



# Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

## **ESTUDIO COMPUTACIONAL DE LAS REGULARIDADES CONFORMACIONALES DE FRAGMENTOS DE HELICE DOBLE DEL ADN CON SECUENCIAS DE PURINA-PURINA.**

Tesis presentada al

**Posgrado en Ciencias Física Aplicada**

Como requisito para obtener el grado de:

**Maestro en Ciencias**

(Física Aplicada)

Presentada por:

**Lic. Edgar Vidal Carrillo**

Asesorado por

**Dr. Valeri Poltev**

**Dr. Eduardo González Jiménez**

Puebla Pue.

Junio 2021

**Título:** ESTUDIO COMPUTACIONAL DE LAS REGULARIDADES CONFORMACIONALES DE FRAGMENTOS DE HELICE DOBLE DEL ADN CON SECUENCIAS DE PURINA-PURINA.

Estudiante: EDGAR VIDAL CARRILLO

## COMITÉ

---

**Dra. Alexandra Deriabina**

**Presidente**

---

**M.C. Ana María Cervantes Tavera**

**Secretario**

---

**Dr. Eduardo Moreno Barbosa**

**Vocal**

---

**Dr. Valeri Poltev**

**Asesor**

---

**Dr. Eduardo González Jiménez**

**Asesor**

A mi amá.

A mis brodis.

## Agradecimientos

Me encuentro infinitamente agradecido con el doctor Valeri Poltev por permitirme ser parte de este proyecto, por su gran apoyo y enseñanzas, y principalmente por su increíble paciencia.

Al doctor Eduardo Gonzáles Jiménez y a la doctora Alexandra Deriabina por su apoyo, guía y comentarios para el avance de este trabajo.

A mi madre, Rosa Natalia Carrillo Mendoza, y a mis hermanos Andrea, Rosario, Rafael y Priscila por el apoyo y la fé que me mostraron.

A Nadia y a Feito por esos 5 años de apoyo, constancia, alegrías, fiestas, conflictos y todas esas cosas que solo los roomies conocen.

A los que fueron y a los son parte importante de mi vida.

## Contenido

<b>RESUMEN</b> .....	ix
<b>OBJETIVO GENERAL:</b> .....	x
<b>OBJETIVOS ESPECÍFICOS:</b> .....	x
<b>Capítulo 1. Introducción</b> .....	11
1.1 Estructura del ADN.....	12
1.3 Estructuras secundarias del ADN.....	18
1.4 Parámetros estructurales del ADN.....	20
<b>Capítulo 2. Metodología</b> .....	23
2.1 Mecánica Molecular .....	23
2.1.1 Términos enlazantes .....	24
2.1.2 Términos no enlazantes.....	26
2.2 Campo de fuerzas AMBER (Assisted Model Building with Energy Refinement).....	27
2.3 Nucleic Acid Database (NDB) y Protein Data Bank.....	28
2.4 Software en línea 3DNA.....	29
2.4 Software en línea DNATCO.....	29
<b>Capítulo 3. Resultados</b> .....	31
3.1 Búsqueda de fragmentos mínimos de hélice doble en NADB y análisis de sus características estructurales. ....	32
3.2 Mínimos locales de energía para fragmentos del esqueleto azúcar-fosfato y anillos desoxirribosa pertenecientes a diferentes conformaciones. ....	34
3.3 Fragmentos mínimos de hélice doble optimizados con los tres campos de fuerza AMBER.....	37
3.3.1 Fragmentos en conformación BB12.....	37
3.3.2 Fragmentos en conformación BB02.....	42
3.3.3 Fragmentos en conformación BA09, y BA13.....	46
<b>Capítulo 4. Conclusiones</b> .....	50
Bibliografía.....	51

## Índice de Figuras

FIGURA 1 NUMERACIÓN DE ÁTOMOS DE PIRIMIDINA Y PURINA.....	13
FIGURA 2 BASES NITROGENADAS PRINCIPALES DE LOS ÁCIDOS NUCLEICOS. ....	13
FIGURA 3 BASES NITROGENADAS Y SUS PARES COMPLEMENTARIOS.....	14
FIGURA 4 AZÚCAR DESOXIRRIBOSA, PRESENTE EN EL ADN.....	14
FIGURA 5 DESOXIRRIBONUCLEÓTIDO PDT. ....	15
FIGURA 6 DESOXIDINUCLEÓSIDO MONOFOSFATO, DAPDA. ....	15
FIGURA 7 DESOXIDINUCLEÓSIDO MONOFOSFATO, DAPDA:DTPDT.....	16
FIGURA 8 ESQUELETO AZÚCAR FOSFATO SPB.....	16
FIGURA 9 CADENA ÚNICA DE ADN, PRIMER NIVEL DE ORGANIZACIÓN.....	17
FIGURA 10 NIVELES DE ORGANIZACIÓN DEL ADN COMENZANDO POR LA HÉLICE DOBLE (SEGUNDO NIVEL) EN LA IZQUIERDA. ....	18
FIGURA 11 MOLÉCULAS DE ADN CON CONFORMACIONES (A) B-ADN, (B) A-ADN Y (C) Z-ADN. ....	19
FIGURA 12 ANGULO DE TORSIÓN FORMADO POR EL PLANO DE LOS ÁTOMOS A, B Y C CON EL PLANO DE ÁTOMOS B, C Y D. ....	20
FIGURA 13 A) DESIGNACIÓN DE LOS ÁNGULOS DE TORSIÓN DE UN dDMP (dCpDC). B) DIFERENTES EMPAQUETAMIENTOS DEL AZÚCAR. ....	21
FIGURA 14 DESCRIPCIÓN GRÁFICA DE LOS PARÁMETROS UTILIZADOS PARA DESCRIBIR LA GEOMETRÍA EN LOS PARES DE BASES COMPLEMENTARIOS Y SECUENCIAS DE PARES DE BASES. ....	22
FIGURA 15 REPRESENTACIÓN GRÁFICA DE LOS TÉRMINOS ARMÓNICOS. A) TÉRMINO DE ENLACE. B) TÉRMINO DE FLEXIÓN ANGULAR. ....	25
FIGURA 16 REPRESENTACIÓN GRÁFICA DEL TÉRMINO DE ÁNGULOS DE TORSIÓN.....	26
FIGURA 17 DIFERENCIAS ENTRE LOS ESQUELETOS AZÚCAR FOSFATO EN DIFERENTES CONFORMACIONES A) BA09 B) BB12 C) BA13 Y D) BB02. ....	34
FIGURA 18 FRAGMENTO SPB 1WET QUE PRESENTA INTERACCIÓN H5'--O4' DESPUÉS DE SER OPTIMIZADO CON CF OL15. ....	36
FIGURA 19 ESTRUCTURA DEL ANILLO DE AZÚCAR DE UN FRAGMENTO EN CONFORMACIÓN BB02. ANILLO INICIAL Y ANILLOS OPTIMIZADOS CON LOS CAMPOS DE FUERZA BSC1, FF99 Y OL15. ....	37
FIGURA 20 COMPARACIÓN ENTRE LOS cDDMP INICIAL Y OPTIMIZADO CON BSC1 PARA EL cDDMP CON CONFORMACIÓN BB12 DE SECUENCIA dGpDg:dCpDC PARA EL CRISTAL 2F5N, EL CUAL NO PRESENTA MUCHOS CAMBIOS DESPUÉS DE SER OPTIMIZADO.....	41
FIGURA 21 COMPARACIÓN ENTRE LOS cDDMP CON CONFORMACIÓN BB12 DE SECUENCIA dGpDg:dCpDC INICIAL Y OPTIMIZADO DEL CRISTAL 1R7M EL CUAL PRESENTA CAMBIOS SIGNIFICATIVOS DESPUÉS DE SER OPTIMIZADO. ....	42
FIGURA 22 FRAGMENTO cDDMP EN CONFORMACIÓN BB02 Y CON SECUENCIA dCpDC:dGpDg DEL CRISTAL 1A1H.....	43
FIGURA 23 FRAGMENTOS cDDMP OPTIMIZADOS CON CONFORMACIÓN BB02 IZQUIERDA) OPTIMIZACIONES FF99 Y OL15 PARA EL FRAGMENTO 1A73 CON SECUENCIA dTPdT:dAPdA EN EL CUAL SE PRESENTA UN BUCKLE GRANDE. DERECHA) OPTIMIZACIONES CON OL15 Y FF99 PARA LOS FRAGMENTOS 1HLV Y 1M19 RESPECTIVAMENTE, EN LOS CUALES SE PRESENTA UNA BASE PERPENDICULAR AL RESTO DE BASES. ....	43

FIGURA 24 FRAGMENTOS CDDMP OPTIMIZADOS PARA EL FRAGMENTO 307D-E-A47 CON LOS TRES CAMPOS DE FUERZA, TODOS COMO HÉLICE DOBLE.....	44
FIGURA 25 FRAGMENTO CDDMP 1VBL EN CONFORMACIÓN BA09.....	47
FIGURA 26 FRAGMENTO CDDMP 1IKK EN CONFORMACIÓN BA13.....	47

## Índice de Tablas

TABLA 1 FORMATO DE UN ARCHIVO *.PDB.....	27
TABLA 2 VALORES PROMEDIOS DE LOS ÁNGULOS DE TORSIÓN DEL ESQUELETO AZÚCAR-FOSFATO PARA SEIS CONFORMACIONES.....	32
TABLA 3 SECUENCIA Y NÚMERO DE LOS FRAGMENTOS ENCONTRADOS EN LAS CONFORMACIONES CONSIDERADAS. ....	33
TABLA 4 VALORES DE LOS ÁNGULOS DE TORSIÓN DEL ESQUELETO AZÚCAR-FOSFATO, PARA LAS CUATRO ESTRUCTURAS OPTIMIZADAS, CADA UNA EN UNA CONFORMACIÓN DIFERENTE PARA LAS FAMILIAS. ....	35
TABLA 5 CARACTERÍSTICAS CONFORMACIONALES INICIALES Y CALCULADOS CON TRES CAMPOS DE FUERZAS DE AMBER PARA LA FAMILIA BB12 DE LOS CDDMPs EN VARIAS SECUENCIAS DE NUCLEÓSIDOS.....	38
TABLA 6 COMPARACIÓN DE LAS CARACTERÍSTICAS CONFORMACIONALES DE LAS ESTRUCTURAS CDDMP, DDMP Y SPB DE LA FAMILIA BB12 CALCULADOS CON EL CAMPO BSC1 Y CON ESTRUCTURAS EXPERIMENTALES.....	39
TABLA 7 COMPARACIÓN DE CARACTERÍSTICAS DE FRAGMENTOS CON SECUENCIA dCpDc DE LA FAMILIA BB12 PARA LA ESTRUCTURA INICIAL Y OPTIMIZACIONES DE DÚPLEX Y CADENA ÚNICA CALCULADOS CON EL CAMPO DE FUERZAS BSC1.....	40
TABLA 8 VALORES DE LOS ÁNGULOS DE TORSIÓN DEL ESQUELETO AZÚCAR FOSFATO Y LA SUPERPOSICIÓN DE VARIOS FRAGMENTOS REPRESENTATIVOS DE LA CONFORMACIÓN BB02 .....	45
TABLA 9 VALORES DE LOS ÁNGULOS DE TORSIÓN DEL ESQUELETO AZÚCAR FOSFATO Y LA SUPERPOSICIÓN DE FRAGMENTOS CON LA CONFORMACIÓN BA09. ....	46
TABLA 10 VALORES DE LOS ÁNGULOS DE TORSIÓN, ENERGÍA Y SUPERPOSICIÓN DE LAS BASES PARA LA ESTRUCTURA INICIAL Y LAS OPTIMIZACIONES DE FRAGMENTOS CON CONFORMACIÓN BA13.....	48

# RESUMEN

El ADN es el biopolímero más importante para la vida. En él vienen inscritas todas las instrucciones para el correcto funcionamiento de las células. Se trata de una molécula químicamente simple, aunque de gran diversidad estructural y sumamente flexible, esto le permite, entre otras cosas, los medios para su empaquetamiento dentro del núcleo de la célula. En los últimos años se han hecho estudios para comprender mejor sus características estructurales y sus mecanismos de regulación.

En concreto la estructura más estudiada es el dúplex de ADN con cadenas antiparalelas del modelo de Watson y Crick (familias A y B). Sin embargo, dada su flexibilidad la molécula de ADN puede presentarse en una gran cantidad de conformaciones que ayuda en las diferentes funciones de los ácidos nucleicos. Por tanto, en este trabajo de tesis se realizó un estudio mecánico computacional sobre las regularidades estructurales de clases conformacionales selectas de estructuras con pares de nucleósidos de Watson-Crick y diferentes regiones de ángulos de torsión utilizando métodos de mecánica molecular (MM) con ayuda del software AMBER y sus tres campos de fuerzas (CF) BSC1, OL15 Y FF99, especializados en ácidos nucleicos.

Para ello se buscaron y optimizaron fragmentos de desoxidinucléosido monofosfato complementario (cdDMP) con secuencias Purina-Purina:Pirimidina-Pirimidina, neutralizando las cargas con iones  $\text{Na}^+$ . También se trabajó con cadenas únicas (dDMP) y fragmentos mínimos de esqueletos azúcar-fosfato (SPB). Los resultados son presentados de forma cuantitativa por medio de Tablas y de forma cualitativa por medio de imágenes que ilustren las diferencias entre conformaciones.

## **OBJETIVO GENERAL:**

Búsqueda de las regularidades estructurales de fragmentos de ADN pertenecientes a diferentes familias conformacionales con secuencias de bases purina-purina: pirimidina-pirimidina vía cálculos con diferentes campos de fuerzas de MM

### **OBJETIVOS ESPECÍFICOS:**

- Realizar la búsqueda en el banco de datos de fragmentos pertenecientes a las familias selectas de los dúplex de ADN, con secuencia purina-purina: pirimidina-pirimidina. Preparar archivos de fragmentos seleccionados para el cálculo usando diferentes CF.
- Realizar la optimización de los fragmentos seleccionados aplicando los campos de fuerza AMBERFF99, AMBER OL15, AMBER BCS1, calcular y comparar las características de las estructuras iniciales y optimizadas.
- Analizar y comparar los cambios en las características conformacionales entre las estructuras experimentales y las optimizadas por los diferentes campos de fuerza, además comparar las diferencias en las regularidades de las características estructurales entre las diferentes familias conformacionales.

# Capítulo 1. Introducción

El modelo de hélice doble del ADN fue propuesto en 1953 por James Watson y Francis Crick [1] gracias a los estudios de rayos X de Rosalind Franklin y Maurice Wilkins. Este modelo lograba explicar muchos de los resultados experimentales ya conocidos. En ella se guardan todas las instrucciones para la producción de proteínas y el correcto funcionamiento de las células.

En el ADN hay segmentos que juegan un papel netamente estructural, es decir, ADN sin código, y otros segmentos cuyas secuencias contienen información para organizar y controlar funciones vitales, a estas secciones con secuencias específicas se les conoce como *genes*. El ADN funciona como molde para la síntesis de ARN, y el ARN cumple una función de traductor entre ADN y proteínas. Así, de forma indirecta, el ADN contiene instrucciones para la síntesis de proteínas. Mientras el ADN se queda en el núcleo, la síntesis de proteínas se lleva a cabo en los ribosomas del citoplasma.

Es por esta información que la estructura del ADN es tan importante y guarda un lugar invaluable dentro de las investigaciones actuales. Los primeros estudios de la hélice doble se realizaron en la década de los 50s con modelos hechos a mano [2,3], después de la invención de la computadora fueron posibles otras herramientas, como modelos mecánico-cuánticos y modelos mecánico-moleculares. Ejemplos de estas herramientas computacionales son el método de la Teoría Funcional de la Densidad, DFT por sus siglas en inglés (Density Functional Theory) y la parametrización de campos de fuerza, como AMBER o CHARMM, que son ajustados a resultados experimentales y consideraciones de mecánica cuántica.

En estudios más recientes [5, 6, 7, 8] el equipo del Dr. Poltev ha logrado demostrar, usando el método DFT, que el fragmento mínimo que aporta información estructural es el desoxidinucléosido monofosfato (dDMP). Además, se mostró la dependencia que existe entre la conformación de la cadena con las secuencias de las bases y que las características principales de las familias más conocidas de dúplex de Watson-Crick, AI y BI, están predefinidas por los mínimos locales de energía de los dDMPs. Estos resultados habrían sido imposibles de determinar solo con las técnicas experimentales disponibles.

Por otro lado, sabemos que la cadena de ADN posee una gran flexibilidad, en la última década investigadores suecos [9, 10, 11] han realizado estudios para lograr clasificar la diversidad estructural del ADN en conformaciones específicas que sirvan a un mejor entendimiento de las características y funciones biológicas del dúplex de ADN.

## 1.1 Estructura del ADN.

La macromolécula del ADN es la estructura más importante para la vida pues en las secuencias de sus bases se encuentran las instrucciones para las funciones biológicas más vitales de los seres vivos. La habilidad de replicarse a sí misma hace a esta molécula única entre el resto, y garantiza la perpetuación de la vida.

El ADN está conformado por dos hélices que contienen cadenas complementarias de nucleótidos. La unidad mínima de información estructural en las hélices dobles del ADN es el desoxidinucleósido monofosfato complementario (cdDMP). Cada nucleótido está formado por tres componentes: (1) una base nitrogenada, (2) un azúcar pentosa y (3) un grupo fosfato. Cada nucleótido del ADN puede tener una de las cuatro diferentes bases nitrogenadas canónicas: dos purinas (Adenina y Guanina,) y dos pirimidinas (Timina y Citosina), en el caso del ARN tenemos Uracilo en lugar Timina. La suma de un nucleósido y un grupo fosfato da como resultado un nucleótido. Cabe mencionar que los nucleótidos tienen varias funciones dentro de las actividades celulares, como lo son el transporte de energía (ATP, ADP) o la respiración celular (NAD, FAD).

Por convención internacional los átomos de las moléculas suelen estar numerados para identificarlos. La Figura 1 muestra la numeración de la pirimidina y de la purina. La Figura 2 muestra las 4 bases nitrogenadas canónicas de ADN, cada una conservando la numeración de la pirimidina y de la purina.

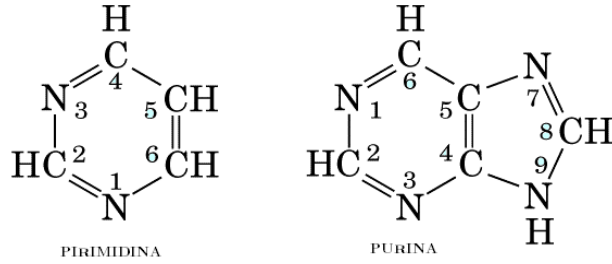


Figura 1 Numeración de átomos de pirimidina y purina.

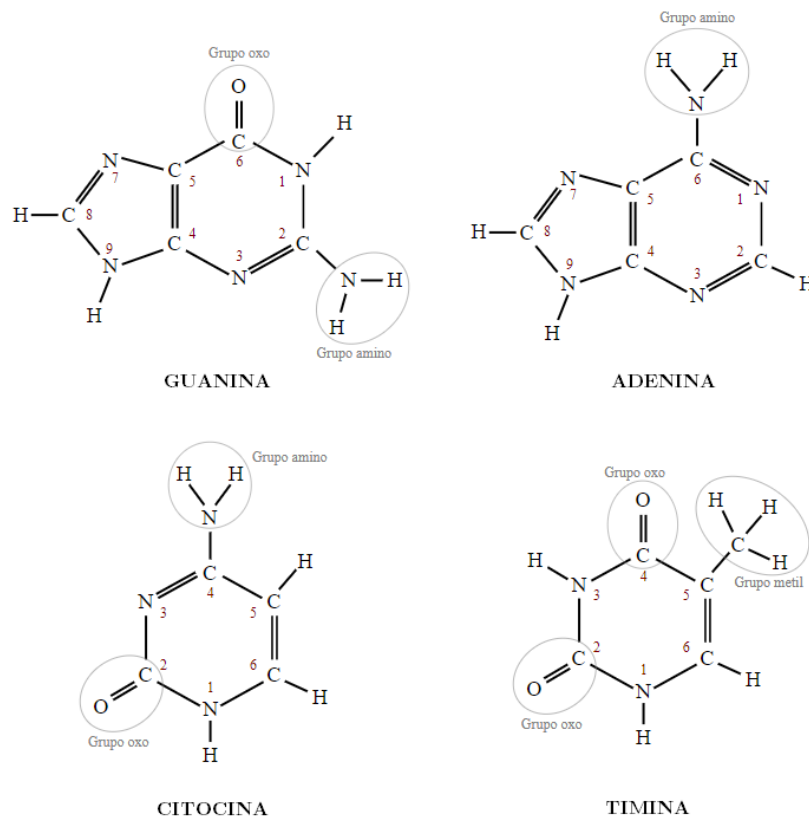


Figura 2 Bases nitrogenadas principales de los ácidos nucleicos.

En el modelo más conocido de Watson-Crick cada nucleótido de una cadena establece una asociación específica y única con la correspondiente base de la otra cadena, pues las

Adeninas solo se unen con Timinas mediante dos puentes de hidrógeno y las Guaninas solo se unen con Citosinas mediante tres puentes de hidrógeno, ver Figura 3.

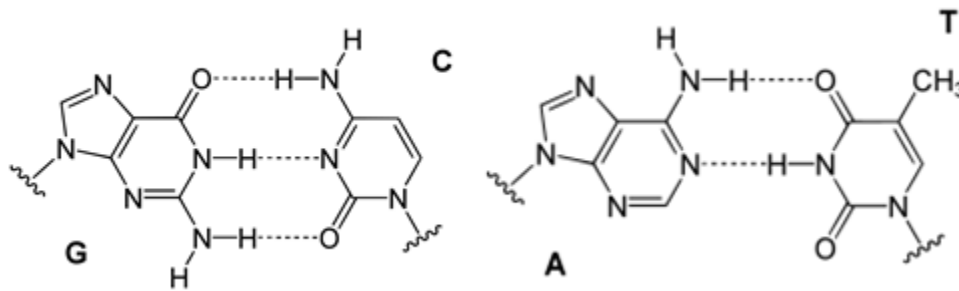


Figura 3 Bases nitrogenadas y sus pares complementarios.

Cabe mencionar que no son las únicas correspondencias entre bases que existen, otra asociación conocida son los pares de Hoosteen que tienen una conformación *syn* para las bases purinas, lo que da como resultado otras posiciones para las bases; o como el caso de Hoogsteen reverso.

El azúcar en el ADN es una pentosa cíclica que contiene un anillo con cuatro carbonos y un oxígeno (C1', C2', C3', C4', O4'), y un quinto carbono (C5') unido al carbono C4'. Esta azúcar recibe el nombre de desoxirribosa por prescindir de un grupo hidroxilo en el carbono C2', ver Figura 4. Los carbonos C5' y C3' son los que están unidos al grupo fosfato mediante enlaces covalentes. El carbono C1' se une con en nitrógeno N9 de las purinas, o con el N1 de las pirimidinas.

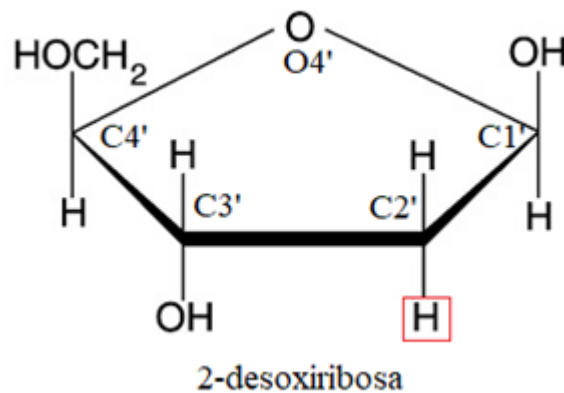


Figura 4 Azúcar desoxirribosa, presente en el ADN

Las moléculas del azúcar desoxirribosa y la base nitrogenada unidas, a través del carbono C1' con el nitrógeno N9 de las purinas o con el N1 de las pirimidinas, forman una molécula nucleósido. El nucleótido resulta de la unión de un nucleósido con un grupo fosfato mediante un enlace covalente desde el carbono C5', ver Figura 5. Cada nucleósido tiene su nombre y una designación dependiendo de la base; desoxiadenosina (dA), desoxiguanosina (dG), desoxitimidina, desoxitimidina (dT) y desoxicitidina (dC); el grupo fosfato se designa con la letra p.

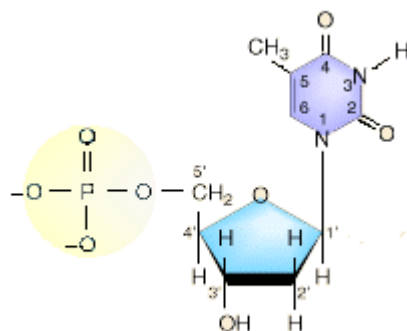


Figura 5 Desoxirribonucleótido pdT.

Desoxidinucleósido monofosfato (dDMP). Se trata del fragmento mínimo que aporta información estructural sobre la cadena de ADN. Como su nombre lo indica se trata de dos nucleósidos unidos cada uno por un enlace covalente con el grupo fosfato, a través del carbono C5' y carbono C3' respectivamente, ver Figura

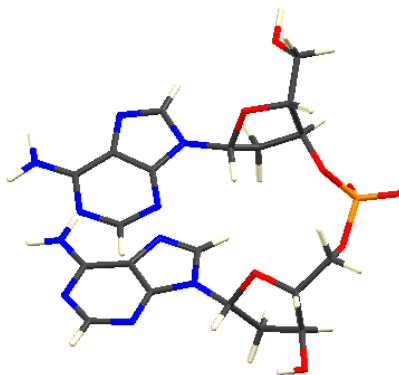
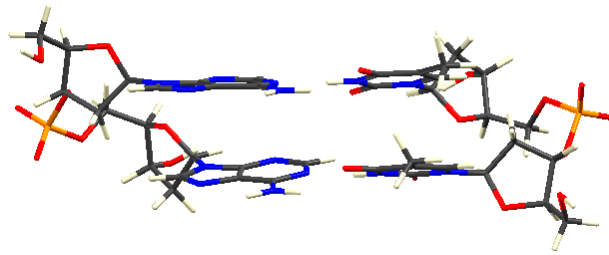


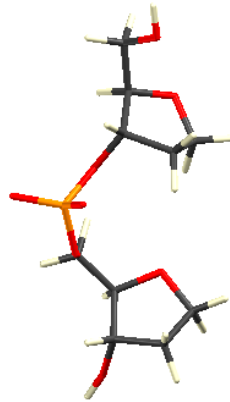
Figura 6 Desoxidinucleósido monofosfato, dApdA.

Desoxidinucleósido monofosfato complementario (cdDMP). Se trata de dos dDMP, los cuales son complementarios, A-T y G-C, unidos por medio de enlaces de hidrogeno.



*Figura 7 Desoxidinucleósido monofosfato, dApdA:dTpdT.*

Esqueleto azúcar-fosfato (Sugar Phosphate Bone SPB). Esta molécula es el resultado de remover las bases nitrogenadas de una molécula de dDMP, es decir que se trata de dos azúcares desoxirribosas unidas a un grupo fosfato.

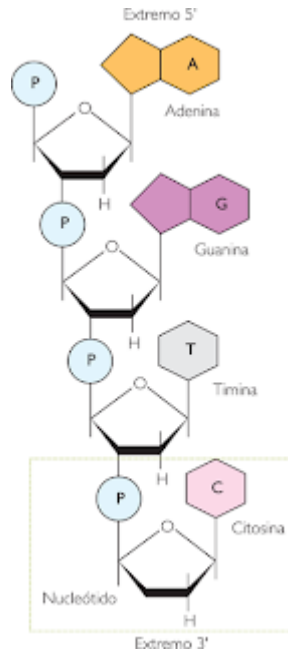


*Figura 8 Esqueleto azúcar fosfato SPB.*

## 1.2 Niveles de organización del ADN.

El primer nivel de organización del ADN consta de secuencias de nucleótidos unidos mediante un enlace fosfodiéster entre el grupo hidroxilo de C3' del primer nucleótido y el

grupo fosfato C5', esto da como resultado una cadena polinucleótida con un grupo fosfato en el extremo C5' y un hidroxilo en el extremo C3', lo que le da la dirección a la cadena de 5' fin a 3' fin.



*Figura 9 Cadena única de ADN, primer nivel de organización.*

La hélice doble es la estructura representativa del segundo nivel de organización del ADN. En el ADN estas cadenas son complementarias, antiparalelas y continuas. El ARN también presenta este nivel de organización, con la diferencia de que solo cuenta con una cadena, teniendo un inicio (5' fin) y un final (3' fin) en una región cercana. Fueron Watson y Crick quienes postularon la hélice doble para describir la estructura tridimensional del ADN.

El tercer y cuarto nivel en la organización están relacionados con la forma en que la hélice doble del ADN se enrolla sobre sí misma para dar lugar a lo que se conoce como superenrollamiento que da como resultado la cromatina y después los cromosomas.

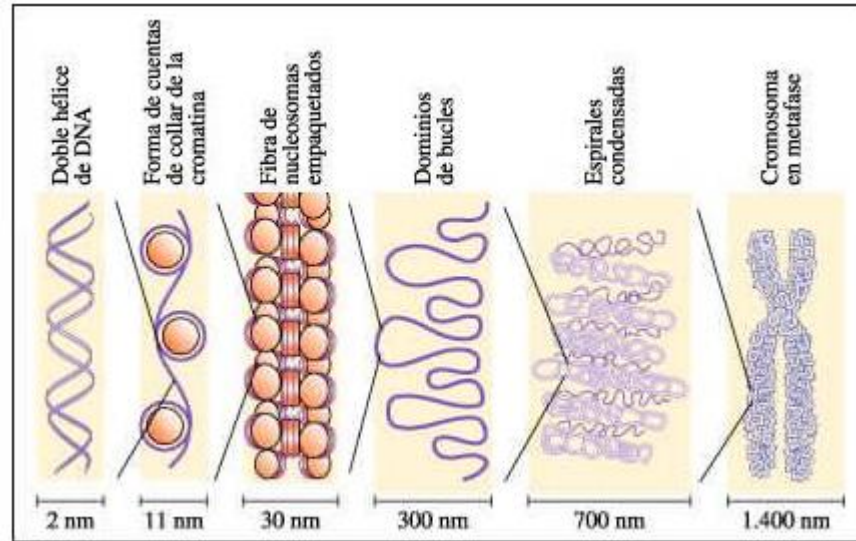


Figura 10 Niveles de organización del ADN comenzando por la hélice doble (segundo nivel) en la izquierda.

### 1.3 Estructuras secundarias del ADN.

La estructura de la hélice doble del ADN presenta algunas diferencias notorias que nos permite catalogarla en familias conformacionales del ADN, se conocen tres tipos principales: A-ADN, B-ADN, Z-ADN. El B-ADN tiene mayor interés biológico pues es, esta forma la que tiene mayor interacción con las proteínas. El A-ADN se presenta cuando se deshidrata levemente el B-ADN. El Z-ADN se diferencia por enrollarse hacia la izquierda en lugar de a la derecha y solo se logra en soluciones salinas.

A-ADN: Se encuentra en condiciones de deshidratación. Es una hélice dextrógira (se enrolla hacia la derecha). La elevación de la hélice es de aproximadamente 2.5 nm. Tiene un surco mayor mucho más estrecho y profundo que la forma B-ADN. Durante la transcripción la forma A resulta más estable que la forma B.

B-ADN: En esta estructura podemos observar un surco mayor apenas más grande que el surco menor, las bases son en general coplanarias y tenemos un eje helicoidal prácticamente perpendicular a las bases y que pasa por el centro de ellas.

Z-ADN: Esta familia presenta más restricciones para obtenerla, se presenta en secuencias específicas como G-C-G-C-G. Esta hélice es en general más delgada y alargada para el mismo número de bases con respecto a las otras dos familias. Una diferencia notoria es su disposición a enrollarse hacia la izquierda (levógira).

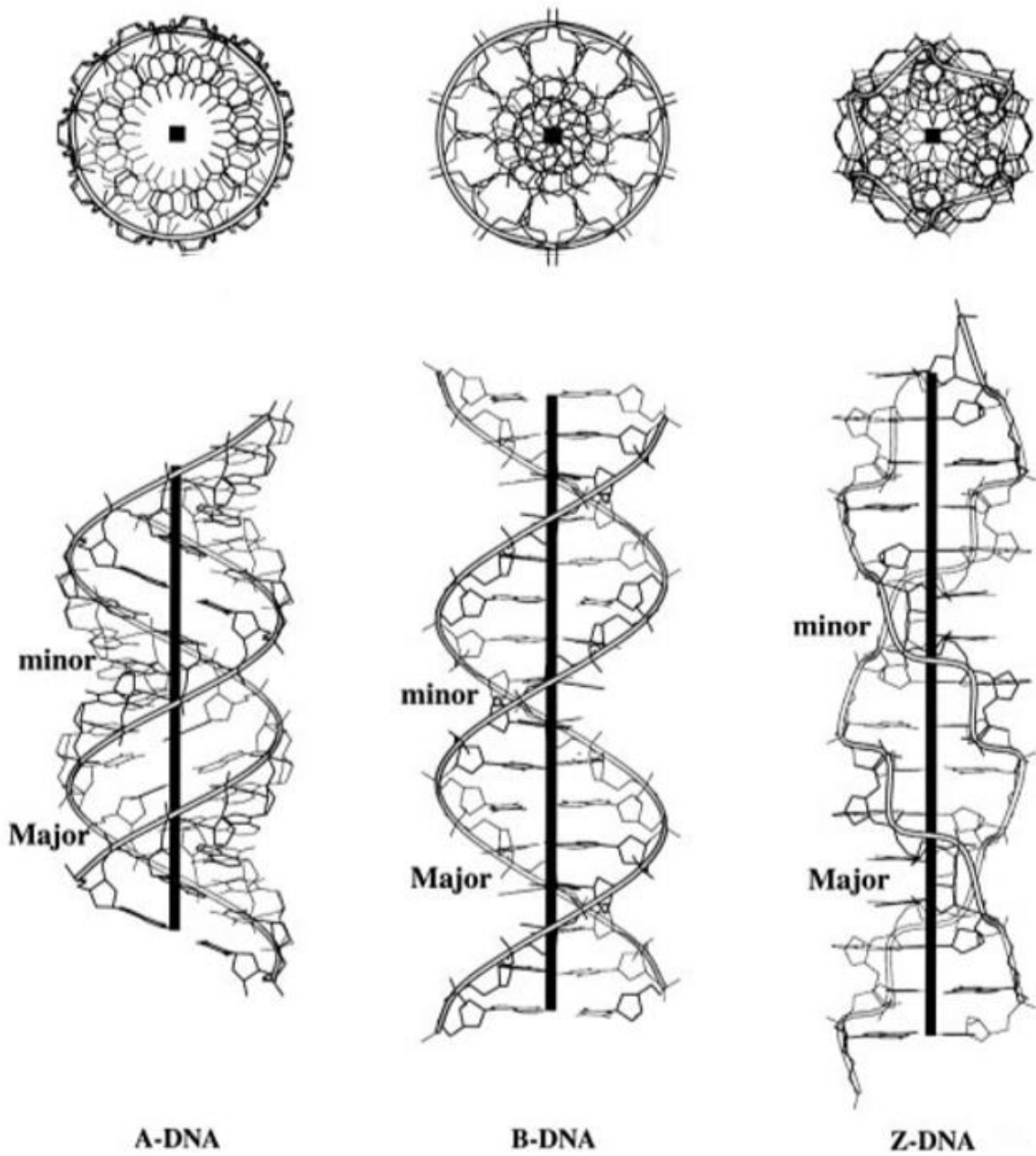


Figura 11 Moléculas de ADN con conformaciones (a) B-ADN, (B) A-ADN y (c) Z-ADN.

## 1.4 Parámetros estructurales del ADN.

Podemos valernos de distintos parámetros para el estudio estructural del ADN. La información estructural se encuentra presente en los ángulos de torsión y en el empaquetamiento del anillo de azúcar. También podemos usar los parámetros como “*Helical Twist*”, “*Helical Rise*”, “*Buckle*”, “*Propeller*” y la superposición de las bases, “*Stacking*”.

El ángulo de torsión se define como el ángulo  $\theta$  que se forma en los planos que contiene a los átomos A, B, C y B, C, D, siendo A, B, C y D átomos consecutivos unidos por enlaces, ver Figura 12. Se definen seis ángulos de torsión principales para describir la configuración del esqueleto azúcar fosfato:  $\delta$ ,  $\epsilon$ ,  $\zeta$ ,  $\alpha$ ,  $\beta$  y  $\gamma$  como se describen en la Figura 13a. En particular el ángulo  $\delta$  contiene información sobre el empaquetamiento del anillo desoxirribosa.

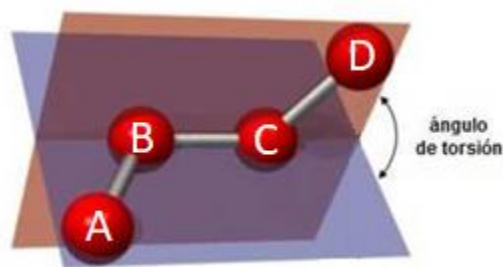


Figura 12 Ángulo de torsión formado por el plano de los átomos A, B y C con el plano de átomos B, C y D.

En el anillo de desoxirribosa uno o varios de sus átomos pueden estar fuera del plano, a estas desviaciones de los átomos del anillo de azúcar se les designa como Xn'-endo si el átomo no es coplanar y sale en la dirección de C5' o Xn'-exo si el átomo sale en sentido contrario (Figura 13b). La conformación C3'-endo y C2'-exo es característica de la familia A-ADN, mientras que las conformaciones C2'-endo y C3'-exo lo son para la familia B-ADN.

Otro ángulo de torsión importante es el ángulo glucosídico  $\chi$ , alrededor del enlace entre el átomo C1' del anillo de azúcar y el nitrógeno de su base adjunta.

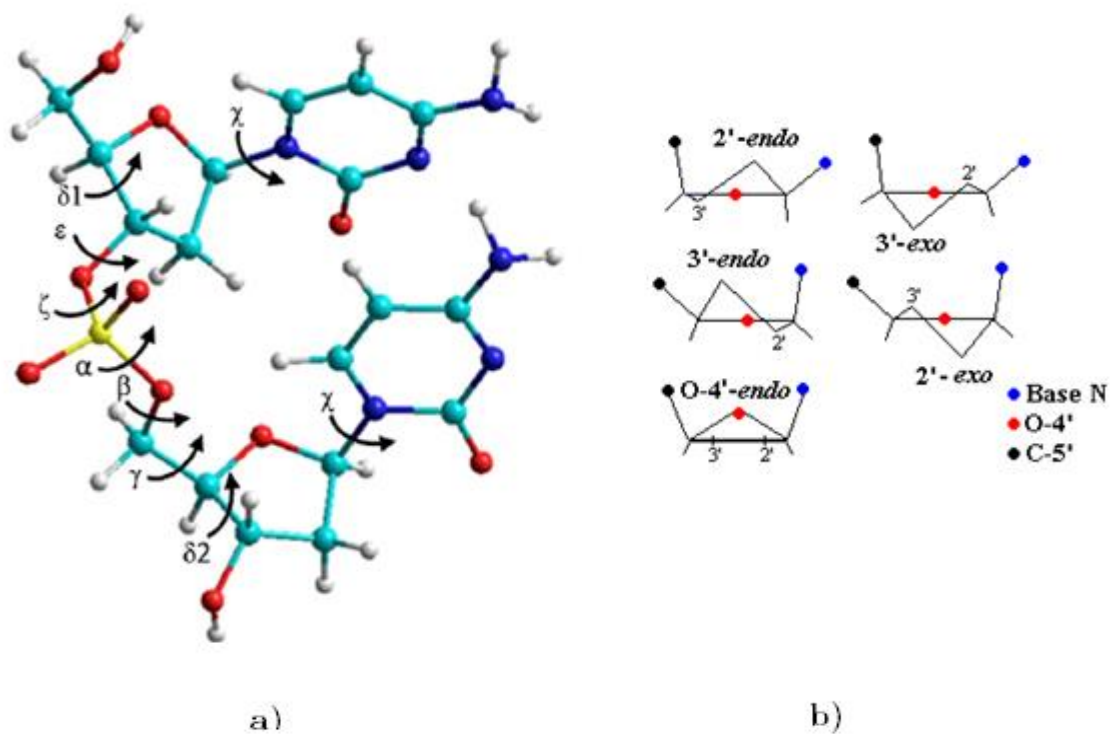


Figura 13 a) Designación de los ángulos de torsión de un dDMP (dCpdC). b) Diferentes empaquetamientos del azúcar.

Hay otros parámetros que definen características sobre la organización de las bases: “*Propeller*” define el ángulo que hay entre los planos de las bases complementarias, cuando una gira respecto de la otra alrededor del eje y (Fig. 10). “*Buckle*” define el ángulo que se produce por la doblez de los planos de las bases (alrededor del eje x). El “*Stacking*” cuantifica el área de traslape entre bases apiladas, este parámetro se mide en  $[\text{Å}^2]$ . “*Helical Twist*” es el ángulo de rotación de pares de bases sucesivos alrededor del eje helicoidal. “*Helical Rise*” es la longitud que separa a pares de bases sucesivas, el promedio de estas separaciones es de 3.4 Å. Estos parámetros permiten una clasificación rápida de las estructuras.

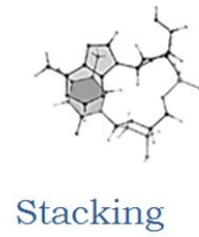
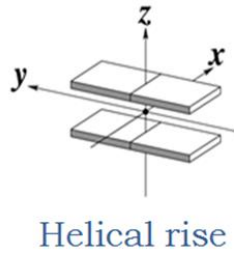
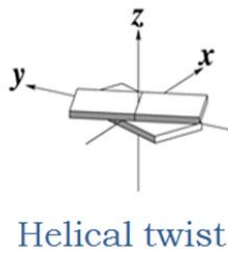
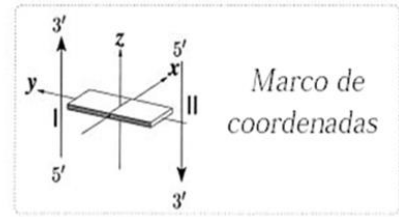
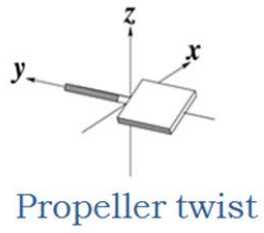
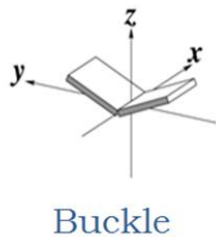


Figura 14 Descripción grafica de los parámetros utilizados para describir la geometría en los pares de bases complementarios y secuencias de pares de bases.

# Capítulo 2. Metodología

En biofísica molecular hay dos formas de estudiar las interacciones que existen entre moléculas. El método mecánico molecular y el modelo mecánico cuántico. El primero permite trabajar con moléculas grandes de una manera bastante sencilla y a bajo costo computacional pues se puede trabajar en cualquier ordenador. El segundo método, el mecánico cuántico, se basa en resolver la ecuación de Schrodinger independiente del tiempo, es decir, encontrar una función de onda que describa las interacciones de la molécula, sin embargo, esto resulta imposible para cualquier molécula que no sea el átomo de hidrógeno, por lo tanto, las soluciones a sistemas más complejos son soluciones numéricas. Este método es más correcto desde un punto de vista teórico, sin embargo, el coste computacional es alto, por lo que solo se utiliza para moléculas de pocos átomos.

## 2.1 Mecánica Molecular

Los métodos Mecánica molecular (MM) permiten el cálculo de energía de sistemas muy grandes mediante la aplicación de funciones de energía potencial átomo-átomo clásicas para la descripción de enlaces, ángulos, torsiones e interacciones no covalentes. Basado en la aproximación de Born-Oppenheimer es posible separar la parte electrónica de la nuclear y permite crear una superficie de energía potencial que depende sólo de la contribución nuclear. De esta forma la energía de la molécula depende sólo de la posición de los núcleos.

Este método permite considerar a los átomos de una molécula como puntos enlazados a los cuales se les asocia una carga dependiendo del tipo de átomo. Los enlaces entre átomos son tratados como resortes cuya posición de equilibrio depende de los átomos que se enlazan. Así, debido a las cargas de los átomos, se crea una superficie de energía potencial.

En los cálculos se considera que los átomos están dentro de un potencial generados por ellos mismos, y por medio del método del gradiente conjugado se generan ligeras

variaciones en las posiciones de los átomos de forma que se pueda llegar a un mínimo de energía.

La función de energía potencial es un conjunto de ecuaciones basadas en la mecánica clásica y con parámetros ajustados para que pueda reproducir mejor las conformaciones de los resultados experimentales. A este conjunto de funciones de energía potencial es lo que conocemos como **Campo de Fuerzas** (CF). La principal aproximación de esta función de energía potencial es la aditividad de varios términos de energía que tienen diferente naturaleza física como se muestra en la Ec. 2.1

$$U = \sum U_{enlace} + \sum U_{angulo} + \sum U_{torsión} + \sum U_{no-enlazante} \quad (1)$$

### 2.1.1 Términos enlazantes

Término de estiramiento de enlaces. Considera los enlaces como resortes que unen dos átomos, la expresión de energía a menudo se describe mediante una ecuación de oscilador armónico, ver Figura 15 a).

$$U_{enlace} = \sum K_l (l_i - l_i^0)^2 \quad (2)$$

Donde  $l_i^0$  es la longitud de enlace de equilibrio,  $l_i$  es longitud del enlace de los átomos y  $K_l$  es la constante de fuerza.

Término de flexión angular. Es la energía relacionada con el ángulo que conecta tres átomos unidos, también se describe mediante una ecuación de oscilador armónico, ver Figura 15 b).

$$U_{angulo} = \sum K_\theta (\theta_i - \theta_i^0)^2 \quad (3)$$

Donde  $\theta_i^0$  es el ángulo de equilibrio,  $\theta_i$  es el ángulo de los átomos considerados y  $K_\theta$  es la constante de fuerza.

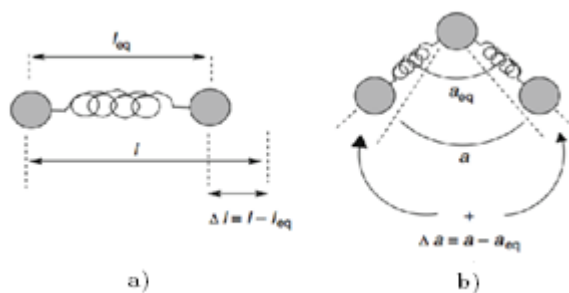


Figura 15 Representación gráfica de los términos armónicos. a) Término de enlace. b) Término de flexión angular.

Término de torsión. Describe la periodicidad de la energía potencial para rotaciones intramoleculares. La expresión de energía generalmente se define como una función seno o coseno o una combinación de ambas. En la Figura 16 se observa que si  $\Delta\Phi = 360^\circ$  se llegaría a la misma posición.

$$U_{torsión} = \sum_i^m \sum_{n=1}^3 K_{\phi_i}^n (1 - \cos n\phi_i) \quad (4)$$

$\Sigma$

Donde  $K_{\phi_i}^n$  es una constante de fuerza,  $\phi_i$  es el ángulo de torsión de los átomos considerados y  $n$  es la multiplicidad, si se trata de enlaces simples es común que  $n = 3$ , con enlaces dobles  $n = 2$ .

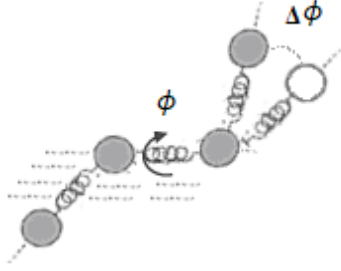


Figura 16 Representación gráfica del término de ángulos de torsión.

### 2.1.2 Términos no enlazantes

Contribución electrostática. Este término representa la atracción o repulsión entre dos átomos no unidos debido a su carga atómica generalmente representada como dos cargas puntuales.

$$U_{no-enlazante} = \sum_i^n \sum_{j < i}^n \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (5)$$

Donde  $q_i$  y  $q_j$  son las cargas de los átomos,  $\epsilon_0$  es la permitividad del vacío y  $r_{ij}$  es la separación entre los átomos  $i$  y  $j$

Contribución de Van der Waals. Este término incluye interacciones de dispersión-repulsión entre átomos no enlazados, de modo que cuando dos átomos se acercan demasiado la energía aumenta considerablemente, mientras que para distancias grandes la energía converge a cero. La relación matemática de esta contribución se expresa típicamente como un potencial de Lennard-Jones 12-6:

$$U_{no-enlazante} = \sum_i^{n-1} \sum_{j > i}^n \epsilon_{ij} \left[ \left( \frac{r_{oij}}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{oij}}{r_{ij}} \right)^6 \right] \quad (6)$$

Aquí  $\epsilon_{ij}$  representa la energía mínima de interacción entre los átomos  $i$  y  $j$ ,  $r_{oij}$  la distancia a la que se da el mínimo de energía y  $r_{ij}$  es la separación entre átomos.

## 2.2 Campo de fuerzas AMBER (Assisted Model Building with Energy Refinement)

El campo de fuerzas AMBER (Assisted Model Building with Energy Refinement) se ha utilizado ampliamente para ácidos nucleicos desde la década de 1990 [12,13]. Este Campo de Fuerzas ha pasado por varias modificaciones desde su creación y diferentes grupos de investigación han descubierto, corregido y reevaluado múltiples aspectos y parámetros [14]. Así, para los fines de este trabajo se utilizaron los campos OL15, BSC1 Y FF99, pues el conjunto de parámetros que tiene cada uno están ajustados para el modelado preciso de la hélice doble de ADN.

Amber necesita un archivo de extensión \*.pdb para poder realizar los cálculos de optimización. Esta extensión es ampliamente usada dentro de la química molecular para describir y visualizar la estructura 3D tanto de ácidos nucleicos como de proteínas. La estructura de un archivo \*.pdb seria:

*Tabla 1 Formato de un archivo \*.pdb.*

COLUMNAS	TIPO DE VALOR	CAMPO	DEFINICIÓN
1 - 6	Record name	"ATOM "	
7 - 11	Integer	serial	Numero serial del átomo
13 - 16	Atom	name	Nombre del átomo.
17	Character	altLoc	Indicador de localización alternativo.
18 - 20	Residue name	resName	Nombre del residuo.
22	Character	chainID	Identificador de cadena.
23 - 26	Integer	resSeq	Numero de secuencia de residuos.
31 - 38	Real	x	Coordenada ortogonal para x en A
39 - 46	Real	y	Coordenada ortogonal para

			y en A
47 - 54	Real	z	Coordenada ortogonal para z en A
55 - 60	Real	occupancy	Ocupación
61 - 66	Real	tempFactor	Factor de temperatura
73 - 76	LString(4)	segID	Identificador de segmento
77 - 78	LString(2)	element	Símbolo del elemento
79 - 80	LString(2)	charge	Carga del átomo.

A continuación, se presenta una serie de pasos para el uso de este software.

1. Se crea un archivo con extensión \*.pdb que contenga los átomos de las moléculas a estudiar y sus coordenadas, para que el programa pueda identificar de qué átomos se trata, estos también deben estar etiquetados con un número, que en este caso corresponden a la convención internacional.
2. Se visualizan las moléculas con ayuda de la aplicación *xleap*, se añaden sodios Na+ e hidrógenos de ser necesarios y se crean por separado dos archivos con extensiones \*.tipo y \*.coor, que contienen los tipos y coordenadas de los átomos respectivamente.
3. Con ayuda de la aplicación *sander*, se le indica al programa que realice los cálculos de las energías en varias posiciones hasta que encuentre la posición de energía optima. Los valores de las energías calculadas en cada paso se muestran en un archivo de salida de extensión \*.out y las coordenadas finales, correspondientes al mínimo, quedan contenidas en un archivo de extensión \*.outcoor.
4. Se crea un archivo \*.pdb con las coordenadas finales para poder visualizar la molécula y ver sus características geométricas.

### 2.3 Nucleic Acid Database (NDB) y Protein Data Bank (PDB)

Nucleic Acid Database y Protein Data Bank [16] son las bases de datos de ácidos nucleicos y proteínas, en las que se archivan las coordenadas atómicas de las estructuras

experimentales obtenidas por métodos como difracción de rayos X en cristales de ADN o Resonancia Magnética Nuclear, NMR por sus siglas en inglés, “*Nuclear Magnetic Resonance*”. PDB es como un almacén de datos mientras que NDB funciona como un catálogo y muestra algunas propiedades de las estructuras, en ese sentido la organización de NDB es más eficiente pues solo contiene estructuras de ADN y ARN, y los algoritmos de búsqueda permiten localizar estructuras con secuencias y conformaciones específicas, y no solo por el nombre del cristal como en PDB. Sin embargo, para poder descargar un archivo en formato PDB hay que seguir un link que nos lleva a PDB para poder descargarlo. Los archivos descargados se designan por cuatro caracteres alfanuméricos iniciando siempre con un número, por ejemplo, 2VBO. En adelante haremos uso de los nombres de los cristales NDB.

En los últimos años un equipo del laboratorio de informática y química de la Universidad de Química y Tecnología en Praga ha creado un alfabeto estructural del ADN, CANA (Conformational Alphabet of Nucleic Acids) analizando casi 120000 dinucleótidos en más de 2000 estructuras cristalinas de ácidos nucleicos; definiendo 96+1 clases conformacionales NtC de CANA, que son usadas como descriptores cualitativos de las diferencias en las características estructurales de los ácidos nucleicos.

## 2.4 Software en línea 3DNA

3DNA es un software en línea [17], versátil e integrado para el análisis, la reconstrucción y la visualización de estructuras tridimensionales que contienen ácidos nucleicos a partir de un archivo PDB, 3DNA maneja todos los pares de bases no canónicos y hélices apiladas coaxialmente. Dentro del análisis que realiza 3DNA identifica y clasifica las interacciones que hay en las bases y el esqueleto azúcar-fosfato. El software arroja un archivo de salida descargable con los valores que describen la geometría del sistema.

## 2.5 Software en línea DNATCO

DNATCO es un software en línea que proporciona una interfaz para algoritmos basados en geometría que asignan clases de conformación a dinucleótidos [18]. Estas conformaciones locales permiten ir más allá de las conformaciones tradicionales en forma de A-ADN, B-

ADN Y Z-ADN. Hoy en día DNATCO asigna 44 clases conformacionales diferentes esto, permite una mejor comprensión de la estructura del ADN. El sitio web es gratuito y requiere de la estructura de ADN cargada en un formato PDB.

## Capítulo 3. Resultados

Como se ha mencionado antes, la estructura del ADN del modelo de WC puede existir en diferentes conformaciones dada su flexibilidad. Para este trabajo de tesis se utilizó la nomenclatura NtC de Svozil [10] para designar las conformaciones de dinucleótidos; así, conformaciones como BI o AI se designa por BB00 y AA00 respectivamente.

Para nuestros estudios fueron seleccionados fragmentos mínimos cdDMP del banco de datos NDB, con conformaciones BB02, BB12, BA09 y BA13 para cálculos vía MM. Estas conformaciones tienen cambios considerables en las regiones de los ángulos de torsión a los de la conformación B1 (BB00), la cual, se ha demostrado, corresponde a un mínimo de energía. Estos fragmentos mínimos fueron escogidos de cristales experimentales tales, que su secuencia fuera Purina-Purina o Pirimidina-Pirimidina.

Estudios recientes [8] han mostrado que el esqueleto SPB en conformación BB12 corresponde a un mínimo local de energía diferente del de B1. Caso contrario sucede con la conformación BB02, la cual tiene ángulos de torsión en regiones muy diferentes a cualquier mínimo de energía conocido, aun así, es una conformación bastante frecuente en los cristales de NDB.

En la Tabla 1 se puede observar las diferencias cuantitativas en los distintos ángulos de torsión del esqueleto azúcar-fosfato para las 4 conformaciones estudiadas y para las conformaciones BB00 y BB07, comenzando por el ángulo  $\delta_1$  del primer nucleósido y terminando en el ángulo delta del segundo nucleósido; además de agregar los dos ángulos glicosídicos de las bases  $\chi_1$  y  $\chi_2$ , uno por cada base nitrogenada.

Tabla 2 Valores promedios de los ángulos de torsión del esqueleto azúcar-fosfato para seis conformaciones.

	$\delta 1$	$\varepsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	$\chi 1$	$\chi 2$
BI (BB00)	138	183	259	303	180	44	138	252	258
BII (BB07)	144	245	171	297	1411	46	141	270	259
BB12	140	196	287	248	73	172	144	263	212
BB02	140	194	246	32	194	297	150	252	253
BA09	135	199	287	254	71	168	86	265	187
BA13	142	231	196	74	232	196	89	266	199

### 3.1 Búsqueda de fragmentos mínimos de hélice doble en NADB y análisis de sus características estructurales.

Se hizo una búsqueda en el banco de datos NDB de fragmentos mínimos de diferentes cristales, obtenidos por difracción de rayos X, que pertenecieran a diferentes familias, como lo son las familias BB12, BB02, BA09 y BA13. En particular basándonos en el trabajo de Poltev [8] y Svozil [9] sabemos que la familia BB02 es la que más estructuras experimentales tiene y que la familia BB12 es una buena candidata para ser un mínimo local de energía. Basados en este hecho se buscaron y analizaron aproximadamente 130 cristales que contenían fragmentos en conformación BB02 y se encontraron 160 fragmentos de cdDMP con diferentes secuencias. Estos fragmentos tienen una cadena con secuencia purina-purina y la otra cadena con secuencia pirimidina-pirimidina.

La mayoría de estas estructuras se encontraban en complejo con proteínas como “zinc fingers” o proteínas de nucleosomas. Sin embargo, también se encontraron estructuras que eran puramente ADN con algunas moléculas de anti-iones y agua; de estas últimas encontramos 12 cristales y en ellos 27 fragmentos de cdDMP. A continuación, en la Tabla 3.1 se presenta la secuencia y el número de cdDMP encontrados para las diferentes conformaciones estudiadas. Podemos observar que, de las secuencias consideradas, la familia BB02 pareciera tener cierta preferencia por la secuencia dApdA pues es la que presenta

mayor número de fragmentos, con un total 35 fragmentos. Y con solo 10 fragmentos, la secuencia dCpdT es la que menor preferencia tiene por esta conformación. Para el resto de las familias no se estudiaron suficientes estructuras para poder concluir que secuencia prefiere cada una.

*Tabla 3 Secuencia y número de los fragmentos encontrados en las conformaciones consideradas.*

Secuencia*	BB12	BB02	BA09	BA13
dApdA	6	35	1	
dApdG	4	18	1	
dGpdG	3	11	1	
dGpdA	5	11	1	4
dCpdC	6	17		
dCpdT	1	10		
dTpdT	4	31		1
dTpdC	4	26		
Total		159		

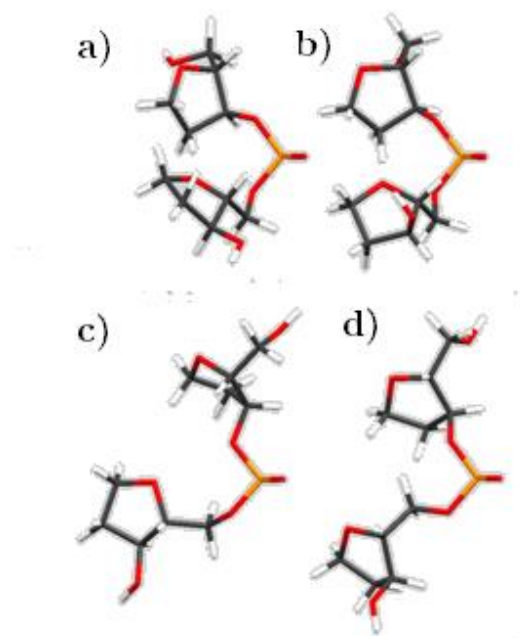
\*Se usa la designación abreviada de las secuencias de la sección 1.1.

Otro resultado que obtuvimos del análisis de los fragmentos fue la preferencia que tiene la cadena complementaria a la cadena BB02. En la mayoría de los casos la cadena complementaria presenta una conformación BI, aunque también se encontraron familias BB02, AA02, AII, BB07, BB04.

También se buscaron fragmentos para otras familias, prestando mayor atención en aquellos cristales que no tuvieran complejo ni alguna otra proteína. Para la familia BB12 se encontraron 40 estructuras de cdDMP. En el caso de la familia BA13 se hallaron tres cristales que no contenían complejo, estos son 1D49, 1IKK, 1SK5. Para la conformación BA09 los resultados experimentales son muy pocos y nos fue imposible encontrar cristales que no estuvieran en complejo con proteínas, sin embargo, aun así, encontramos 4 cristales, los cuales serían 1LMB, 1WTE, 2VBL.

### 3.2 Mínimos locales de energía para fragmentos del esqueleto azúcar-fosfato y anillos desoxirribosa pertenecientes a diferentes conformaciones.

Después de la búsqueda y selección de fragmentos, y la preparación de los archivos de coordenadas. Se seleccionaron fragmentos representativos de cada familia para optimizar y estudiar su esqueleto azúcar-fosfato y su anillo desoxirribosa; esto se hizo para conocer el aporte energético que estas puedan tener a la estructura final. La Figura 17 muestra de forma cualitativa las diferencias entre las cuatro conformaciones de fragmentos SPB.



*Figura 17 Diferencias entre los esqueletos azúcar fosfato en diferentes conformaciones a) BA09 b) BB12 c) BA13 y d) BB02.*

Para los esqueletos azúcar-fosfato se hicieron optimizaciones para las 4 familias antes mencionadas, es decir, las familias BB12, BB02, BA09 y BA13.

Para la familia BB02 se hicieron optimización de 4 esqueletos azúcar-fosfato cuyos cristales no contaban con complejos. Los resultados mostraron una gran similitud con el esqueleto de la estructura inicial para los campos BSC1 y FF99, sin embargo, el campo de fuerzas

OL15 generaba una interacción entre el sodio y el oxígeno O5'. Esta interacción con el sodio se repitió para los SPB de las estructuras con familia BA13.

Al igual que con la familia BB02, para los SPB de familia BA13 solo se hicieron optimizaciones con fragmentos de cristales en los cuales no había interacciones con otras moléculas. Sin embargo, los SPB optimizados se alejan bastante de la conformación de la estructura inicial, esto sucede con cada campo de fuerzas utilizado.

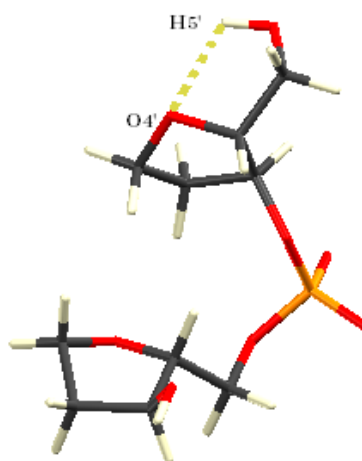
En cuanto a las familias BA09 y BB12 (que solo cambian en el ángulo de torsión  $\delta 2$ ) las optimizaciones del esqueleto azúcar-fosfato no se alejan mucho de la conformación de la estructura inicial. Sin embargo, los campos de fuerza OL15 y FF99 cambian la región del ángulo  $\delta 2$  para la familia BA09, de forma que el mínimo local para ambas familias es el mismo; esto no pasó con el campo BSC1 que si respetó la conformación inicial. La tabla 4 muestra valores de los ángulos de torsión de estructuras SPB iniciales y optimizadas, con diferente conformación.

Tabla 4 Valores de los ángulos de torsión del esqueleto azúcar-fosfato, para las cuatro estructuras optimizadas, cada una en una conformación diferente para las familias.

		$\delta 1$	$E$	$\zeta$	$\alpha$	$\beta$	$\Gamma$	$\delta 2^*$
<b>BB12</b>	dCpdT	139.9	199	282.1	240	73.4	164.4	149.5
	OL15	108.7	181.5	269.3	291.5	88.3	181.5	110.8
<b>ILMB</b>	BSC1	129.5	193.4	275	283	73.9	185.3	137.8
	FF99	110.8	187.3	273.7	286.4	73.6	190	111.8
<b>BA09</b>	dApdG	139.2	202.8	306.4	250.9	68.2	168	78.6
	OL15	108.7	181.5	269.3	291.5	88.3	181.6	110.8
<b>IWTE</b>	BSC1	131.4	195.4	275.3	281.5	73	182.7	78.3
	FF99	110.7	187.3	273.8	286.3	73.5	190	111.8
<b>BA13</b>	dApdA	141	225.1	189.6	75.8	237.2	194.8	87.5
	OL15	107.9	197.9	226	72.1	251.7	185.1	102.8
<b>IHKK</b>	BSC1	132.6	211.7	179.7	75.3	294.7	199.9	77.6
	FF99	116.2	202.6	192.8	72.7	290	210.7	98.3
<b>BB02</b>	dTpdC	135.7	195.4	249.3	31.2	197.4	295.2	147.1
	OL15	81.1	266.8	201.1	68.1	189.4	300.8	136.3
<b>428F</b>	BSC1	130.2	192.3	236.7	62.2	179.1	300.9	147.2
	FF99	110.1	196.1	229.4	63.1	178.9	298	136.5

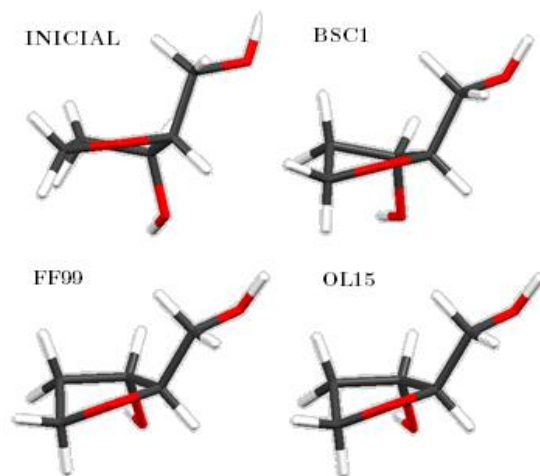
\*Todos los ángulos están expresados en sistema hexadecimal.

Una falla común en las optimizaciones del esqueleto azúcar-fosfato fue la interacción tipo enlace de hidrógeno entre el átomo de hidrogeno HO5' y el oxígeno O4', ver Figura 18.



*Figura 18 Fragmento SPB 1wet que presenta interacción H5'---O4' después de ser optimizado con CF OL15.*

Además, se hicieron optimizaciones del anillo de azúcar con los tres campos de fuerza. Tomando solo el anillo de la familia BB02, y obteniendo como resultado la conformación ya conocida del azúcar para la familia BI, ver Figura 19, en donde C2' se encuentra en conformación *endo*, ver Figura 13. Aquí hay que señalar que, aunque los tres campos de fuerza dieron el mismo resultado, los campos de fuerza OL15 y FF99 también daban esa conformación para el carbono C3', lo cual es poco frecuente en los resultados experimentales.



*Figura 19 Estructura del anillo de azúcar de un fragmento en conformación BB02. Anillo inicial y anillos optimizados con los campos de fuerza BSC1, FF99 y OL15.*

### 3.3 Fragmentos mínimos de hélice doble optimizados con los tres campos de fuerza AMBER

#### 3.3.1 Fragmentos en conformación BB12

Se continuo con la búsqueda de fragmentos en configuración BB12 ya que en un trabajo más reciente de Svozil [10] sobre la clasificación de diferentes conformaciones de ADN se reporta la existencia de más fragmentos en conformación BB12 en comparación con su trabajo anterior. Logrando encontrar 40 estructuras de dinucleótido con esa conformación en todas las secuencias Purina-Purina y Pirimidina-Pirimidina, la mayoría de estas estructuras se encontraban en complejos con proteínas. De las estructuras encontradas se optimizaron solo de algunos cristales seleccionados: 2VBJ, 2VBN, 2VBL, 2VBO, 3BAM. Encontramos que estos fragmentos tenían en su mayoría cadenas complementarias en conformación BB12 a excepción de la secuencia dGpdG cuya cadena complementaria es BB00 (BI). A continuación, se presenta la Tabla 5 de resultados, después de optimizar con tres campos de fuerzas (BSC1, OL15 y FF99) del software Amber. En la Tabla 5 se presentan los ángulos de torsión, la superposición de bases apiladas, el empaquetamiento de azúcar desoxirribosa y dos parámetros extra,  $\Sigma\Delta$  que representa la suma de valores

absolutos de las diferencias entre el valor de ángulos y el valor promedio de todas estructuras en el banco de datos.  $SD$ , es la raíz cuadrada de la suma de los cuadrados de las diferencias en  $\Sigma\Delta$ .

Tabla 5 Características conformacionales iniciales y calculados con tres campos de fuerzas de AMBER para la familia BB12 de los cdDMPs en varias secuencias de nucleósidos.

[11]		$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	$\chi 1$	$\chi 2$	$P^5$	$P^3$	St1	St2	$\Sigma\Delta$	SD
		<b>140</b>	<b>196</b>	<b>287</b>	<b>248</b>	<b>73</b>	<b>172</b>	<b>144</b>	<b>263</b>	<b>212</b>						
dApdA BB12	ini	151	189	268	265	77	176	141	283	210	165	158	2.84	0.46	87	36
2VBN-	BSC1	130	196	275	277	65	182	144	269	184	137	152	3.92	0.46	103	45
E-A13	OL15	119	185	270	284	79	179	133	260	190	123	137	3.72	0.43	134	53
	FF99	139	200	286	271	63	185	141	268	186	127	140	3.6	0.37	86	39
dApdG BB12	ini	140	200	278	263	65	175	146	275	201	153	157	4.74	0.03	65	26
2PI0-F-	BSC1	141	201	282	271	64	181	145	250	181	153	154	2.72	0.0	97	43
A22	OL15	121	187	270	282	77	177	136	266	184	125	142	3.53	0.0	126	53
	FF99	131	194	281	276	66	188	140	260	177	139	146	2.38	0.0	108	48
dGpdA BB12	ini	148	185	272	265	77	172	137	284	210	125	180	2.98	0.59	85	35
2VBL-	BSC1	132	196	275	276	66	183	145	262	191	159	152	3.7	0.4	89	40
C-G13	OL15	121	187	270	283	79	178	135	268	185	155	139	3.51	0.43	133	54
	FF99	124	190	275	281	68	187	137	269	188	153	143	3.37	0.41	124	49
dGpdG BB12	ini	148	179	295	243	80	180	145	262	195	171	178	3.45	0.0	114	61
3BAM	BSC1	141	200	282	273	67	185	147	250	172	154	154	0.97	0.12	147	76
_C_G4	OL15	123	185	271	282	80	180	139	266	172	128	146	0.85	0.0	154	74
	FF99	125	189	276	280	69	191	142	264	173	131	149	0.74	0.0	170	76
dCpdC BB12	ini	150	194	285	254	73	171	151	273	212	176	172	2.2	0.44	40	17
2O61-	BSC1	143	203	286	267	64	180	147	243	192	155	156	0.54	1.2	89	37
F-C2	OL15	138	198	280	272	72	178	138	252	188	150	143	0.99	1.86	84	37
	FF99	137	199	286	270	63	184	143	247	192	147	150	0.76	1.8	88	37
dCpdT BB12	ini	140	199	282	240	73	164	150	260	213	157	181	0.16	2.64	34	14
1PUF-	BSC1	124	204	284	266	62	178	147	251	192	160	157	0.35	3.86	97	37
D-C2	OL15	139	199	280	271	68	177	140	244	197	158	147	0.34	3.73	82	35
	FF99	138	200	285	270	61	182	144	246	187	155	152	0.2	3.46	94	41
dTpdC	Ini	117	195	274	273	65	171	148	262	210	190	126	0.62	2.84	78	38
2VBJ-	BSC1	143	204	284	268	63	180	145	244	194	152	159	0.4	3.7	90	36
C-G13	OL15	140	200	281	271	72	180	135	249	186	139	155	0.43	3.51	91	40
	FF99	139	200	286	270	63	186	139	247	192	143	153	0.41	3.37	93	38
dTpdT 2VBN-	Ini	118	197	263	278	67	165	152	262	214	129	188	0.56	0.56	101	46
E-T11	BSC1	143	203	284	268	63	179	146	244	194	160	155	0.46	3.92	89	36
	OL15	140	200	281	271	71	178	138	249	188	155	144	0.43	3.72	85	38
	FF99	139	200	286	271	63	185	141	248	190	154	147	0.36	0.36	92	39

$\delta 1$ ,  $\epsilon$ ,  $\zeta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta 2$ . Son los ángulos de torsión del esqueleto azúcar fosfato SPB.  $\chi 1$ ,  $\chi 2$ . Son los ángulos glucosídicos.  $P^5$ ,  $P^3$  Es el ángulo de pseudorrotación del anillo de azúcar. St1, St2. Es el Steacking entre pares de bases, se mide en Angstroms cuadrados [ $\text{\AA}^2$ ].

Se puede notar que en mayoría de los casos el campo de fuerzas BSC1 tiene el parámetro  $\Sigma\Delta$  menor que el resto de los campos. Esto nos sugiere que el campo BSC1 reproduce mejor las características conformacionales iniciales de la estructura experimental.

En la Tabla 6 se presentan la comparación de resultados para los fragmentos cdDMP, dDMP.y SPB calculados con el campo de fuerzas BSC1 con los mismos parámetros que en la Tabla 5.

Tabla 6 Comparación de las características conformacionales de las estructuras cdDMP, dDMP y SPB de la familia BB12 calculados con el campo BSC1 y con estructuras experimentales.

		$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	$\chi 1$	$\chi 2$	$P^5$	$P^3$	St1	St2	$\Sigma\Delta$	SD
[11]		<b>140</b>	<b>196</b>	<b>287</b>	<b>248</b>	<b>73</b>	<b>172</b>	<b>144</b>	<b>263</b>	<b>212</b>					77	40
SPB		<b>131</b>	<b>194</b>	<b>277</b>	<b>282</b>	<b>74</b>	<b>186</b>	<b>137</b>								
dApdA	Ini	151	189	268	265	77	176	141	283	210	165	158	2.84	0.46	87	36
2VBN-E-A13	cdDMP	130	196	275	277	65	182	144	269	184	137	152	3.92	0.46	103	45
	dDMP	141	202	281	271	63	178	146	252	193	155	156	4.28		84	35
dApdG	Ini	138	197	304	224	76	166	162	145	217	183	157	2.64	0.16	264	143
1PUF-D-C2	cdDMP	142	201	281	273	64	182	142	253	188	149	158	3.86	0.35	93	39
	dDMP	141	201	281	272	64	182	142	254	191	156	149	3.98		87	37
dGpdA	Ini	148	185	272	265	77	172	137	284	210	125	180	2.98	0.59	85	35
2VBL-E-T11	cdDMP	132	196	275	276	66	183	145	262	191	159	152	3.7	0.4	89	40
	dDMP	144	202	283	269	62	180	146	255	193	165	153	3.34		83	34
dGpdG	Ini	148	179	295	243	80	180	145	262	195	171	178	3.99	0.0	114	61
2VBN-T-G16	cdDMP	141	200	282	273	67	186	147	250	172	154	154	2.92	0.0	147	76
	dDMP	142	202	279	271	70	163	158	248	230	151	197	0.5		118	58
dCpdT	Ini	140	199	282	240	73	164	150	260	213	157	181	0.16	2.64	34	14
1PUF-D-C2	cdDMP	124	204	284	266	62	178	147	251	192	160	157	0.35	3.86	97	37
	dDMP	142	205	283	268	64	180	144	235	198	156	151	0.30		94	40
dTpdC	Ini	117	195	274	273	65	171	148	262	210	190	126	0.62	2.84	78	38
2VBJ-C-G13	cdDMP	143	204	284	268	63	180	145	244	194	152	159	0.4	3.7	90	36
	dDMP	143	204	284	267	63	176	147	242	204	158	159	1.40		79	33
dTpdT	Ini	118	197	263	278	67	165	152	262	214	129	188	0.56	0.56	101	46
2VBN-E-T11	cdDMP	143	203	284	268	63	179	146	244	194	160	155	0.46	3.92	89	36
	dDMP	142	201	284	266	64	171	158	245	220	150	189	3.92		78	32

$\delta 1$ ,  $\epsilon$ ,  $\zeta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta 2$ . Son los ángulos de torsión del esqueleto azúcar fosfato SPB.  $\chi 1$ ,  $\chi 2$ . Son los ángulos glucosídicos.  $P^5$ ,  $P^3$ . Es el ángulo de pseudorrotación del anillo de azúcar.  $St1$ ,  $St2$ . Es el Steacking entre pares de bases, se mide en Angstroms cuadrados [ $\text{\AA}^2$ ]

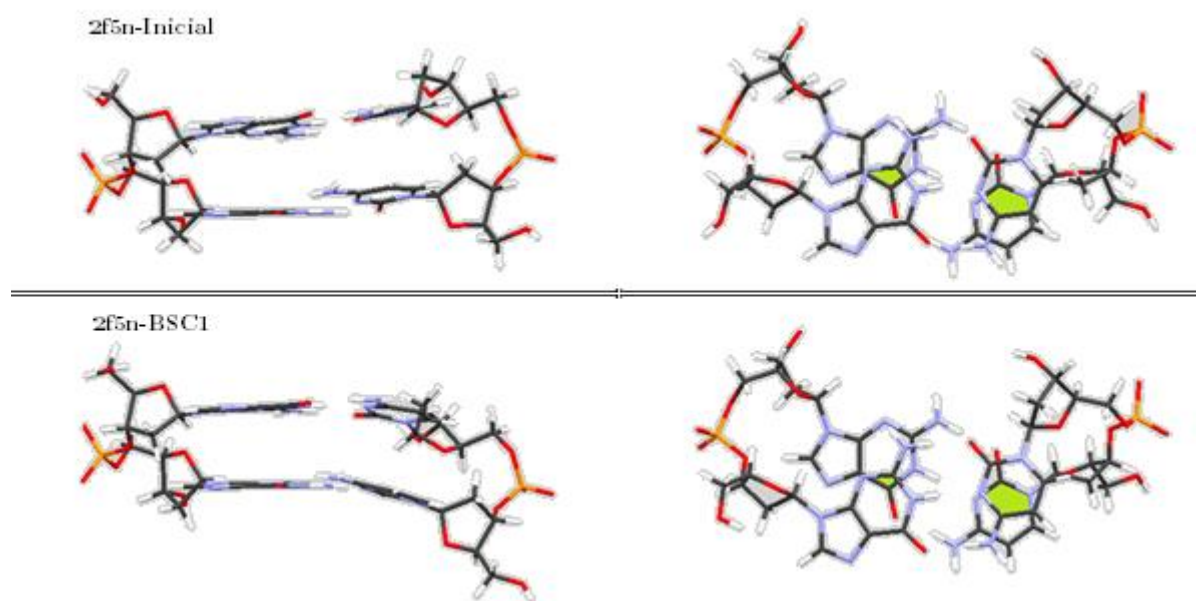
Finalmente presentamos una comparación de resultados en la Tabla 7 para secuencias dCpdC para fragmentos de diferentes cristales los cuales tienen una conformación BB12 para la cadena principal y diferente para la otra cadena. Cabe mencionar que los dúplex cuya cadena complementaria no pertenecen a la conformación BB00 (Familia B1) ya que no respetan las reglas generales de los fragmentos mínimos del modelo de Watson y Crick

Tabla 7 Comparación de características de fragmentos con secuencia dCpdC de la familia BB12 para la estructura inicial y optimizaciones de dúplex y cadena única calculados con el campo de fuerzas BSC1.

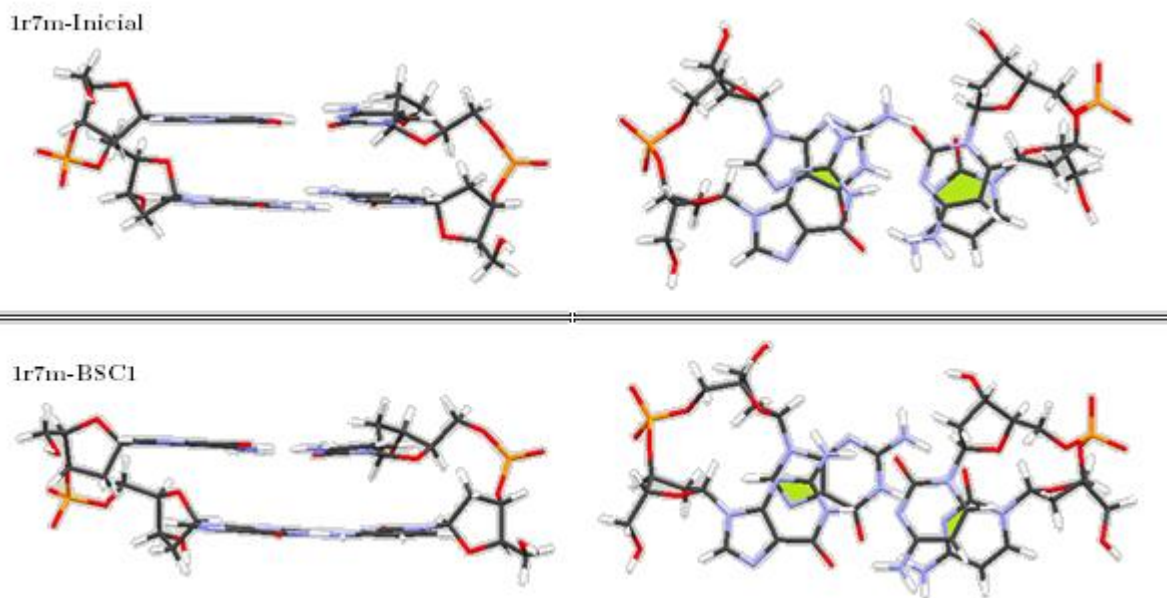
1R7M	Familia	$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	St	$\chi 1$	$\chi 2$
Inicial	BB12	150	194	285	254	73	171	151		2.20	273	212
	BB00	150	179	252	307	190	40	151		0.44	251	260
BSC1	BB12	143	203	286	266	64	179	147	-859.7	0.54	242.7	191.7
	BB00	143	185	272	290	183	47	138		1.20	224.4	237.6
CadUni	BB12	140	201	286	268	67	179	146	-435.6	1.60	243	192
1ZR4	Familia	$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	S	$\chi 1$	$\chi 2$
Inicial	BB12	139	197	293	231	68	171	118		0.10	260.1	195.5
	NANT	138	222	253	63	228	237	126		4.29	234.6	214.4
BSC1	BA09	142	205	294	256	58	167	81	-848.1	0.66	243.2	188.5
	NANT	148	215	288	89	190	191	139		1.36	189.1	236.2
CadUni	BA09	141	206	293	256	58	168	81	-431.3	0.56	238	188
2F5N	Familia	$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	S	$\chi 1$	$\chi 2$
Inicial	BB12	146	212	290	248	70	163	156		2.17	270.7	220.7
	BB07	138	264	165	288	137	46	143		0.35	262.9	248
BSC1	BB12	138	203	278	270	62	174	149	-862.1	2.77	256	203.9
	BB07	130	269	160	285	130	48	141		0.12	260.8	242.5
CadUni	BB12	140	200	281	270	68	178	148	-434.9	1.80	248	193
2PI0	Familia	$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	S	$\chi 1$	$\chi 2$
Inicial	BB12	121	197	269	267	62	171	134		2.30	276.8	200.1
	BA05	131	179	268	307	174	39	105		1.11	242.5	224.7
BSC1	BB12	143	203	286	266	64	179	147	-859.7	0.53	242.7	191.8
	BB00	143	185	272	290	183	47	138		1.20	224.3	237.6
CadUni	BB12	140	201	286	268	67	179	146	-435.6	1.60	243	192
3AZH	Familia	$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	S	$\chi 1$	$\chi 2$
Inicial	BB12	150	174	299	226	81	176	147		0.55	276.2	202.2
	BB03	154	154	272	173	148	181	146		2.11	270.7	230
BSC1	BB12	143	203	286	266	64	179	147	-852.9	0.48	242.8	191.8
	NANT	156	210	289	92	186	193	141		0.78	186.4	229.5
CadUni	BB12	140	201	286	268	67	179	146	-435.6	1.60	243	192
2O61	Familia	$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	S	$\chi 1$	$\chi 2$
Inicial	BB12	139	196	290	243.1	63	171	141		2.41	275.9	210.7
	BB02	151	189	244	40.4	184	296	152		1.13	247.9	248.6
BSC1	BB12	142	201	287	266	66.5	180	148	-856	0.34	245.6	188
	NANT	150	195	237	57	166.5	299	160		0.00	205.3	246.7
CadUni	BB12	140	201	286	268	67	179	146	-435	1.60	243	192

$\delta 1$ ,  $\epsilon$ ,  $\zeta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta 2$ . Son los ángulos de torsión del esqueleto azúcar fosfato.  $\chi 1$ ,  $\chi 2$ . Son los ángulos glucosídicos. E es la energía mínima se mide en kilocalorías por mol [KCal/Mol] St. Es el Steacking entre secuencia de bases, se mide en Angstroms cuadrados [ $\text{\AA}^2$ ]

De forma cualitativa en las Figuras 20 y 21 se presentan dos imágenes para los fragmentos de los cristales 1R7M y 2F5N, como ejemplos de estructuras que sufren cambios grandes y de estructuras que casi no cambian después de la optimización. Cabe mencionar que la estructura que sufre los cambios más grandes tiene cadena complementaria en conformación B1 y que los cambios sufridos son a favor de las reglas generales, en las cuales las secuencias Pirimidina-Pirimidina tiene una menor superposición que las secuencias Purina-Purina.



*Figura 20 Comparación entre los cdDMP inicial y optimizado con BSC1 para el cdDMP con conformación BB12 de secuencia dGpdG:dCpdC para el cristal 2F5N, el cual no presenta muchos cambios después de ser optimizado.*



*Figura 21 Comparación entre los cdDMP con conformación BB12 de secuencia dGpdG:dCpdC inicial y optimizado del cristal 1R7M el cual presenta cambios significativos después de ser optimizado.*

### 3.3.2 Fragmentos en conformación BB02

Después de optimizar fragmentos en conformación BB02, con los tres campos de fuerza de AMBER encontramos dificultades para reproducir la estructura de la doble hélice del ADN. En la mayoría de los casos, la estructura tenía grandes cambios en su conformación. La Figura 22 muestra la optimización de un fragmento de cdDMP con los 3 campos de fuerza de AMBER. Se puede observar que los tres campos de fuerza tienen problemas al reproducir la conformación de este fragmento en particular,

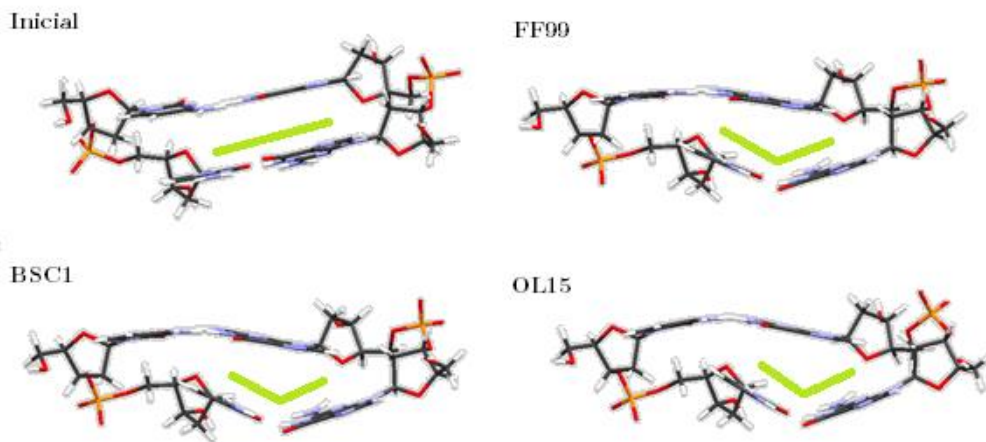


Figura 22 Fragmento cdDMP en conformación BB02 y con secuencia dCpdC:dGpdG del cristal 1A1H.

Otras optimaciones también presentan cambios en la estructura de la hélice, como lo es un “Buckle” grande después de optimizar. También hay mínimos en el cual la base complementaria presenta un “Propeller” grande llegando a los 90 grados, es decir, se trata de bases perpendiculares. En la Figura 23 se presentan ejemplos de estructuras que después de ser optimizadas presentan este tipo de cambios.

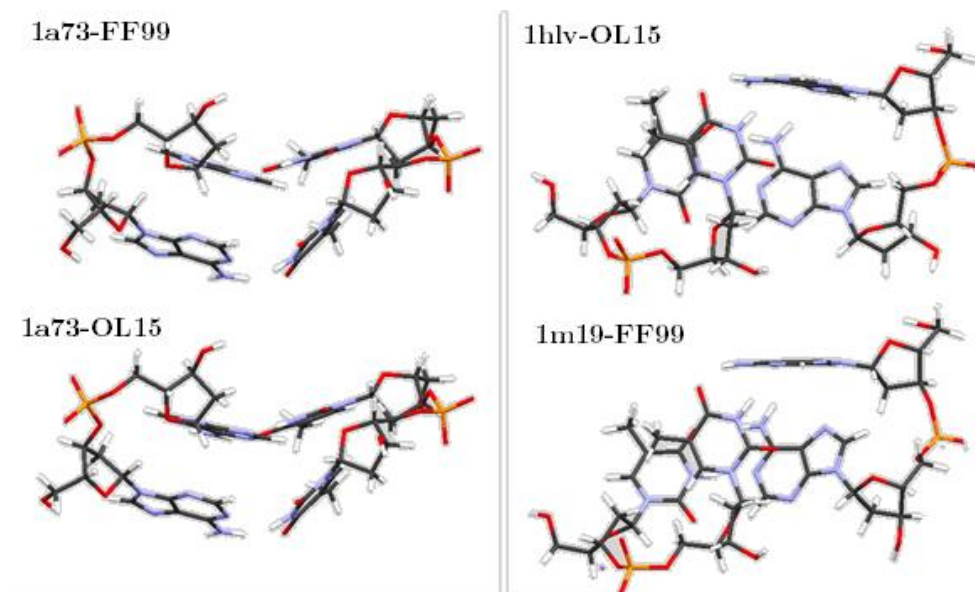
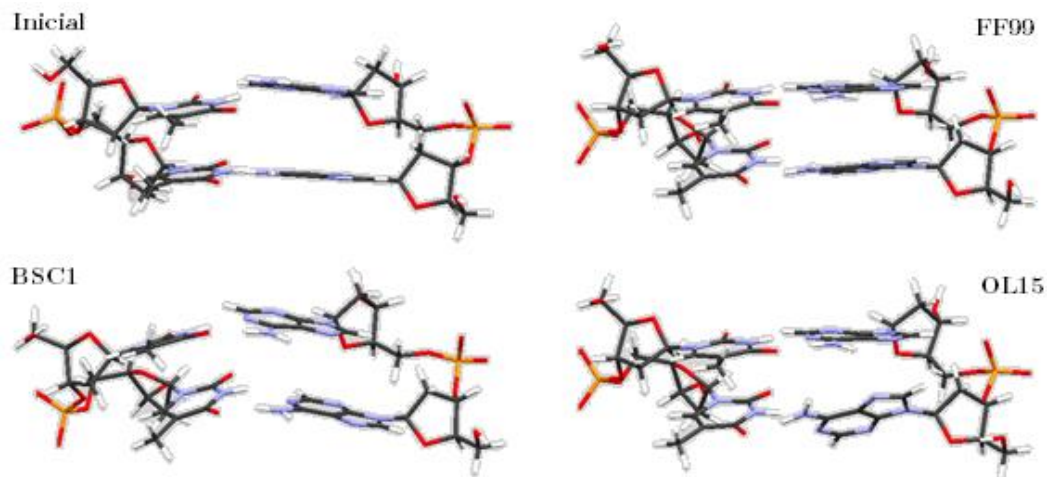


Figura 23 Fragmentos cdDMP optimizados con conformación BB02 Izquierda) Optimizaciones FF99 y OL15 para el fragmento 1A73 con secuencia dTpdT:dApdA en el cual se presenta un buckle grande. Derecha) Optimizaciones con OL15 y FF99 para los fragmentos 1HLV y 1M19 respectivamente, en los cuales se presenta una base perpendicular al resto de bases.

Sin embargo, no todos los fragmentos tienen este inconveniente, en particular la secuencia 307D-E-A47 puede reproducir de forma general la conformación de la hélice doble. Esta estructura, y casi todas, presentan una interacción entre el hidrógeno HO5' y el oxígeno O4' del anillo del azúcar, lo cual no corresponde a estructuras experimentales pues en la posición de HO5' se encuentra un grupo fosfato y esta interacción es imposible. La Figura 24 muestra la estructura inicial 307D-E-A47 y las estructuras que se obtuvieron tras optimizar con los tres campos de fuerza AMBER.



*Figura 24 Fragmentos cdDMP optimizados para el fragmento 307D-E-A47 con los tres campos de fuerza, todos como hélice doble.*

La mayoría de las optimizaciones presentaron fallas como las mostradas en el primer ejemplo, aun así, hubo resultados diferentes e interesantes. A continuación, se presenta la Tabla 8 con algunos resultados representativos de la familia BB02.

Tabla 8 Valores de los ángulos de torsión del esqueleto azúcar fosfato y la superposición de varios fragmentos representativos de la conformación BB02

		$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	St1	St2
<b>[BB02]</b>		<b>140</b>	<b>194</b>	<b>246</b>	<b>32</b>	<b>194</b>	<b>297</b>	<b>150</b>		
<b>dCpdC</b>	ini	142	203	265	18	211	293	146	1.02	0.82
<b>BB02</b>	BSC1	140	191	237	58	170	296	158	0.0	0.08
<b>1A1K-</b>	OL15	136	189	233	61	173	297	157	0.0	0.2
<b>C-C56</b>	FF99	134	189	236	61	172	292	156	0.0	0.12
<b>dGpd</b>	ini	153	176	253	44	178	301	148	3.57	0.0
<b>A</b>	BSC1	147	194	236	53	169	293	160	0.35	0.0
<b>BB02</b>	OL15	147	191	228	60	146	307	160	1.25	0.24
<b>1A1K-</b>	FF99	142	193	237	59	172	286	157	0.16	0.0
<b>B-G8</b>										
<b>dCpdC</b>	ini	142	201	265	21	203	295	145	1.11	0.77
<b>BB02</b>	BSC1	140	191	237	58	170	296	158	0.0	0.08
<b>1A1H-</b>	OL15	136	189	233	61	173	297	157	0.0	0.02
<b>C-C56</b>	FF99	134	190	236	61	172	292	156	0.0	0.12
<b>dTpdC</b>	ini	135	203	253	24	205	291	146	0.67	1.36
<b>BB02</b>	BSC1	140	189	240	57	171	296	158	0.0	0.05
<b>1DOU-</b>	OL15	136	187	237	60	173	297	156	0.0	0.08
<b>B-T20</b>	FF99	135	188	239	60	172	293	155	0.0	0.07
<b>dApdA</b>	ini	149	193	256	31	211	287	138	3.13	0.0
<b>BB02</b>	BSC1	133	188	239	58	176	297	157	0.0	0.0
<b>1G9Z-</b>	OL15	129	185	236	59	182	297	154	0.0	0.0
<b>C-</b>	FF99	128	186	238	61	177	295	155	0.0	0.0
<b>A503</b>										
<b>dApdA</b>	ini	120	177	250	55	163	305	141	2.3	0.0
<b>BB02</b>	BSC1	138	187	238	52	171	296	159	1.96	0.0
<b>307D-</b>	OL15	136	196	197	49	139	319	153	0.39	0.16
<b>E-A47</b>	FF99	136	191	208	37	152	332	159	0.51	0.33
<b>dTpdT</b>	ini	136	201	248	25	216	293	153	0.0	3.13
<b>BB02</b>	BSC1	141	188	241	60	174	300	160	0.0	0.0
<b>1G9Z-</b>	OL15	137	186	238	63	177	300	158	0.0	0.0
<b>C-</b>	FF99	136	187	240	63	175	299	158	0.0	0.0
<b>A503</b>										

$\delta 1$ ,  $\epsilon$ ,  $\zeta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta 2$ . Son los ángulos de torsión del esqueleto azúcar fosfato.

St1, St2. Es el Stacking entre secuencia de bases, se mide en Angstroms cuadrados [ $\text{\AA}^2$ ].

### 3.3.3 Fragmentos en conformación BA09 y BA13

Otra conformación estudiada fueron las familias BA09. Esta conformación tiene ángulos de torsión en las mismas regiones que la familia BB12 a excepción del ángulo  $\delta 2$ , la familia BA09 tiene el ángulo  $\delta 2$  en una región cercana a  $90^\circ$  como en la familia A.

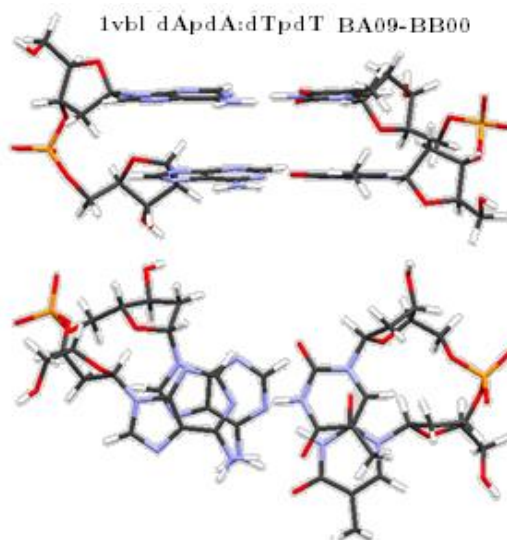
Tras la búsqueda de esta conformación, sólo se encontraron tres dúplex de las secuencias deseadas con una cadena en conformación BA09.

Después de optimizar esas estructuras con los tres campos de fuerza de AMBER encontramos que las estructuras con familia BA09 dan resultados diferentes dependiendo con que campo de fuerzas se optimice. Para los campos de fuerzas FF99 y OL15 cambian la región del ángulo  $\delta 2$  a la misma región que la familia BB12, pero eso no sucede para el campo de fuerza BSC1. En la Tabla 9 se presentan los resultados de los tres fragmentos con sus respectivas optimizaciones. La Figura 25 muestra el fragmento cdDMP 1VBL como ejemplo de la conformación BA09.

Tabla 9 Valores de los ángulos de torsión del esqueleto azúcar fosfato y la superposición de fragmentos con la conformación BA09.

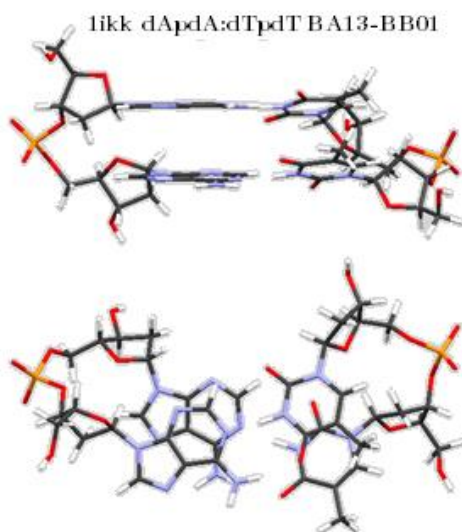
		$\delta 1$	$\epsilon$	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	$\chi 1$	$\chi 2$	St1	St2
<b>[BA09]</b>		<b>135</b>	<b>199</b>	<b>287</b>	<b>254</b>	<b>71</b>	<b>168</b>	<b>86</b>	<b>265</b>	<b>187</b>		
<b>dGpd</b>	ini	136	201	277	259	69	169	82	272	186	3.87	0.06
<b>G</b>	BSC1	135	194	282	264	57	175	82	268	175	1.63	0.0
<b>BA09</b>	OL15	129	194	281	269	63	175	82	273	184	2.8	0.0
<b>1LMB-</b>	FF99	131	193	282	272	62	183	82	281	185	2.18	0.0
<b>C-G13</b>												
<b>dApd</b>	ini	139	203	306	251	68	168	79	266	192	3.29	0.0
<b>G</b>	BSC1	133	198	284	264	56	164	89	268	189	4.9	0.0
<b>BA09</b>	OL15	123	187	277	267	59	162	100	270	188	4.97	0.0
<b>1WTE</b>	FF99	123	191	281	268	55	167	96	268	188	4.8	0.0
<b>-B-A4</b>												
<b>dApdA</b>	ini	129	201	283	253	62	169	87	263	193	3.93	0.0
<b>BA09</b>	BSC1	139	203	288	261	55	169	82	255	184	3.8	0.04
<b>2VBL-</b>	OL15	122	179	272	268	59	165	104	269	185	2.92	0.58
<b>C-A16</b>	FF99	132	198	286	267	56	177	83	264	184	4.02	0.19

$\delta 1$ ,  $\epsilon$ ,  $\zeta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta 2$ . Son los ángulos de torsión del esqueleto azúcar fosfato. **St1**, **St2**. Es el Stacking entre secuencia de bases, se mide en Angstroms cuadrados [ $\text{\AA}^2$ ].  **$\chi 1$** ,  **$\chi 2$** . Son los ángulos glucosídicos



*Figura 25 Fragmento cdDMP 1VBL en conformación BA09*

También se optimizaron estructuras en conformación BA13, estos resultados conservaron las características estructurales de la hélice doble con los tres campos de fuerzas AMBER, pero tenían cambios significativos en las regiones de los ángulos de torsión comparados con su estructura inicial, y por tanto cambiaban de conformación después de la optimización. La Figura 26 muestra el fragmento inicial de cdDMP 1IKK con conformación BA13. A continuación, se presenta la Tabla 10 con algunos resultados representativos de la familia BA13.



*Figura 26 Fragmento cdDMP 1IKK en conformación BA13.*

Tabla 10 Valores de los ángulos de torsión, energía y superposición de las bases para la estructura inicial y las optimizaciones de fragmentos con conformación BA13

1IKK-B-A16			$\delta 1$	E	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	St
Inicial	BA13	dApdA	141	225.1	189.6	75.8	237.2	194.8	87.5		1.57
	BB01	dTpdT	129.7	184.2	263.7	304.6	168.3	52.8	108.4		0.28
OL15		dApdA	125.9	273.4	160	59.1	232.1	174.3	88.1	-508	0.25
		dTpdT	136.3	182.2	274.5	296.8	161.8	59.1	119		1.37
BSC1		dApdA	130.6	270.6	162	63.7	228.1	172.5	88.1	-510.7	0.16
		dTpdT	140.3	189	277.2	292.6	166.4	55.5	130.5		1.26
FF99		dApdA	131.3	198.6	175.5	77.6	266.5	184.5	93	-517.4	1.47
		dTpdT	134.8	185.8	275.3	295.6	164.6	56.7	117.6		0.58
ISK5-A-A6			$\delta 1$	E	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	St
Inicial	BA13	dApdA	142.5	220	195.4	79.4	239.1	198.7	80.7		1.89
		dTpdT	118.9	180.1	269.9	298.5	170.7	56.8	105.3		0.21
OL15		dApdA	130.4	208	173.8	79.5	255.4	180.3	93.7	-506.8	0.79
		dTpdT	135.3	181.8	273.7	296.9	162.1	59.2	120		1.05
BSC1		dApdA	136.2	205.4	169.6	80.9	266.6	177.7	88	-510.5	1.07
		dTpdT	139.2	188.7	275.5	292.5	167	55.9	131.2		0.72
FF99		dApdA	131.3	198.6	175.5	77.6	266.4	184.5	93.6	-517.5	1.47
		dTpdT	134.8	185.9	275.3	295.6	164.6	56.7	117.5		0.58
1D49			$\delta 1$	E	$\zeta$	$\alpha$	$\beta$	$\gamma$	$\delta 2$	E	St
Inicial	BA13	dApdA	143.3	239.8	195.6	72.7	233.4	195.5	90		1.62
		dTpdT	115.6	191.8	253.4	319.5	171.9	40.5	115.6		0.07
OL15		dApdA	125.9	273.4	160	59.2	232.1	174.2	88.1	-508.1	0.25
		dTpdT	136.3	182.2	274.5	296.7	161.8	59.2	119		1.37
BSC1		dApdA	130.6	270.6	162	63.6	228.1	172.5	88.1	-510.7	0.16
		dTpdT	140.2	189.1	277.2	292.5	166.4	55.5	130.6		1.26
FF99		dApdA	124.2	296.8	91.9	72.5	207.3	184.3	89.6	-519.7	0.99
		dTpdT	135.2	185.8	275.7	295.7	164.2	56.7	166.5		0.88

$\delta 1$ ,  $\epsilon$ ,  $\zeta$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta 2$ . Son los ángulos de torsión del esqueleto azúcar fosfato.  $\chi 1$ ,  $\chi 2$ . Son los ángulos glucosídicos. **E**. Es la energía mínima se mide en kilocalorías por mol [KCal/Mol]. **St**. Es el Steacking entre secuencia de bases, se mide en Angstroms cuadrados [ $\text{Å}^2$ ].

Cabe mencionar que, en todos los cálculos obtenidos, sin importar el campo de fuerza, se generaba una interacción tipo enlace de hidrógeno entre el hidrógeno HO5' y el oxígeno O4' del azúcar fosfato, si bien la contribución energética era muy pequeña este resultado no corresponde a la estructura del ADN por lo que se buscó la forma de restringir la posición del hidrógeno HO5' con el mismo software AMBER.

## Capítulo 4. Conclusiones

1. Fueron optimizados fragmentos mínimos de estructuras seleccionadas pertenecientes a 4 clases conformacionales con pares de nucleósidos de Watson y Crick de cadena única, hélice doble y esqueleto azúcar fosfato utilizando 3 campos de fuerzas AMBER.
2. Entre las diferentes familias conformacionales del ADN hay una familia, denominada como BB12, que tiene características generales que permiten la formación de una estructura espacial dependiente de la secuencia de nucleósidos, como sucede con las familias canónicas BI y AI.
3. Los fragmentos mínimos de cadena única de conformación BB12 tienen regiones de ángulos de torsión de azúcar-fosfato que corresponden al mínimo local del fragmento SPB separado, que fue obtenido con los tres campos de fuerzas mencionados.
4. La optimización de la geometría de los cdDMPs, dDMPs y SPB, partiendo de estructuras experimentales de la familia BB12, con diferentes secuencias de nucleósidos Purina-Purina:Pirimidina-Pirimidina muestra que el campo de fuerzas BSC1 reproduce mejor las estructuras del banco de datos que los otros dos campos de AMBER.
5. Los fragmentos mínimos de cadena única en conformación BB02 tienen ángulos de torsión en regiones que no corresponden a mínimo local de energía del SPB optimizado de forma separada. Es por esto que los dDMP y cdDMP de esta conformación dan como resultado de la optimización, estructuras que no corresponden a la geometría de hélice doble.

## Bibliografía.

- [1] J. D. Watson and F. H. Crick. Molecular structure of nucleic acids. *Nature*, 171, No. 4356, pp. 737-738, 1953..
- [2] Krebs, J.E.; Goldstein, E. S.; and Kilpatrick, S.T. (2014), *Lewin's Genes XI* (11a ed.). Jones Barlett Edition.
- [3] Blackburn, G.M.; Gait, M.J.; Loakes, D.; and Williams, D.M. (2006). *Nucleic Acids in Chemistry and Biology* (3rd ed.). UK, RSC Publishing.
- [4] Poltev, V. I.; Anisimov, V.M.; Danilov, V. I.; Garcia, D.; Deriabina, A.; González, E.; Salazar, R.; Rivas, F.; and Polteva, N. (2011). *DFT study of DNA sequence dependence at the level of dinucleoside monophosphates*. *Computational and Theoretical Chemistry*, **975**(1-3), 69-75.
- [5] Poltev, V. I; Anisimov, V. M.; Danilov, V. I.; Garcia, D.; Sanchez, C.; Deriabina, A.; González, E.; Rivas, F.; and Polteva, N. (2014). *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**(6), 640-650.
- [6] Poltev, V. I; Anisimov, V. M.; Sanchez, C.; Deriabina, A.; González, E.; Garcia, D.; Rivas, F.; and Polteva, N. (2016). *Analysis of the conformational features of Watson-Crick duplex fragments by molecular mechanics and quantum mechanics methods*. *Biophysics BIOPHYSICS*, **61**(2), 217-226.
- [7] J. V. Poltev (2012). *Molecular Mechanics: Method and Applications*. In Leszczynski, (Ed.), *Handbook of computational chemistry* (pp. 259-291). Dordrecht: Springer.
- [8] Poltev, V., Anisimov, V. M., Dominguez, V., Deriabina, A., Gonzalez, E., Garcia, D., & Rivas, F. (2018, September). Current Problems in Computer Simulation of Variability of Three-Dimensional Structure of DNA. In *International Workshop on Quantum Systems in Chemistry, Physics, and Biology* (pp. 233-253). Springer, Cham.
- [9] Čech, P., Kukul, J., Černý, J., Schneider, B., & Svozil, D. (2013). Automatic workflow for the classification of local DNA conformations. *BMC bioinformatics*, *14*(1), 205
- [10] Schneider, B., Božíková, P., Nečasová, I., Čech, P., Svozil, D., y Černý, J. (2018). Un alfabeto estructural de ADN proporciona una nueva visión de la flexibilidad del ADN. *Acta Crystallographica Sección D: Biología Estructural*, *74* (1), 52-64.
- [11] J. Cerney, P. Bozikov a, J. Svodoba & B. Schneider. (2020). A unified dinucleotide alphabet describing both RNA and DNA structures. *Nucleic Acid Research* *48* (11), 6367.
- [12] Cornell W. D., Cieplak P., Bayly C. I., Gould I. R. Merz K. M. Jr., Ferguson D. M., Spellmeyer D. C., Fox T., Caldwell J. W. and Kollman P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids and Organic Molecules, *Journal of the American Chemical Society*, Vol. 117, No. 19 (1995), pp. 5179-5197.

- [13] D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O’Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman (2021), Amber 2021, University of California, San Francisco.
- [14] Galindo-Murillo, R., Robertson, J. C., Zgarbova, M., Sponer, J., Otyepka, M., Jurečka, P., & Cheatham III, T. E. (2016). Assessing the current state of amber force field modifications for DNA. *Journal of chemical theory and computation*, 12(8), 4114-4127.
- [15] Hestenes, M. R.; and Stiefel, E. (1952). *Methods of Conjugate Gradients for Solving Linear Systems*. Journal of Research of the National Bureau of Standards, 9(6), 409-436.
- [16] Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I. N.; and Bourne, P.E. (2000). *The Protein Data Bank*. Nucleic Acids Research, 28, 235-242. Recuperado de <http://www.rcsb.org/pdb/home/home.do>.
- [17] Lu, X. J.; and Olson, W.K. (2003). *3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures*. NucleicAcidsResearch. 31(17), 5108-21. Recuperado de <http://x3dna.org/citations/>.
- [18] Černý, J., Božíková, P., Malý, M., Tykač, M., Biedermannová, L., & Schneider, B. (2020). Structural alphabets for conformational analysis of nucleic acids available at dnatco.datmos.org. *Acta Crystallographica Section D: Structural Biology*, 76(9). Recuperado de: <https://dnatco.datmos.org/>