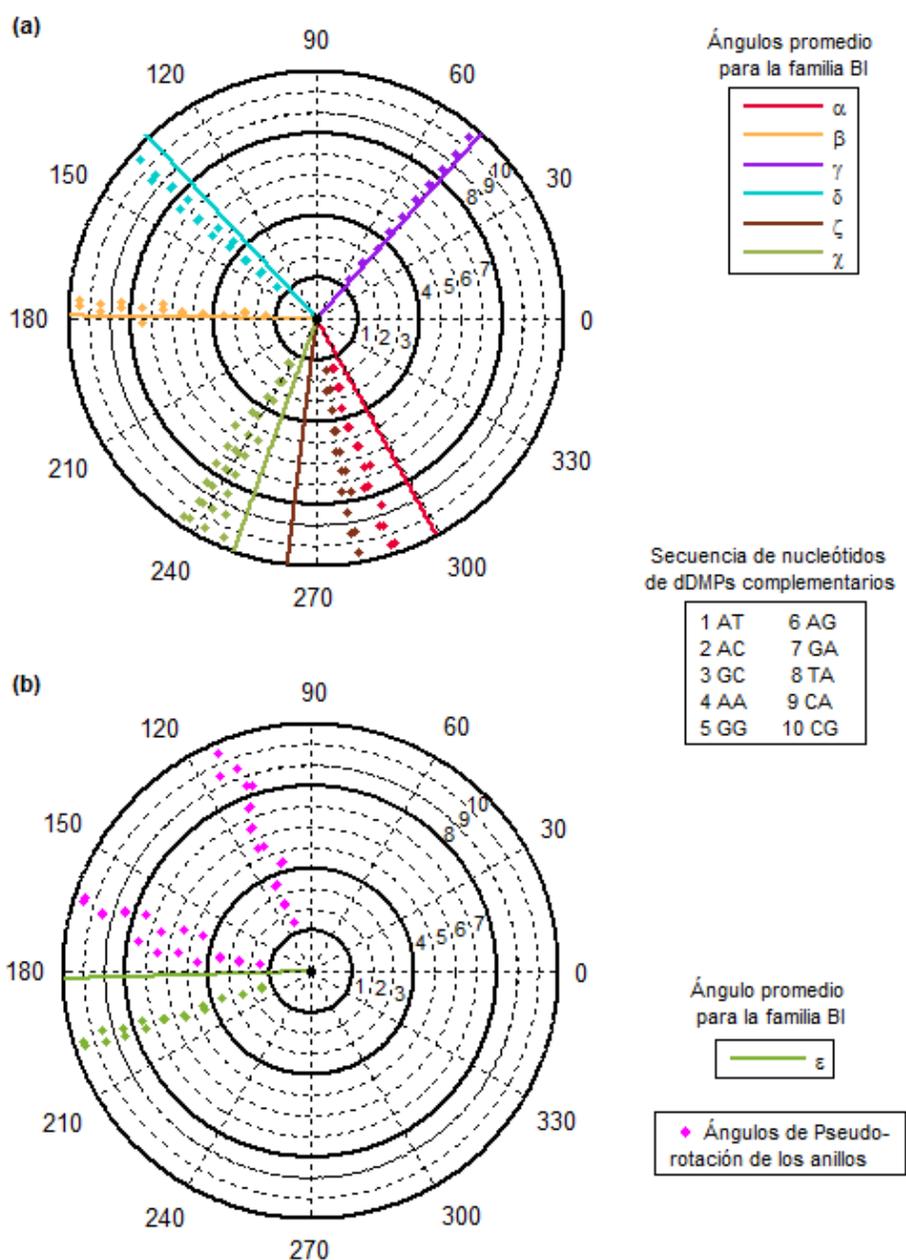


Figura 4.2: Distribución de ángulos para las estructuras minimizadas de los dDMPs complementarios (con diferentes secuencias) de la familia estructural BI. **a)** (ARRIBA) Ángulos glucosídicos (χ) y de torsión. **b)** (ABAJO) Ángulos de pseudorotación y ángulos ϵ . NOTA: En la designación de secuencia de nucleótidos sólo se menciona la secuencia de un dDMP, no del complementario.

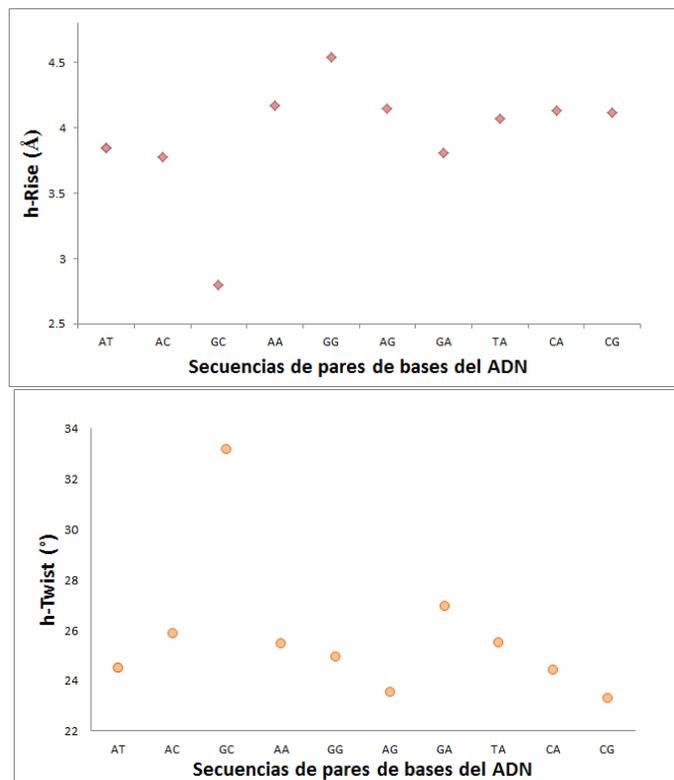


Figura 4.3: *Parámetros de los pares de bases de los dDMPs complementarios minimizados para la familia estructural BI, manteniendo el mismo esqueleto, con respecto al eje de la hélice.*

Además para dichas estructuras optimizadas, en la Fig. 4.3 se encuentran graficados los parámetros helicoidales, donde podemos notar valores de h -Rise un poco más grandes que los que se tienen en el ADN, pues tenemos distancias un poco más grandes entre pares de bases; por otro lado, tenemos valores de h -Twist menores que los que se tienen en la hélice doble, ésto como consecuencia de las distancias más grandes entre pares de bases. Cabe mencionar que se tienen estos resultados debido a que el método empleado no calcula exactamente las interacciones entre pares de bases.

Por otro lado, en la Figura 4.4 se encuentran los valores de los parámetros *Buckle* y *Propeller* para ambos pares de bases de los dDMPs complementarios, donde se observa que no tenemos cambios muy grandes, pues tenemos ambos parámetros como se tienen normalmente en la familia B; sabiendo que, en la familia A existen valores más grandes. Finalmente, en la Figura 4.5 tenemos las gráficas de la suma de la superposición de las bases de los dDMPs complementarios.

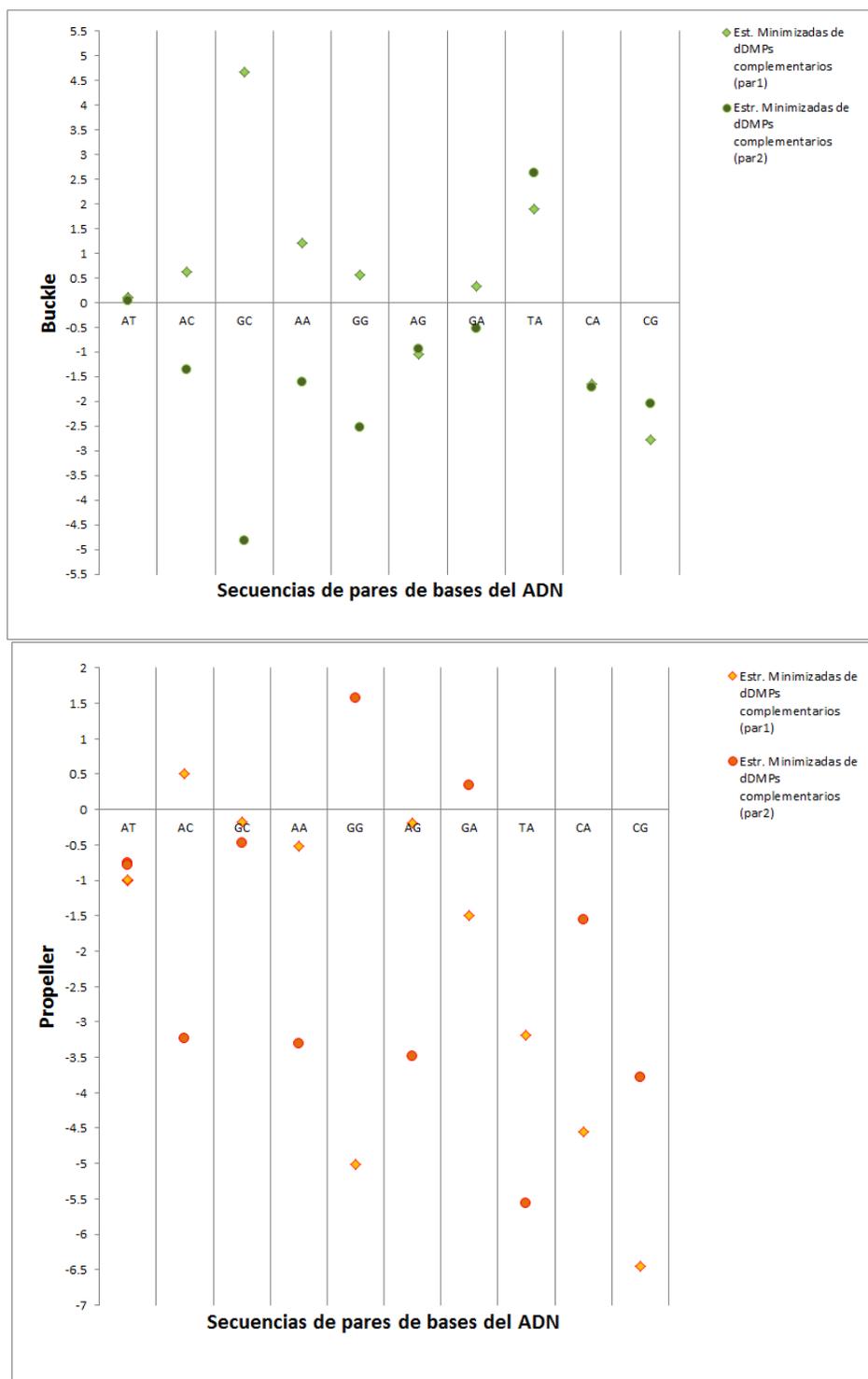


Figura 4.4: Valores de los parámetros *Buckle* y *Propeller* para los pares de bases en los dDMPs complementarios previamente minimizados, de la familia estructural BI.

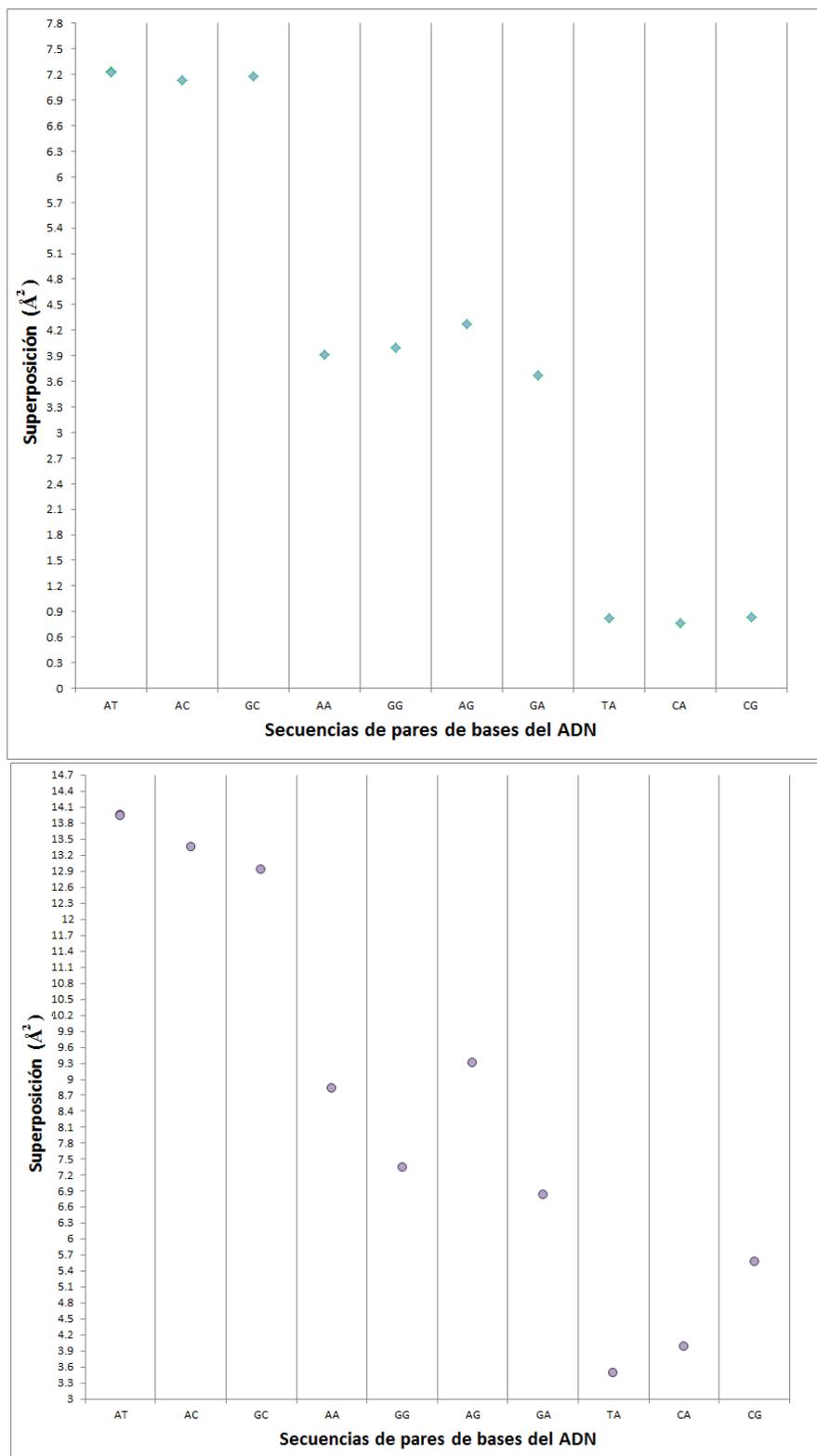


Figura 4.5: *Superposición de las bases, para los dDMPs complementarios minimizados de la familia estructural BI, con el mismo esqueleto inicial, debida al Stacking entre ellas. a) (ARRIBA) Área de superposición de los anillos de las bases. b) (ABAJO) Área de superposición de los anillos de las bases y sus radicales.*

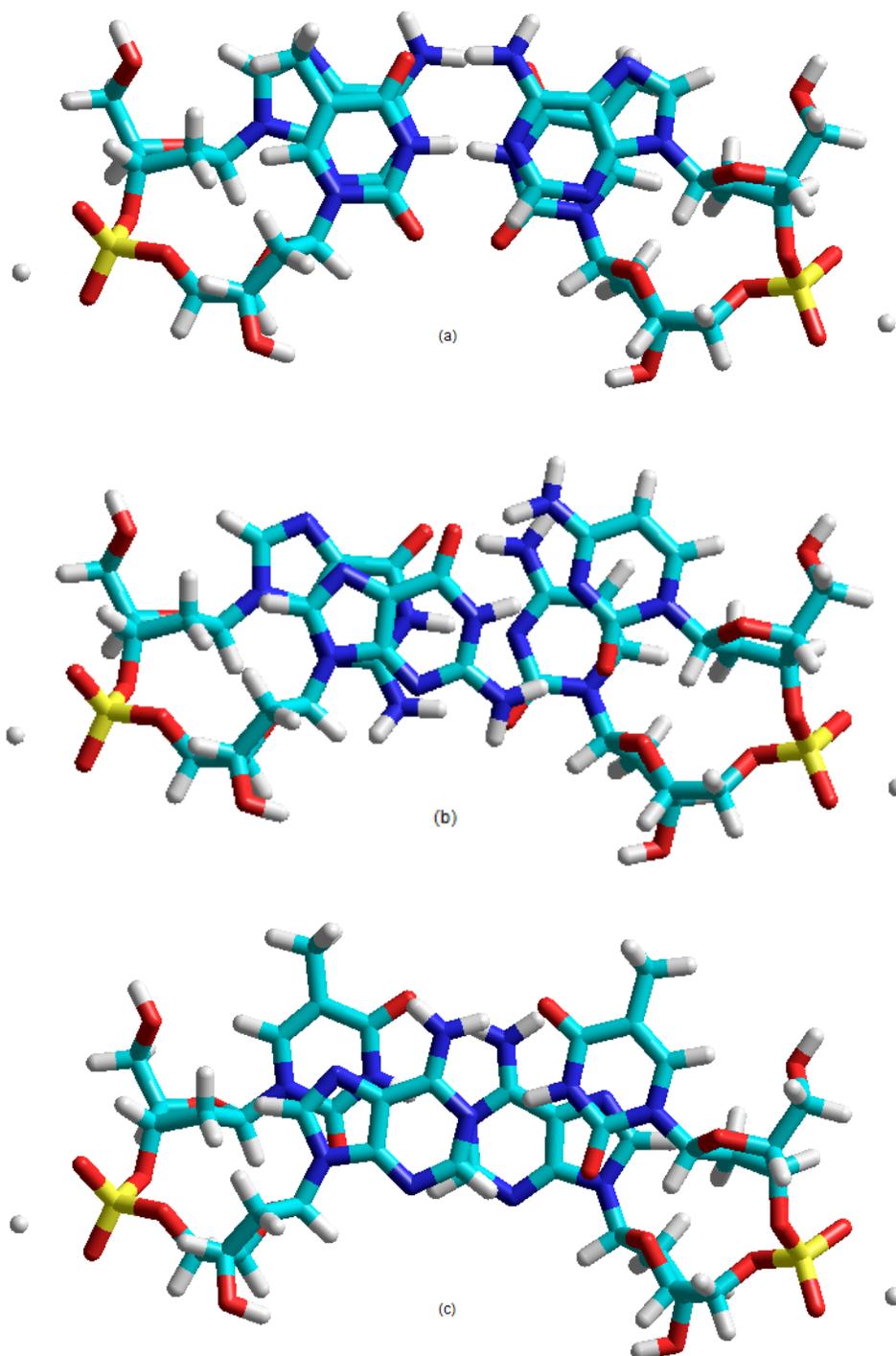


Figura 4.6: Estructuras optimizadas de *dDMPs* complementarios de la familia estructural *BI*, ejemplos característicos: a) *dAdpT* b) *dGpdG* y *dCpdC* c) *dTpdA*.

El análisis de las estructuras optimizadas de los dDMPs complementarios con distintas secuencias de nucleósidos, pertenecientes a la familia estructural BI, permitió constatar interesantes regularidades en las posiciones mutuas de las bases. Resultó que para todas las secuencias *Purina-Pirimidina* y *Purina-Purina* en sus cuatro conformaciones existe un evidente traslape de los planos de los anillos de las bases involucradas (ver Figuras 4.6 *a*) y *b*). En contraste, con las secuencias *Pirimidina-Pirimidina* y *Pirimidina-Purina* este traslape es muy pequeño o inexistente (ver Figuras 4.6 *b*) y *c*). Cabe mencionar que en estudios anteriores[11] para dDMPs en la cadena única se pudo constatar esto cualitativamente.

Una aportación del presente trabajo, es presentar un estudio cuantitativo y comparar con lo reportado anteriormente (ver Figura 4.5). Esta regla fue inicialmente derivada por Dickerson a partir del análisis de datos estructurales de cristales pertenecientes a la familia B obtenidos por medio de rayos X[15]. Nuestros estudios sugieren que este patrón es común a todas las estructuras de doble hélice diestras con pares de Watson-Crick.

Cabe mencionar que las estructuras optimizadas contienen *enlaces de Hidrógeno* correspondientes a la hélice doble del ADN. Como ejemplo tenemos las estructuras con las secuencias Adenina-Timina y Guanina-Citosina: para los pares A-T tenemos las distancias 1.7Å entre los átomos N1 (A) e H3 (T) y 1.8Å entre los átomos H6 (A) y O4 (T); mientras que para los pares G-C tenemos las distancias 1.7Å entre O6 (G) e H4 (C), 1.8Å entre H1 (G) y N3 (C), 1.8Å entre H2 (G) y O8 (C).

Como observación extra, en las estructuras optimizadas los pares de bases se encuentran en planos casi paralelos entre sí, lo cual se puede observar en la Fig. 4.7.

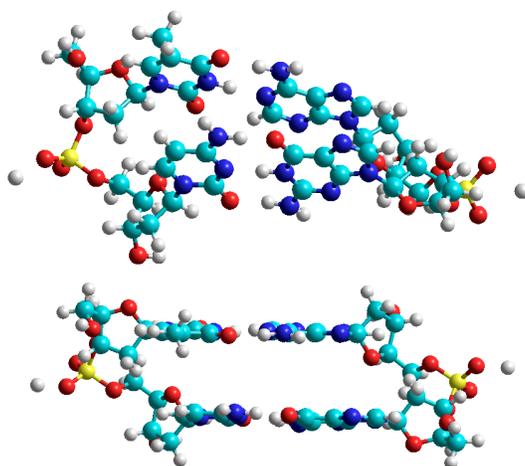


Figura 4.7: Estructura optimizada de dDMPs complementarios con las secuencias Timina-Citosina (dDMP a la izquierda) y Guanina-Adenina (dDMP a la derecha), vista de distintos ángulos.

4.2.2. Comparación de resultados obtenidos con los Funcionales *PWB6K* y *PW91* para dDMPs-Na⁺ complementarios con la secuencia Adenina-Timina

Como se ha mencionado anteriormente, son importantes las interacciones de *Stacking* entre las bases nitrogenadas, ya que depende en parte de éstas la estabilidad de la hélice doble del ADN así como el plegamiento de éste.

La Teoría del Funcional de Densidad es muy eficiente y actualmente se considera como una herramienta importante para el estudio de los sistemas biológicos. Sin embargo, la metodología DFT empleada en los cálculos anteriores, de los diez dDMPs complementarios con distinta secuencia de nucleósidos, no puede describir interacciones de apilamiento pues tiene fallas enormes en los términos de dispersión.

Como continuación de nuestros estudios se analizan las optimizaciones bajo el funcional *PWB6K*, éste es conocido como un funcional meta híbrido, diseñado para cinética. Para éste trabajo se utilizó un conjunto base 6-31G, además de los conjuntos base 3-31G y 3-31G++**.

Otros trabajos han mostrado que DFT-PW91 subestima las energías de estabilización pero sobreestiman las energías de deformación en los monómeros. Yan Zhao y Donald Truhlar[19] dicen que con el funcional *PWB6K* se pueden obtener resultados más realistas para las interacciones del apilamiento en pares de bases del ADN.

Los cálculos DFT se llevaron a cabo en el programa *Northwest Computational Chemistry Package* (NWChem) 6.3, éste es un software que permite seleccionar el funcional, la base y el sistema que se emplearán para la optimización.

Cabe mencionar que los resultados obtenidos son cálculos de prueba, preliminares, pues para realizar estos cálculos no fue posible emplear una *base zeta* más extendida, ya que el programa NWChem y nuestros recursos computacionales no lo permitieron. Sin embargo, con *PWB6K/6-31G* (y las demás bases), no se tienen muchos cambios comparando con los resultados que se obtienen con *PW91*, pero sí se necesitan muchos más recursos computacionales.

En la Tabla 4.8 podemos observar que no existe una gran diferencia entre los ángulos obtenidos con distintas funcionales y dobles bases zeta. De igual manera, se puede ver en la Tabla 4.9 que no se obtiene una diferencia notable en el empaquetamiento entre las estructuras optimizadas. Sin embargo, en la Tabla 4.10, al comparar los valores de *h-Rise* para los diferentes resultados obtenidos, podemos observar que el Funcional *PWB6K* nos da unas distancias más cortas entre los átomos de las bases nitrogenadas, posiblemente por el uso de una base zeta corta, y por ende, cambio de valores de posición mutua de las bases, *h-Twist* y *h-Rise*. Además de que nos da una mayor superposición entre las bases nitrogenadas (Adenina, Timina).

Tabla 4.8: *Ángulos de torsión del esqueleto azúcar-fosfato* (véase designación en la figura 2.5) *y valores promedio del ángulo entre los planos de las bases (ϕ), para los mínimos de dApdT-Na+ complementarios, obtenidos con distintas funcionales.*

Funcional	$\delta 1$	ϵ	ζ	α	β	γ	$\delta 2$	ϕ
<i>estruct1</i>	143.5	195	280	288	175	52	106	7
	143	195	280	288	174	52	106	6
PWB6K/3-21G	148.5	195	268	309	156	44.5	103	4
	148	195	268	309	156	44.5	103	10
PWB6K/-6-31G	144	195	277	291	173	47	105	19.5
	144	195	277.5	291	173	47	105	2
PWB6K/3-31G ^{++**}	142	197	269	302	169	49	100.5	10
	144	195	271	300	169	49.5	101	5
PW91/DZVP	140	198	280	288	175	51	101	4
	140	199	280	288.5	175	51	101	4

estruct1: ángulos para la estructura inicial de los cálculos con el funcional PWB6K, ésta es el resultado de optimizar una estructura experimental con el funcional PW91.

Los ángulos para la estructura optimizada con el funcional PW91 fueron tomados de los resultados de la sección anterior.

Ángulos en grados (°).

Tabla 4.9: Empaquetamiento del anillo de azúcar y ángulo glucosídico de torsión, χ , entre los mínimos $dApdT-Na+$ complementarios, obtenidos con distintas funcionales.

Funcional	Nucleósido 5'-extremo				Nucleósido 3'-extremo			
	Empaquetamiento	P	v_{max}	χ	Empaquetamiento	P	v_{max}	χ
<i>estruct1</i>	C2'-endo	176	30	-123	C1'-exo	115	44	-118
	C2'-endo	176	30.5	-123	C1'-exo	115	44	-119
PWB6K/3-21G	C2'-endo	179	29	-114	C1'-exo	109	48	-123
	C2'-endo	178	29	-114	C1'-exo	110	48	-123
PWB6K/6-31G	C2'-endo	173	29	-122	C1'-exo	112	45	-121
	C2'-endo	173	29	-122	C1'-exo	111	45	-121
PWB6K/3-31G++**	C2'-endo	175	26	-116	O4'-endo	106	45	-118
	C2'-endo	117.5	27	-112	O4'-endo	107	44.5	-119
PW91/DZVP	C2'-endo	172.5	30	-123	C1'-exo	110	43	-120
	C2'-endo	171	30	-123	C1'-exo	110	43	-120

estruct1: empaquetamiento para la estructura inicial de los cálculos con el funcional PWB6K, ésta es el resultado de optimizar una estructura experimental con el funcional PW91.

El empaquetamiento para la estructura optimizada con el funcional PW91 fue tomada de los resultados descritos en la sección anterior.
 P , v_{max} y χ en grados (°).

Tabla 4.10: *Parámetros de las bases y pares de bases de los dApdT complementarios, minimizados con distintas funcionales* (véase designación en la figura 2.7).

Funcional	Superposición de las bases					Suma	<i>h-Twist</i>	<i>h-Rise</i>	<i>Buckle</i>		<i>Propeller</i>	
	i1-i2	i1-j2	j1-i2	j1-j2	par1				par2	par1	par2	
<i>estruct1</i>	3.66	0.00	0.00	3.68	7.34	24.48	3.31	-3.59	3.62	2.28	2.28	
PWB6K/3-21G	3.57	0.00	0.00	3.64	7.21	20.89	2.75	-3.78	3.95	-1.53	-1.58	
PWB6K/6-31G	3.58	0.00	0.00	3.57	7.15	23.34	3.38	-2.33	2.34	3.58	3.62	
PWB6K/3-31G++**	2.61	0.00	0.00	3.26	5.87	23.18	4.04	-3.88	0.44	4.97	2.16	
PW91/DZVP	3.59	0.00	0.00	3.64	7.24	24.52	3.85	0.10	0.06	-1.00	-0.76	

estruct1: parámetros para la estructura inicial de los cálculos con el funcional PWB6K, ésta es el resultado de optimizar una estructura experimental con el funcional PW91.

Los parámetros para la estructura optimizada con el funcional PW91 fueron tomados de los resultados de la sección anterior.

Para designación de las bases de los dDMFs complementarios, empleada en la superposición, ver Fig. ??

Superposición en (\AA^2), *h-Twist* en grados ($^\circ$), *h-Rise* en (\AA), *Buckle* y *Propeller* en grados ($^\circ$).

4.3. Comparación de resultados obtenidos por los métodos MP2 y DFT

En ésta sección se muestran los resultados obtenidos al optimizar algunas estructuras por los métodos MP2 y DFT-PW91, con objeto de comparar ambos métodos, las estructuras iniciales fueron tomadas del banco de datos, además se anexan los resultados obtenidos en la Sección 4.2.1, con el objeto de comparar los resultados. Es importante mencionar que el método MP2 necesita muchos recursos computacionales, por lo que los resultados obtenidos son de colaboradores de EUA [7].

4.3.1. dDMPs complementarios con secuencia Timina-Adenina

En la Tabla 4.11 podemos observar que no se tiene gran diferencia entre los ángulos de torsión de las estructuras optimizadas mediante MP2 y DFT, por otro lado, en la Tabla 4.12 puede verse que se obtienen mejores resultados para las estructuras optimizadas con MP2, aunque no son grandes las diferencias entre los resultados de ambos métodos.

Además en la Tabla 4.13 es notorio el hecho de que al optimizar con MP2, no se muestra una superposición mayor entre las bases de los dDMPs complementarios, lo cual nos indica que ambos métodos dan resultados semejantes. Remarcando el hecho de que MP2 modela distancias más cortas entre los átomos de las bases nitrogenadas, lo cual podemos constatar cuantitativamente con el parámetro *h-Rise*, y cualitativamente en la Figura 4.8.

Tabla 4.11: *Ángulos de torsión del esqueleto azúcar-fosfato y valores promedio del ángulo entre los planos de las bases (ϕ), para los mínimos de dTpdA-Na+ complementarios, obtenidos con MP2 y DFT.*

Métodos	$\delta 1$	ϵ	ζ	α	β	γ	$\delta 2$	ϕ
<i>inicial</i> **	125	191	247	299.5	181	43	139	5
	133	183	244.5	316	165	50	124	5.5
MP2	139	190.5	278	292	168	49	117.5	13
	140	190	278	293	167.5	49	117	12
DFT	116.5	182	268	294	187	52	134	10
	120	179	268	299	179	56	123	5
MP2-DFT	137	196	280.5	289	177	49	113	23
	137	196	281	288	176	48	114	24
DFT*	137	199	279	289	176.5	50	102	0.3
	138	197	279	289	175	50	100	1

** Estructura experimental tomada del banco de datos.

* Los ángulos de la estr. optim. con el método DFT son tomados de los resultados de la Sec. 4.2.1.

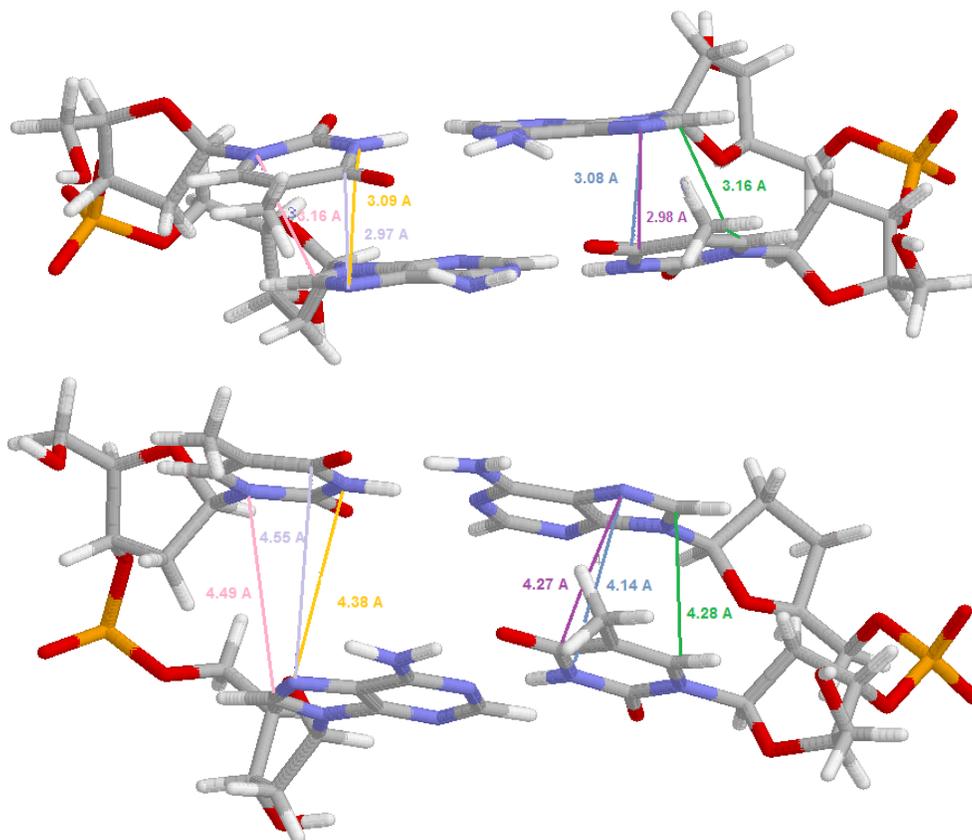


Figura 4.8: Estructuras optimizadas de dDMPs complementarios de la familia estructural BI, con la secuencia Timina-Adenina. **a)** (ARRIBA) dDMPs complementarios optimizados mediante el método MP2. **b)** (ABAJO) dDMPs complementarios optimizados mediante el método DFT-PW91.

En la Figura 4.8 podemos observar una clara diferencia entre las distancias de los dDMPs complementarios minimizados por los métodos MP2 y DFT. Es evidente que se tiene distancias más cortas entre *N1* de Timina en 5'-fin con *C8* de Adenina en 3'-fin, *C4* de Timina con *N7* de Adenina y *N3* de Timina con *N7* de Adenina en los dDMPs optimizados por el método MP2, comparando con las que las que obtienen en los dDMPs optimizados por el método DFT.

Es necesario mencionar que distancias menores que 3.2Å no son posibles, pues tendríamos contactos cortos entre los átomos. Ésta es una gran desventaja para el método MP2, a pesar de ser conocido como excelente método de primeros principios, estos no son mejores que los obtenidos con DFT.

Tabla 4.12: *Empaquetamiento del anillo de azúcar y ángulo glucosídico de torsión, χ , para los mínimos de dTpdA-Na+ complementarios, obtenidos con MP2 y DFT.*

Métodos	Nucleósido 5'-extremo				Nucleósido 3'-extremo			
	Empaquetamiento	P	v_{max}	χ	Empaquetamiento	P	v_{max}	χ
<i>inicial</i>	C1'-exo	133	41	-118	C2'-endo	148	36.5	-113
	C2'-endo	144	39	-108	C1'-exo	132	35	-119
MP2	C2'-endo	158.5	32	-116	C1'-exo	121	48	-127
	C2'-endo	159	33	-116	C1'-exo	120	48	-129
DFT	C1'-exo	123	38	-119	C2'-endo	149	36	-111
	C1'-exo	129	34.5	-113	C1'-exo	132.5	36	-120
MP2-DFT	C2'-endo	158	32	-121	C1'-exo	120	43	-125
	C2'-endo	158	32	-120	C1'-exo	120	43	-125
DFT*	C2'-endo	162.5	31	-124	C1'-exo	109	42	-117
	C2'-endo	162	32	-122	O4'-endo	107.5	42	-122

** Estructura experimental tomada del banco de datos.

*El empaquetamiento para la estructura optimizada con el método DFT fue tomado de los resultados de la Sección 4.2.1.

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

Tabla 4.13: *Parámetros de las bases y pares de bases de los dTpdA complementarios, minimizados con MP2 y DFT (véase designación en la figura 2.7).*

Métodos	Superposición de las bases					Suma	<i>h-Twist</i>	<i>h-Rise</i>	<i>Buckle</i>		<i>Propeller</i>	
	i1-i2	i1-j2	j1-i2	j1-j2	par1				par2	par1	par2	
<i>inicial</i>	0.00	0.00	0.08	0.27	0.35	34.42	3.19	-0.56	-0.48	-7.89	-3.69	
MP2	0.28	0.00	0.41	0.20	0.89	27.64	2.52	8.46	-9.04	-14.56	-15.50	
DFT	0.00	0.00	0.08	0.14	0.21	34.36	3.73	5.28	-0.43	-10.05	-5.90	
MP2-DFT	0.00	0.00	0.89	0.00	0.89	36.17	2.42	9.87	-10.37	-12.68	-13.36	
DFT*	0.26	0.00	0.19	0.37	0.82	25.52	4.07	1.90	2.62	-3.19	-5.56	

** Estructura experimental tomada del banco de datos.

* Los parámetros para la estructura optimizada con el método DFT fueron tomados de los resultados de la Sección 4.2.1.

Para designación de las bases de los dDMFs complementarios, empleada en la superposición, ver Fig. ??

Superposición en Å², *h-Twist* en grados (°), *h-Rise* en Å, *Buckle* y *Propeller* en grados (°).

4.3.2. dDMPs complementarios con secuencia Guanina-Citosina

Podemos observar en la Tabla 4.14 que los ángulos obtenidos para los resultados de MP2 no son mejores que los obtenidos con DFT, además de que con MP2 tenemos los ángulos α y β fuera de la región de la familia estructural B, para uno de los dGpdC.

De igual forma, se puede observar en la Tabla 4.15 que como resultado de la optimización con DFT tenemos algunos valores para el ángulo de pseudorotación, P , fuera de la región para la familia B.

Además, observemos en la Tabla 4.16 que con MP2 obtenemos una mayor superposición entre las bases, sin embargo, tenemos distancias más cortas que las obtenidas con DFT.

Hay que recalcar que la diferencia de resultados entre los dDMPs complementarios radica en que son de distinta familia estructural.

Tabla 4.14: *Ángulos de torsión del esqueleto azúcar-fosfato (véase designación en la figura 2.5) y valores promedio del ángulo entre los planos de las bases (ϕ), para los mínimos de dGpdC-Na⁺ complementarios, obtenidos con distintos métodos.*

Métodos	$\delta 1$	ϵ	ζ	α	β	γ	$\delta 2$	ϕ
<i>inicial</i>	150	250	181	300	142	28	109	4
	137	167	290.5	292	183	50	113	3
MP2	147	189.5	172	26	83.5	45	111	4
	149	186	273	297	163.5	58	94	8
DFT(M06)	135	270	156	287	130	47	122	10
	147	186	272	298	161	59	76	3
DFT(PW91)	133	262	162.5	289	136.5	54	132	37
	144	199	267	298	160	56.5	83.5	18
DFT*	141	197	279	286	178	49	103	20
	140	197	280	286	178	49	104	18

inicial: ángulos para la estructura inicial con los métodos MP2 y DFT, dicha estructura es un fragmento de la estructura BDJ039 tomada del banco de datos³, cabe mencionar que ésta es una estructura híbrida pues la primer cadena es de la familia estructural BII, mientras que la segunda cadena es de la familia BI.

* Los ángulos para la estructura optimizada con el método DFT fueron tomados de los resultados de la Sección 4.2.1.

Ángulos en grados ($^{\circ}$).

Tabla 4.15: Empaquetamiento del anillo de azúcar y el ángulo glucosídico de torsión, χ , para los mínimos de $dGpdC-Na^+$ complementarios, obtenidos con distintos métodos.

Métodos	Nucleósido 5'-extremo				Nucleósido 3'-extremo			
	Empaquetamiento	P	v_{max}	χ	Empaquetamiento	P	v_{max}	χ
<i>inicial</i>	C2'-endo	159	48	-96	C1'-exo	116	25	-116
	C3'-exo	186	24	-95	C1'-exo	118	35	-120
MP2	C2'-endo	158	40	-101	C1'-exo	120	47	-132
	C2'-endo	170	34	-106	O4'-endo	101	48	-139
DFT(M06)	C1'-exo	132	46	-106	C1'-exo	130.5	44	-129
	C2'-endo	174	31	-108	O4'-endo	73	45	-146
DFT(PW91)	C1'-exo	132	44	-117	C2'-endo	152	36	-116
	C2'-endo	174	30	-111	O4'-endo	83	41	-123
DFT*	C2'-endo	170	30	-118	C1'-exo	113	42	-122
	C2'-endo	172	30	-118	C1'-exo	112	42	-122

inicial: empaquetamiento para la estructura inicial con los métodos MP2 y DFT, dicha estructura es un fragmento de la estructura BDJ039, estructura híbrida pues la primer cadena es de la familia estructural BI1, mientras que la segunda cadena es de la familia BI.

* El empaquetamiento para la estructura optimizada con el método DFT fueron tomados de los resultados de la Sección 4.2.1.

P, v_{max} y χ en grados ($^\circ$).

Tabla 4.16: *Parámetros de las bases y pares de bases de los dGpdC complementarios, minimizados con distintos métodos (véase designación en la figura 2.7).*

Métodos	Superposición de las bases			<i>h-Twist</i>		<i>h-Rise</i>		<i>Buckle</i>		<i>Propeller</i>	
	i1-i2	i1-j2	j1-j2	Suma	par1	par2	par1	par2	par1	par2	
<i>inicial</i>	2.63	0.00	0.00	2.93	36.93	3.64	6.32	-7.08	-8.22	-11.68	
MP2	1.30	0.00	0.00	4.06	30.54	3.08	3.24	4.51	9.47	-2.77	
DFT(M06)	0.65	0.00	0.00	3.92	32.20	3.08	2.04	8.08	10.64	-4.05	
DFT(PW91)	2.21	0.00	0.00	2.21	44.12	3.19	-4.65	0.38	-11.77	11.50	
DFT*	3.51	0.00	0.00	3.68	32.97	2.81	4.54	-4.73	-0.18	-0.42	

inicial: parámetros para la estructura inicial con los métodos MP2 y DFT, dicha estructura es un fragmento de la estructura BDJ039, estructura híbrida pues la primer cadena es de la familia estructural BII, mientras que la segunda cadena es de la familia BI.

* Los parámetros para la estructura optimizada con el método DFT fueron tomados de los resultados de la Sección 4.2.1.

Para designación de las bases de los dDMPs complementarios, empleada en la superposición, ver Fig. ??

Superposición en Å², *h-Twist* en grados (°), *h-Rise* en Å, *Buckle* y *Propeller* en grados (°).

Capítulo 5

Conclusiones

*Todos los sueños pueden hacerse realidad si
tenemos el coraje de perseguirlos.*

WALT DISNEY

El análisis de los resultados obtenidos para los *Trinucleósidos Difosfatos Complementarios*, muestra que éstos no son un buen modelo como se propone para los dDMPs-Na⁺ complementarios, para el estudio de las características de la hélice doble. Aún con el modelo con Na⁺ y Agua, que es el que mejor reproduce las características de la hélice doble, se tiene existencia de contactos cortos, los cuales no describen bien la estructura para la hélice doble del ADN.

Como consecuencia del análisis de los dTDPs complementarios, elegimos como modelo para estudiar las características de la hélice doble a los *Dinucleósidos Monofosfatos Complementarios*. Las diez estructuras de dDMPs complementarios optimizadas y descritas en ésta tesis, con todas las posibles secuencias de nucleósidos, muestran valores para los ángulos de torsión semejantes a los ya bien definidos para la cadena única de la familia estructural BI. Además, corroboran cuantitativamente los resultados obtenidos anteriormente para la superposición de las bases, dependiendo la secuencia de nucleósidos (Purinas, Pirimidinas). Los resultados muestran dependencia de la estructura en la secuencia de las bases, dDMPs que contienen bases iguales pero secuencia distinta pueden presentar diferentes características.

Por otro lado, el análisis de los resultados obtenidos con el Funcional PWB6K, muestran que éste mantiene las características generales hasta ahora descritas por el Funcional PW91, por las cuales optamos, ya que no demandan un exceso en el costo computacional.

Los resultados obtenidos de primeros principios, con Teoría de Perturbaciones (MP2), no mejoran ampliamente las características estructurales para la descripción de los dDMPs-Na⁺ complementarios, pues muestran contactos cortos entre las bases apiladas, lo cual no puede ser.

De la comparación de los diferentes métodos y funcionales, todos nos dan resultados generales parecidos, así que éstos no dependen de los métodos empleados. Por lo tanto, hasta ahora no es posible hacer conclusiones sobre qué método o funcional es necesario para reproducir las características de fragmentos de la hélice doble del ADN.

Apéndice A

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

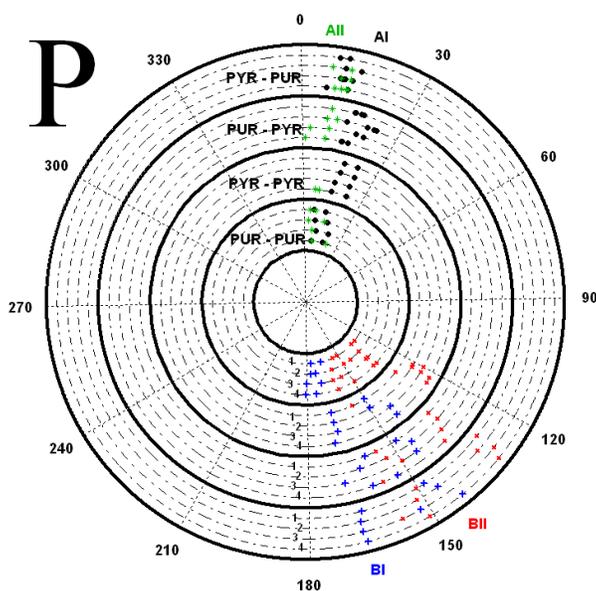


Figura A.1: Distribución del ángulo pseudorotación P del anillo de azúcar en las estructuras optimizadas de dDMPs, correspondientes a dúplex diestros. Cada secuencia de bases se localiza en una circunferencia punteada. Cuatro grupos de secuencias están separados por círculos sólidos. Estos son Pur-Pur: 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.[11]

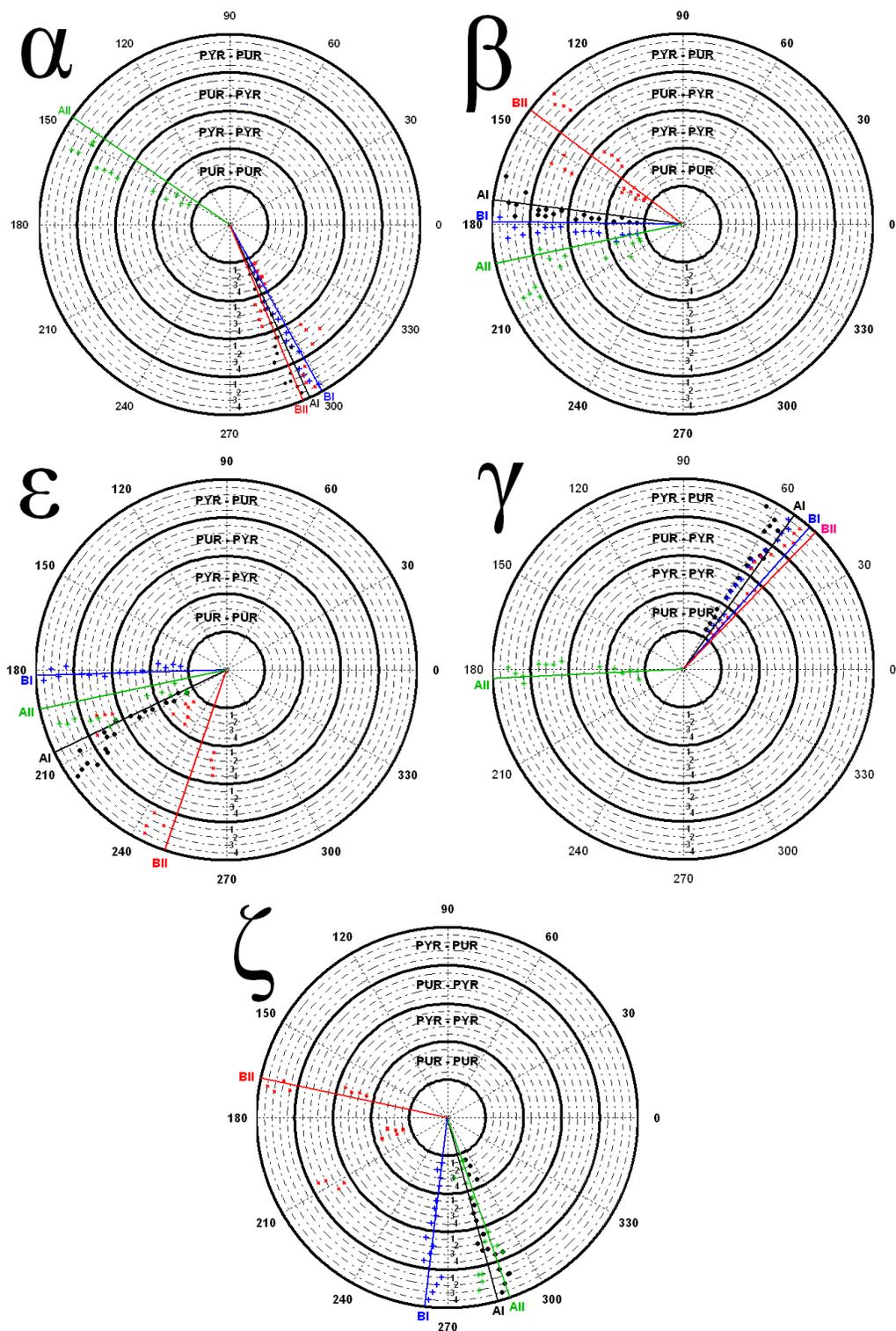


Figura A.2: Distribución de los ángulos de torsión del esqueleto azúcar-fosfato para estructuras optimizadas de dDMPs. Los valores promedio de los ángulos de torsión para cada una de cuatro familias de dúplex en cristales se indican mediante líneas rectas.

Bibliografía

*Y así, del mucho leer y del poco dormir, se
le secó el cerebro de manera que vino a
perder el juicio.*

Miguel de Cervantes Saavedra

- [1] B. Bryson. *Una Breve Historia de Casi Todo*. RBA bolsillo, séptima edition, 2008.
- [2] 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.
- [3] D. Svozil, J. Kalina, M. Omelka, and B. Schneide. Dna conformations and their sequence preferences. *Nucleic Acids Research*, 36, pp. 3690-3706, 2008.
- [4] 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.
- [5] V. I. Poltev, V. M. Anisimov, V. I. Danilov, A. Deriabina, E. Gonzalez, D. Garcia, F. Rivas, A. Jurkiewicz, and A. L. 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.
- [6] V. I. Poltev, V. M. Anisimov, V. I. Danilov, T. V. Mourik, A. Deriabina, E. González, M. Padua, D. García, F. Rivas, and N. 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.

- [7] V. Poltev, V. M. Anisimov, V. I. Danilov, D. García, C. Sánchez, A. Deriabina, E. González, F. Rivas, and N. 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, 2013.
- [8] V. R. Cooper, T. Thonhauser, A. Puzder, E. Schroder, B. I. Lundqvist, and D. C. Langreth. Stacking interactions and the twist of dna. *Journal of the American Chemical Society (JACS)*, 130, pp. 1304-1308, 2008.
- [9] X.-J. Lu and W. 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.
- [10] R. E. Dickerson, H. R. Drew, B. N. Conner, R. M. Wing, A. V. Franiti, and M. L. Kopka. The anatomy of a-, b-, and z-dna. *Science*, 216, pp. 475-485, 1982.
- [11] 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.
- [12] 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.
- [13] V. 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.
- [14] 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.
- [15] R. Dickerson, R. Sarma, and M. Sarma. Structure, motion, interaction and expression of biological macromolecules. *Adenine Press, Albany, NY*, 1, pp. 17-36, 1998.
- [16] T. A. Zubatiuk, O. V. Shishkin, L. Gorb, D. M. Hovorun and J. Leszczynski B-DNA Characteristics are Preserved in Double Stranded d(A)₃-d(T)₃ and d(A)₃-d(T)₃ Mini-Helices: Conclusion from DFT/M06-2X Study stacking interactions in biological systems? *Phys. Chemistry Chemical Physics*, 15, no. 15, pp. 18155-18166, 2013.

-
- [17] Y. Zhao and D. G. Truhlar. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Phys. Theoretical Chemistry Accounts*, 120(1-3), pp. 215-241, 2008.
- [18] G. Sun, J. H. Voigt, I. V. Filippov, V. E. Marquez, and M. C. Nicklaus. Prosit: Pseudo-rotational online service and interactive tool, applied to a conformational survey of nucleosides and nucleotides. *J. Chemical Inf. Comput. Sci.*, 44, pp. 1752-1762, 2004.
- [19] Y. Zhao and donald G. Truhlar. How well can new-generation density functional methods describe stacking interactions in biological systems? *Phys. Chem*, 7, pp. 2701-2705, 2005.

*Pregúntate si lo que estás haciendo hoy
te acerca al lugar en el que quieres estar mañana.*

WALT DISNEY

*En los sueños encontramos un mundo enteramente nuestro,
que se sumerge en su profundo océano,
que sobrevuela la más alta nube...*

ALBUS DUMBLEDORE

*Vive como si fueras a morir mañana,
aprende como si fueras a vivir
para siempre.*

JAMES DEAN

