



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Estudio de diferentes conformaciones para fragmentos
simples del ADN, desoxirribosa y nucleósidos, por métodos
de mecánica molecular

Tesis presentada al

Colegio de Física

como requisito parcial para la obtención del grado de

LICENCIADO EN FÍSICA

por

José Antonio Piceno Martínez

Asesores de Tesis:

Dr. Valeri Poltev

y

Dra. Alexandra Deriabina

Puebla Pue.
Noviembre 2021



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Estudio de diferentes conformaciones para fragmentos
simples del ADN, desoxirribosa y nucleósidos, por métodos
de mecánica molecular

Tesis presentada al

Colegio de Física

para obtener el grado de:

LICENCIADO EN FÍSICA

Presenta:

José Antonio Piceno Martínez

Asesores de Tesis:

Dr. Valeri Poltev

y

Dra. Alexandra Deriabina

Puebla Pue.
Noviembre 2021

Título: Estudio de diferentes conformaciones para fragmentos simples del ADN, desoxirribosa y nucleósidos, por métodos de mecánica molecular

Estudiante: JOSÉ ANTONIO PICENO MARTÍNEZ

COMITÉ

Dr. Eduardo González Jiménez
Presidente

Dra. Dolores García Toral
Secretario

Dr. Carlos Robledo Sánchez
Vocal

Dr. Roberto Ramírez Sánchez
Suplente

Dr. Valeri Poltev
Asesor

Dra. Alexandra Deriabina
Asesor

Agradecimientos

Agradezco enormemente a mis padres Male y Toño, por su apoyo incondicional, por su constante compromiso con mi bienestar y por toda su paciencia y atención.

A mis hermanos Neni y Neto, por siempre estar dispuestos a ayudarme.

Agradezco también a mi pareja Paola por hacer mis días felices y darme la motivación necesaria para seguir.

A mi asesor Poltev por su paciencia y dedicación en este trabajo, y al equipo de trabajo que conocí en esta área.

A mi primo Juan Carlos y a mis amigos de toda la vida, Dani y Mike, sé que estarán orgullosos.

A todos ustedes les dedico este trabajo.

Resumen

El estudio de la estructura del ácido desoxirribonucleico (ADN), de los nucleótidos, nucleósidos y de los ácidos nucleicos (AN) tiene una gran importancia biológica. Los nucleótidos tienen muchas funciones en organismos vivos y para entender su función biológica debemos conocer sus características estructurales. A partir de la estructura de los nucleótidos llegamos a la hélice doble del ADN, la cual forma parte de un principio estructural que cambió el mundo de la bioquímica, y de la doble hélice pasamos a los cromosomas, los cuales contienen la mayor parte de la información genética de un ser vivo.

En el presente trabajo se hará un análisis estructural de subunidades de hélice doble del ADN, en particular de nucleósidos y desoxirribosa. Se usarán diferentes campos de fuerzas implementados en el software AMBER (Assisted Model Building with Energy Refinement) para realizar la optimización geométrica de estas estructuras, destacando los campos de fuerzas FF99, BSC1 y OL15, utilizando el método de Mecánica Molecular. Con base en los resultados, se determinará la viabilidad de cada campo de fuerzas usado para reproducir las características estructurales de los fragmentos analizados.

Índice general

Agradecimientos	V
Resumen	VII
Índice general	IX
Índice de figuras	XI
Índice de tablas	XIII
Objetivos	1
Introducción	3
1. Estructura del ADN	9
1.1. Bases Nitrogenadas	9
Bases del ADN	10
Reglas de Chargaff:	12
1.2. Azúcares	12
1.3. Nucleósidos y Nucleótidos	14
1.4. Niveles de Organización de Estructuras del ADN	16
1.5. Familias de ADN	18
1.6. Parámetros Estructurales	19
2. Optimización Geométrica del ADN	21
2.1. Mecánica Molecular	21
2.2. Campo de Fuerza	23
2.3. Optimización Geométrica con AMBER	26
3.3. Optimización Geométrica con AMBER	26
2.3.1. Procedimiento de Optimización Usando el Software AMBER	26
2.4. Mediciones y Cálculos Finales	29
2.4.1. Mediciones con Mercury	29
2.4.2. Cálculos de Parámetros de Pseudorrotación y Distancias C2', C3'	30
3. Resultados	31
3.1. Mínimos de energía de desoxicitidina	32
3.2. Mínimos de energía de desoxitimidina	36
3.3. Mínimos de energía de desoxiadenosina	40
3.4. Mínimos de energía de desoxiguanosina	43
3.5. Mínimos de energía de desoxirribosa	46
Conclusión	50

A. Preparación de Estructuras	53
A.1. Búsqueda y Selección	53
A.2. Recorte del Fragmento	57
B. Cambio de Base de un Nucleósido	59
C. Restricción de Ángulos de Torsión	61
C.0.1. Cambiar Ángulos de Torsión Antes de Optimizar	62
C.0.2. Crear Archivo RTS	63
Bibliografía	65

Índice de figuras

1.	Imagen 51.	5
1.1.	Estructura de purina y pirimidina	9
1.2.	Bases nitrogenadas del ADN	10
1.3.	Emparejamiento de Watson y Crick	11
1.4.	Estructura de 2'-desoxirribosa	12
1.5.	Empaquetamiento del azúcar	13
1.6.	Círculo pseudorrotacional	13
1.7.	Estructura de nucleósidos y nucleótidos	15
1.8.	Cadena del ADN	16
1.9.	Hélice doble del ADN	17
1.10.	Niveles de organización del ADN	18
1.11.	Familias de conformaciones de ADN	19
1.12.	Ángulos de torsión para dDMP	20
1.13.	Parámetros sobre la organización de las bases	20
2.1.	Cargas para ADN usadas en el paquete AMBER	24
2.2.	Potenciales usados en MM	25
2.3.	Unit Editor de XLeap, AMBER	27
2.4.	Interfaz del software Mercury	29
3.1.	Ángulos de torsión para desoxinucleósidos y desoxirribosa.	31
3.2.	Mínimos locales de desoxicitidina, BSC1.	33
3.3.	Mínimos globales de desoxicitidina	35
3.4.	Mínimos locales de desoxitimidina, BSC1.	37
3.5.	Mínimos globales de desoxitimidina	39
3.6.	Mínimos locales de desoxiadenosina, BSC1.	41
3.7.	Mínimos globales de desoxiadenosina	42
3.8.	Mínimos locales de desoxiguanosina, BSC1.	44
3.9.	Mínimos globales de desoxiguanosina	45
3.10.	Mínimos locales de desoxirribosa, BSC1.	47
3.11.	Mínimos globales de desoxirribosa	48
A.1.	Sitio web DNATCO	53
A.2.	Estructura en DNATCO.	54
A.3.	Fragmento en DNATCO.	54
A.4.	Sitio web NDB.	55
A.5.	Sitio web RCSB PDB.	56
A.6.	Recorte en el Archivo.pdb con el procesador de texto.	57
B.1.	Archivo .pdb de un desoxinucleósido	60

C.1. Potencial del ángulo de torsión con restricción	62
C.2. Software Chimera	63

Índice de tablas

3.1. Mínimos locales encontrados con el campo BSC1 para desoxicitidina.	32
3.2. Mínimos locales encontrados con el campo FF99 para desoxicitidina.	34
3.3. Mínimos locales encontrados con el campo OL15 para desoxicitidina.	34
3.4. Mínimos locales encontrados con el campo BSC1 para desoxitimidina.	36
3.5. Mínimos locales encontrados con el campo FF99 para desoxitimidina.	38
3.6. Mínimos locales encontrados con el campo OL15 para desoxitimidina.	38
3.7. Mínimos locales encontrados con el campo BSC1 para desoxiadenosina.	40
3.8. Mínimos locales encontrados con el campo FF99 para desoxiadenosina.	41
3.9. Mínimos locales encontrados con el campo OL15 para desoxiadenosina.	42
3.10. Mínimos locales encontrados con el campo BSC1 para desoxiguanosina.	43
3.11. Mínimos locales encontrados con el campo FF99 para desoxiguanosina.	44
3.12. Mínimos locales encontrados con el campo OL15 para desoxiguanosina.	45
3.13. Mínimos locales encontrados con el campo BSC1 para desoxirribosa.	46
3.14. Mínimos locales encontrados con el campo FF99 para desoxirribosa.	46
3.15. Mínimos locales encontrados con el campo OL15 para desoxirribosa.	47

Objetivos

Objetivo General

Encontrar los mínimos locales de energía para subunidades simples del ADN, de desoxirribosa y desoxinucleósidos, utilizando tres campos de fuerza AMBER y evaluar la eficacia de cada campo para la simulación de dichos fragmentos.

Objetivos Particulares

- Buscar y seleccionar estructuras de cristales de desoxinucleósidos, desoxinucleótidos y fragmentos de ADN, los cuales tienen diferentes parámetros conformacionales de subunidades del ADN.
- Preparar los archivos con las características geométricas de desoxirribosa y nucleósidos, cortados de las estructuras seleccionadas.
- Hacer la optimización geométrica utilizando tres campos de fuerza AMBER de las estructuras de desoxirribosa y nucleósidos previamente preparadas, y medir los parámetros conformacionales que corresponden a mínimos locales.
- Hacer el análisis comparativo de estructuras optimizadas con los tres campos y de estructuras iniciales, y sacar conclusiones sobre la posibilidad que tiene cada campo de fuerza para reproducir las estructuras de subunidades.

Introducción

La doble hélice del ácido desoxirribonucleico (ADN) es probablemente la más emblemática de todas las moléculas biológicas hoy en día, en 1953 se descubrió la estructura de doble hélice del ADN, con el trabajo de los científicos James Watson y Francis Crick [1], y gracias a previas observaciones de Rosalind Franklin y otros investigadores. Con este descubrimiento se empezaron a realizar varias investigaciones en biofísica molecular que han contribuido en muchas áreas del conocimiento como la genética, medicina, biotecnología, etc. Sin embargo, quedan sin resolver una serie de cuestiones relacionadas con el esclarecimiento de los principios físico-químicos de la formación de la estructura espacial de diferentes secuencias de los nucleótidos que forman las cadenas del ADN y del papel que tienen las interacciones entre los componentes de los biopolímeros. Por esta razón se continúa con el estudio intensivo de los ácidos nucleicos (AN) por diferentes métodos.

Es importante conocer los principios físicos de la construcción de los AN, la dependencia de la secuencia y las preferencias conformacionales de la estructura del ADN en tres dimensiones, ya que esto nos dará una mejor comprensión de los mecanismos moleculares de los procesos del ADN. Por lo tanto, en este trabajo se estudiarán las subunidades de fragmentos de ADN, desoxirribosa y nucleósidos, se analizarán sus dependencias y diferencias, así como sus características comunes y para ello utilizaremos diferentes campos de fuerza [2] implementados en el paquete computacional AMBER [3].

Existen trabajos anteriores de los asesores de este trabajo [4, 5] en los cuales se usan métodos de mecánica cuántica de alto costo computacional. En estos trabajos se observó que las características principales de la estructura de hélice doble, tales como los ángulos de torsión del esqueleto azúcar-fosfato, el empaquetamiento del anillo de azúcar y el orden de las bases paralelas, están predeterminados en el mínimo local de energía de los dDMPs (deoxiDinucleosid Mono Phosphate por sus siglas en inglés).

Los primeros estudios de la hélice doble se realizaron en la década de 1950 con modelos hechos a mano, después de la invención de la computadora fue posible el uso de otras herramientas para estudiar la hélice doble, tales como modelos mecánico-cuánticos y modelos mecánico-moleculares. Un ejemplo de estas herramientas computacionales es la parametrización de campos de fuerza en los programas como AMBER (Assisted Model Building with Energy Refinement) o CHARMM (Chemistry at HARvard Macromolecular Mechanics), que son ajustados a resultados experimentales y consideraciones de mecánica cuántica y son utilizados para calcular la energía potencial del sistema de átomos.

En años anteriores se ha encontrado por métodos computacionales que el fragmento mínimo de hélice doble (dDMP) puede existir en conformaciones diferentes a las canónicas de Watson-Crick [6], es nuestra tarea encontrar la contribución de las subunidades (nucleótidos y nucleósidos) a estas conformaciones. Sin embargo no es posible obtener resultados experimentales con buenos grados de exactitud, debido a que las estructuras son muy complejas.

Existen muchas conformaciones diferentes de fragmentos simples para estudiar, para ello es necesario utilizar el método semiempírico de la Mecánica Molecular (MM), en dicho método hay diferentes campos de fuerza que podemos utilizar para analizar las propiedades energéticas y geométricas de las diferentes estructuras. Por lo tanto es necesario aprobar la reproducción de dichos campos usando subunidades y fragmentos simples del ADN.

Un Poco de Historia

El ADN fue descubierto en 1869, por el científico suizo Johann Friederich Miescher [7]. Cuando él examinaba pus de vendajes quirúrgicos con el microscopio, encontró una sustancia no identificada a la que llamó nucleína (ya que residía en el núcleo de la célula).

A principios de la década de 1950, el biólogo estadounidense James Watson y el físico británico Francis Crick propusieron su famoso modelo de la doble hélice del ADN [6]. En ese entonces Linus Pauling, quien descubrió la estructura secundaria de las proteínas [8], y otros científicos también trataban de encontrar el modelo correcto. En lugar de realizar nuevos experimentos en el laboratorio, Watson y Crick principalmente recolectaron y analizaron fragmentos de información y los juntaron de formas novedosas y reveladoras. Algunas de sus pistas más importantes fueron producto del trabajo de Rosalind Franklin [9], una química que trabajaba en el laboratorio del físico Maurice Wilkins, y los descubrimientos del bioquímico austriaco Erwin Chargaff [10].

Franklin era experta en cristalografía de rayos X, técnica usada para la determinación de la estructura de moléculas. Cuando la forma cristalizada de una molécula, como el ADN, se expone a rayos X, los átomos en el cristal desvían algunos de los rayos y forman un patrón de difracción que da pistas sobre la estructura de la molécula.

La cristalografía de Franklin dio a Watson y Crick importantes pistas sobre la estructura del ADN. Algunas de estas provenían de la famosa *imagen 51* [11], una imagen de difracción de rayos X del ADN extraordinariamente clara que produjeron Franklin y su estudiante. El patrón de difracción en forma de "x" de la imagen le sugirió a Watson una estructura helicoidal de dos cadenas, Figura (1).

En 1962, James Watson, Francis Crick y Maurice Wilkins recibieron el premio Nobel de medicina [12]. Desafortunadamente, Franklin había fallecido y los Nobel no se otorgan póstumamente.

Por otro lado, desde la década de 1960 se ha visto una rápida expansión del método de Mecánica Molecular (MM), provocado por la introducción de métodos computacionales en todas las ramas de la ciencia [13].

El método de MM, iniciado en 1940, se basaba en cálculos para conformaciones seleccionadas de moléculas simples hechos manualmente o usando calculadoras de escritorio. Algunos investigadores construyeron modelos hechos a mano de acero o madera para hacer cuidadosas mediciones de los parámetros geométricos de dichas moléculas. Estos cálculos, bastante aproximados, demostraron claramente la utilidad del enfoque de la MM para problemas de química orgánica, así como la necesidad de cálculos más extensos y la búsqueda de parámetros más confiables. Todo esto llegaría unos años después gracias al gran avance computacional.

El primer artículo con un estudio computacional de conformaciones de moléculas orgánicas, en específico para hidrocarburos saturados, fue escrito por Hendrickson, en 1961 [14]. Así, durante los 60's y a comienzo de los 70's se llevaron a cabo muchos cálculos computacionales por diversos grupos de investigadores (destacando Hendrickson et al. [15][16] y Allinger et al. [17]). En 1973 se hizo una extensión del método de MM enfocado a hidrocarburos más complejos [18].

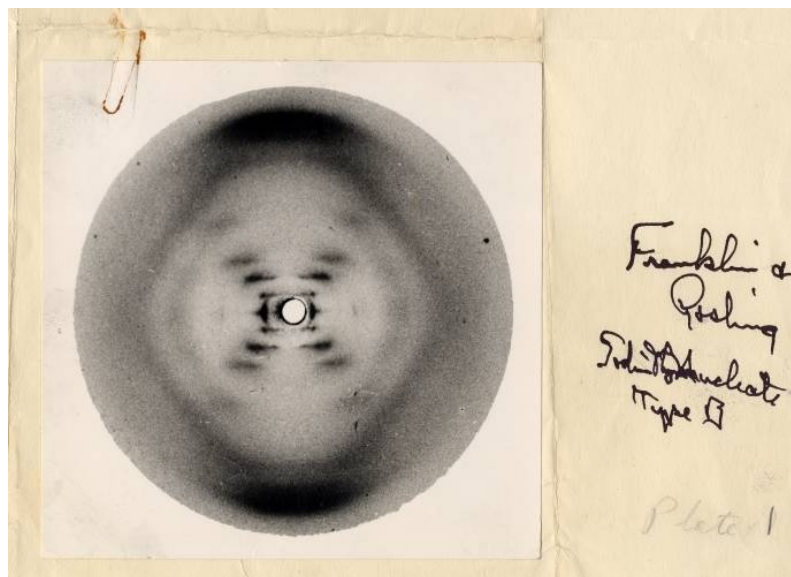


Figura 1: Foto cristalográfica de timonucleato de sodio, tipo B, Mayo 1952 [11].
A partir de esta foto Watson y Crick pudieron rehacer un modelo en tres dimensiones del ADN que habían construido.

El resultado más importante en los inicios de la MM con computación está ilustrado en el artículo titulado *Evaluación Crítica de la Mecánica Molecular* publicado por Engler et al., 1973 [19]. En este artículo se hicieron cálculos para varios hidrocarburos usando dos campos de fuerza, los cuales tienen diferentes parámetros atendiendo a distintos conjuntos de características experimentales. Esto resultó en una diferencia significativa en términos de energía, que puede variar mucho e incluso ser positiva para un campo y negativa para otro. Sin embargo la energía total y los valores relativos de varias conformaciones calculadas con dos campos de fuerza son muy parecidos entre ellos para la mayoría de compuestos considerados. Esto lo podemos observar también en este trabajo.

A mediados de la década de los 60's, el asesor de este trabajo, V. Poltev, propuso el primer esquema átomo-átomo y los parámetros numéricos para calcular interacciones entre bases de ácidos nucleicos [13]. Este esquema tenía tres puntos clave, las estimaciones cuantitativas de subunidades de AN, la aproximación del *monopolo* para interacciones electrostáticas entre bases, y un enfoque átomo-átomo para los cálculos de interacciones de van der Waals en moléculas de cristales. En ese entonces no había un esquema de la MM que encajara con las consideraciones cuantitativas para las interacciones y estructuras de AN [13].

En cuanto a los programas usados para cálculos computacionales, el paquete de software y campos de fuerzas AMBER es el más popular, diseñado para una amplia gama de biomoléculas incluyendo proteínas y AN.

La primera versión del programa AMBER apareció en 1981 [20] y la primera especificación detallada del campo de fuerzas fue publicada 3 años más tarde [21]. Los parámetros del campo de fuerzas habían sido ajustados tanto a los datos experimentales como a los obtenidos por mecánica cuántica. En los próximos años, se publicaron cientos de artículos usando el programa AMBER y su primer campo de fuerza, debido a las diferencias que éste presentaba. A partir de estas publicaciones, se lanzaron nuevas versiones del programa y del campo de fuerzas [13].

En 1995, Cornell et al. [22], publicaron la segunda generación del campo de fuerzas AMBER

(ff94), en la que mejoraron las interacciones energéticas entre proteínas, AN y muchas moléculas orgánicas. De esta versión se siguen otros conjuntos de parámetros con iniciales *ff* seguidos del año en 2 dígitos en el que se lanzaron. Algunas versiones difieren muy poco del campo *ff94*, pero otras, como la versión mejorada *ff99*, reproduce los datos para arreglos moleculares con una precisión muy superior a sus versiones anteriores. El campo de fuerzas *ff99* será muy usado en éste trabajo.

Actualmente, el software AMBER y sus campos de fuerzas son bastante usados para la simulación de varios sistemas biológicos con gran importancia científica. Muchos investigadores siguen buscando formas de mejorar sus parámetros para sistemas más específicos.

Con esto, el método de MM nos ha acercado al estudio de sistemas moleculares desde hace más de 70 años, y su avance computacional lo ha convertido en una útil y poderosa herramienta de investigación para muchas ramas de las ciencias naturales.

Importancia del ADN

Por muchos años, los científicos debatieron qué molécula portaba las instrucciones biológicas para la vida. La mayoría pensaba que el ADN era una molécula demasiado sencilla para desempeñar un papel tan importante, argumentaban que era más probable que las proteínas llevaran a cabo esta función vital debido a su mayor complejidad y más amplia variedad de formas. Su importancia se aclaró en la década de 1950 como vimos en la sección anterior. Hasta ahora seguimos descubriendo y estudiando sus razgos más interesantes.

Los nucleótidos tienen muchas funciones en organismos vivos. Los nucleótidos y nucleósidos interactúan con proteínas en todo el proceso metabólico. Además, el material hereditario, el ácido desoxirribonucleico (ADN), es un polímero construido de unidades monoméricas, los nucleótidos. Incluso si estas unidades no constituyeran el ADN, seguirían estando entre las moléculas más importantes para la biología.

Por su parte, el ADN es la molécula biológica más importante, pues es la portadora de la información genética de todos los seres vivos, incluso de algunos virus. En las secuencias de sus bases, Sección (1.1), se encuentran las instrucciones para las funciones biológicas vitales de los seres vivos. Además cuenta con la habilidad de replicarse a sí misma, haciendo única a esta molécula y garantizando el ciclo de la vida.

El ADN se encuentra en el núcleo de la célula. Como la célula es muy pequeña y hay muchas moléculas de ADN por célula, cada una de estas moléculas de ADN debe estar empaquetada de forma muy compacta y precisa. Esta forma de empaquetamiento del ADN se denomina cromosoma [23].

En el ADN hay segmentos que juegan un papel netamente estructural (ADN sin código), y otros segmentos cuyas secuencias contienen información para organizar y controlar funciones vitales (genes). A nivel funcional se organiza en genes, que son piezas de ADN que generan características fisiológicas específicas; estas características no vienen directamente del propio ADN, sino de una molécula llamada ARN (Ácido Ribonucleico). El ADN funciona como molde para la síntesis de ARN, y el ARN a su vez cumple una función de traductor entre ADN y proteínas. Así, de forma indirecta, el ADN contiene instrucciones para la síntesis de proteínas, pues el ADN se queda en el núcleo y la síntesis de proteínas se lleva a cabo en el ribosoma del citoplasma. Durante el proceso de replicación del ADN, éste se desenrolla para que pueda ser copiado. En otros puntos del ciclo, algunas secciones también se desenrollan para ser usadas en la fabricación de proteínas y para otros procesos biológicos. Pero durante la división celular, el ADN se encuentra en su forma compacta de cromosoma para hacer posible la transferencia a nuevas células.

Es por toda esta información que la estructura del ADN es tan importante y en las investigaciones actuales tiene un lugar invaluable, que dará pie a muchos nuevos descubrimientos importantes para el desarrollo de nuestra especie. Sin embargo, para entender bien la función biológica del ADN y de los nucleósidos, necesitamos conocer sus características estructurales, las cuales veremos en el siguiente capítulo.

Capítulo 1

Estructura del ADN

Como vimos al final del capítulo anterior, para entender el funcionamiento del ADN tenemos que entender su estructura, y para entender su estructura es necesario tener una apreciación completa de sus componentes.

El ADN está conformado por dos hélices que contienen cadenas complementarias de nucleótidos. La unidad de la que se componen las hélices del ADN es el desoxidinucleótido monofosfato complementario (cdDMP)[24]. Cada cadena de ADN es un polímero, cuyos monómeros, conectados entre ellos con enlaces covalentes, consisten en una base nitrogenada, una molécula de azúcar y una molécula de ácido fosfórico (B-S-P), lo que llamamos nucleótido. También puede verse como la suma de un nucleósido (B-S) y un grupo fosfato (P).

Este capítulo se enfocará en explicar la estructura del ADN, desde las características de sus bases, su relación con el azúcar y el empaquetamiento de esta, pasando por las estructuras que forman (nucleósidos y nucleótidos) y llegando hasta su formación como cadenas.

1.1. Bases Nitrogenadas

Para el ADN existen cuatro bases nitrogenadas diferentes: adenina (A), timina (T), guanina (G) y citosina (C). En el caso del ARN se tiene al uracilo (U) en lugar de timina. Las bases C, T y U, que sólo tienen un anillo en su estructura molecular, se llaman **pirimidinas**, mientras que las bases A y G, que tienen dos anillos, se llaman **purinas**, Figura (1.1).

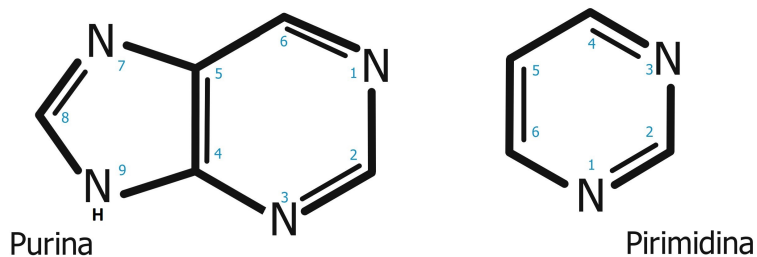


Figura 1.1: Estructura de purina y pirimidina.

Las purinas tienen 2 anillos formados de carbonos y nitrógenos, las pirimidinas sólo tienen 1 anillo.

Podemos ver en la figura la convención para la numeración de las posiciones de cada átomo.

La estructura de las purinas es de dos anillos, uno de seis átomos y otro de cinco, a cada posición de los átomos que forman el anillo se les da un número. En general estos anillos de purinas presentan cuatro átomos de nitrógeno (en las posiciones 1, 3, 7, 9). Las pirimidinas son mucho más grandes que las pirimidinas, ya que estas últimas tienen un sólo anillo de seis átomos. Las purinas cuentan con dos átomos de nitrógeno (1 y 3), Figura (1.1).

Las bases son las que contienen la información genética, estas presentan una estructura cíclica (con átomos de C, N, H, O). Todos los átomos están en un mismo plano espacial (lo que clasifica a estas moléculas como aromáticas). Este aplanamiento es importante en la organización de las bases dentro de la hélice, pues les permite apilarse uniformemente. Las bases se unen a la ribosa con un enlace N-glucosídico en N1 para las pirimidinas y en N9 para las purinas, véase la Figura (1.1). Los planos de la base y el azúcar son casi perpendiculares, pero existen dos conformaciones predominantes: **anti** y **syn**, Sección (1.3).

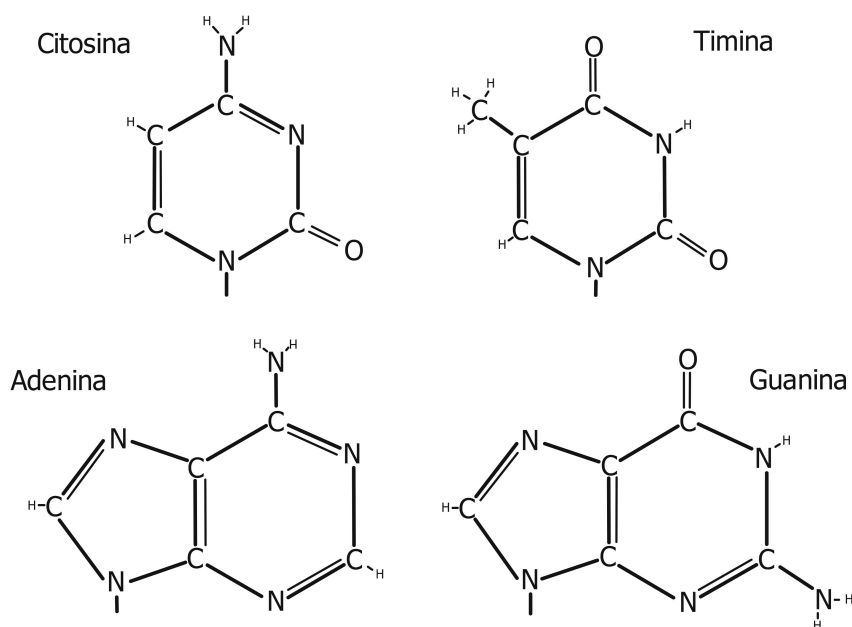


Figura 1.2: Bases nitrogenadas del ADN.
Se describen las bases pirimidinas (citocina y timina) y las purinas (adenina y guanina).

Bases del ADN

- **Adenina:** [6-amino-pur]($C_5H_5N_5$) Representada con la letra **A**. En esta base púrica el hidrógeno 6 es sustituido por un grupo amino (NH_2). Forma los nucleósidos *adenosina* (Ado) y *desoxiadenosina* (dAdo); y los nucleótidos *adenilato* (AMP) y *desoxiadenilato* (dAMP).^{1 2}
- **Guanina:** [2-amino-6-oxo-pur]($C_5H_5N_5O$) Representada con la letra **G**. Cuenta con un oxígeno en lugar del hidrógeno 6 y con un grupo amino en la posición 2. Forma los nucleósidos

¹La adenina y la timina fueron descubiertas en 1885 por el bioquímico alemán Albrecht Kossel [25].

²En 1959 Juan Oro pudo sintetizar adenina a partir del ácido cianhídrico [25].

guanósina (Guo) y *desoxiguanósina* (dGuo); y los nucleótidos *guanilato* (GMP) y *desoxiguanilato* (dGMP).³

- **Citosina:** [2-oxo-4-amino-pyr]($C_4H_5N_3O$) Representada con la letra **C**. Tiene un grupo amino en posición 4 y un grupo cetónico en 2. Forma los nucleósidos *citidina* (Cyd) y *desoxicitidina* (dCyd); y los nucleótidos *citidilato* (CMP) y *desoxicitidilato* (dCMP).⁴
- **Timina:** [2,4-oxo-5-metil-pyr]($C_5H_6N_2O_2$) Representada con la letra **T**. Tiene oxígeno en la posición 2 y 4, y un grupo metil en la posición 5. Forma los nucleósidos *timidina* (Thd) y *desoxitimidina* (dThd); y los nucleótidos *timidilato* (TMP) y *desoxitimidilato* (dTMP). En el ARN el Uracilo toma el lugar de la Timina, Figura (1.2).

En el ADN, cada nucleósido (Base + Azúcar) de una cadena establece una asociación específica con su correspondiente de la otra cadena. Los átomos de hidrógeno de los grupos amino funcionan como donadores de electrones, estableciendo enlaces covalentes, mientras los oxígenos de los grupos carbonilos y los nitrógenos de los anillos tienen el papel de receptores. Las Adeninas sólo se unen con Timinas mediante dos puentes de hidrógeno y las Guaninas sólo se unen con las Citosinas mediante tres puentes de hidrógeno. A estas uniones se les llama *apareamiento de Watson-Crick*, Figura (1.3). Estos pares se conocen a partir de estudios físicos hechos con rayos X [1], y del número de dichos pares depende el cómo será ese ser vivo en específico (tipos de animales o plantas).

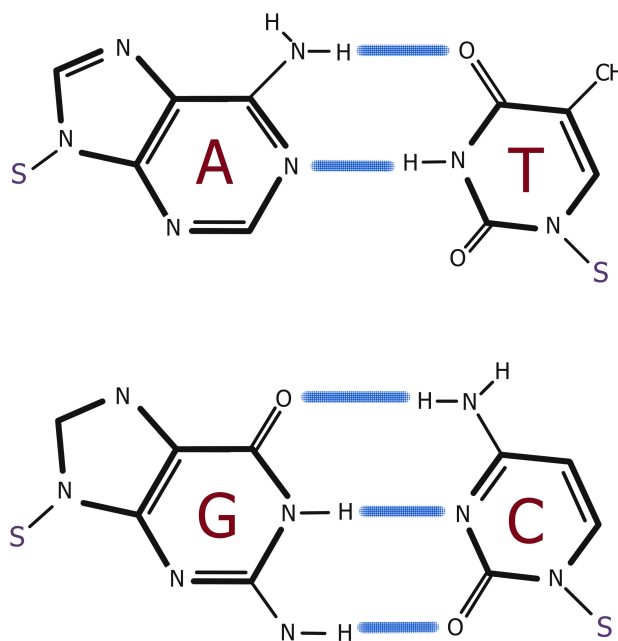


Figura 1.3: Emparejamiento de Watson y Crick.

La adenina se une a la timina mediante dos puentes de hidrógeno; N1-H3 y H6-O4 respectivamente. Mientras, la guanina se une a la citosina con tres puentes de hidrógeno; O6-H4, H1-N3 y H2-O2.

³Fue aislada por primera vez en 1844 a partir de los excrementos de aves marinas (guano), alrededor de 1900 Emil Fischer determinó su estructura y mostró que el ácido úrico se puede convertir en guanina [25].

⁴Fue descubierta por Albrecht Kossel en 1894 cuando fue aislada en tejido de carnero, en 1903 se propuso su estructura y fue sintetizada [25].

Reglas de Chargaff: Otra pieza clave de información relacionada a la estructura del ADN, la proporcionó el bioquímico Erwin Chargaff, quien hizo varias observaciones clave [10], llamadas reglas de Chargaff:

- A, T, C y G no se encontraban en cantidades iguales.
- La cantidad de bases variaba entre especies, pero no entre individuos de la misma especie.
- La cantidad de A era siempre igual a la cantidad de T, y la cantidad de C era siempre igual a la cantidad de G.

1.2. Azúcares

El azúcar en el ADN es una molécula orgánica con forma de pentosa cíclica (furanosa) que contiene un anillo con cuatro carbonos (C1', C2', C3', C4') y un oxígeno (O4'), y un quinto carbono unido al C4' con un grupo hidroxilo (O5'-H)⁵. Este azúcar recibe el nombre de desoxirribosa por prescindir de un grupo hidroxilo en el carbono C2', Figura (1.4).

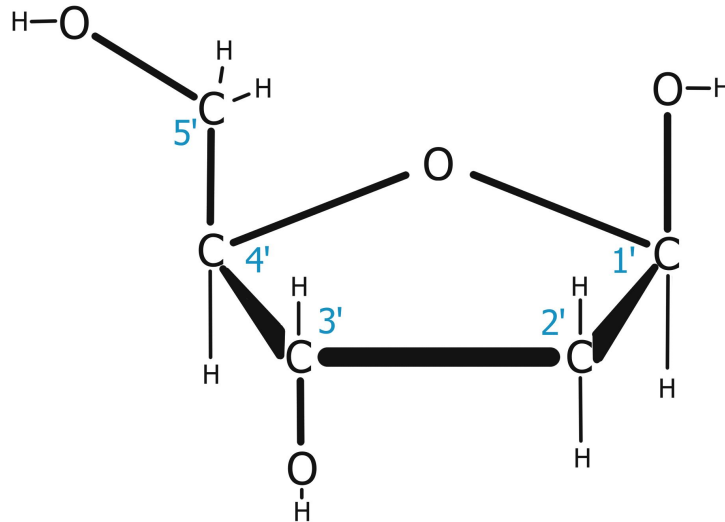


Figura 1.4: Estructura de 2'-desoxirribosa, azúcar del ADN. Los carbonos C5' y C3' son los que están unidos al grupo fosfato mediante enlaces covalentes formando la cadena azúcar-fosfato. El carbono C1' va unido a la base nitrogenada en lugar del grupo OH.

En el anillo de desoxirribosa uno o varios de sus átomos pueden estar fuera del plano que forman otros tres átomos, ya que es una de las partes más flexibles y dinámicas de la molécula de ADN, Figura (1.5).

Otro concepto que tenemos que revisar es la transición entre los tipos de torsión del azúcar, la cual se da siguiendo el denominado *Círculo Pseudorrotacional* [26], Figura (1.6), parametrizado por P. Los 20 tipos de conformaciones pueden ser encontradas en dicho círculo, cada conformación

⁵Las posiciones en el azúcar tienen notación primada para discernir entre los átomos de la base y del azúcar.

Existen dos parámetros que se deben calcular para caracterizar estas formas del azúcar: la fase del ángulo (P) y la amplitud de pseudorrotación (τ_M). Estos parámetros se determinan por cinco ángulos de torsión del anillo (τ_i) y se relacionan según la ecuación:

$$\tau_i = \tau_M \cos[P + 144(i - 1)], i = 0, \dots, 4, \quad (1.1)$$

donde τ_0 es el ángulo de torsión entre $C4'-O4'-C1'-C2'$, τ_1 entre $O4'-C1'-C2'-C3'$, τ_2 entre $C1'-C2'-C3'-C4'$, τ_3 entre $C2'-C3'-C4'-O4'$ y τ_4 entre $C3'-C4'-O4'-C1'$.

El ángulo de pseudorrotación P se calcula usando los ángulos de torsión τ_i (con $i = 0, \dots, 4$) de acuerdo a la siguiente expresión:

$$\tan P = \frac{(\tau_4 + \tau_1) - (\tau_3 + \tau_0)}{2(\tau_2)(\sin 36^\circ + \sin 71^\circ)}. \quad (1.2)$$

A la forma de *empaquetamiento* del azúcar se le llama **puckering**. ¿Por qué sucede esto en el anillo de azúcar? Teóricamente, un anillo de azúcar plano es energéticamente desfavorable y el sistema reduce su energía mediante el puckering; como las barreras de rotación en los enlaces C-O son más bajas que en los enlaces C-C, el anillo adopta conformaciones con ángulos de torsión C-O cercanos a 0° y con ángulos de torsión C-C máximamente escalonados, resultando mayormente en puckerings C2'endo, C2'exo, C3'exo y C3'endo (dependiendo de los otros ángulos de torsión).

Las formas más abundantes de los nucleósidos en la naturaleza son las **N** (C3'endo) y **S** (C2'endo); la forma **E** (O4'endo) se encuentra a veces, pero nunca la **W** (O4'exo), Figura (1.6); las formas **B** del ADN presentan puckerings S y en algún caso E y las formas **A** del ADN y ARN presentan siempre puckerings N, Sección (1.5).

La distribución de los ángulos internos (exocíclicos) en el anillo del azúcar es similar a la presentada en otras moléculas orgánicas, como el tetrahidrofurano, con el cuál se hicieron muchas investigaciones para describir la geometría de sus enlaces [27]. La suma de sus ángulos exocíclicos es alrededor de 523° ; igual a la suma de los ángulos exocíclicos del tetrahidrofurano, esta similitud sugiere que los efectos de empaquetamiento influyen más que las sustituciones de los átomos de carbono. En general, los ángulos exocíclicos en C1' y C4' son casi tetraédricos ($\sim 109^\circ$) mientras que en C2' y C3' incrementan, siendo entre 112° y 117° . Incluso se pueden observar cambios en los ángulos C-O dependiendo del empaquetamiento del azúcar.

1.3. Nucleósidos y Nucleótidos

En un nucleósido, el azúcar está unida a la base nitrogenada (A, T, G o C) por un enlace con el carbono C1' del azúcar y el nitrógeno de la base (N9 de las purinas o N1 de las pirimidinas) llamado enlace glucosídico, Figura (1.7).

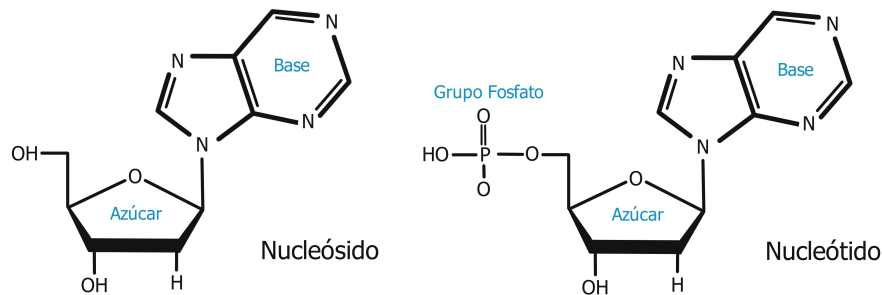


Figura 1.7: Estructura de nucleósidos y nucleótidos.

Un nucleósido es un azúcar unido a una base nitrogenada, mientras que el nucleótido está dado por la unión de un nucleósido con un grupo fosfato.

A su vez, un nucleótido está formado por la unión del nucleósido con un grupo fosfato P. El grupo fosfato es un ión poliatómico (fórmula PO_4^{3-}), aquí el fósforo está unido a 4 átomos de oxígeno mediante un enlace covalente doble (P=O) y tres enlaces covalentes sencillos (P-O) teniendo así una estructura tetraédrica.

Para formar las cadenas polinucleotídicas, los nucleótidos se unen mediante enlaces fosfodiéster (para la cadena 5'fin-3'fin, el enlace es entre el hidroxilo C3' del primero y el fosfato C5' del segundo). El enlace covalente para unir al azúcar con el grupo fosfato se da con los carbonos C5' y C3' del azúcar (C5'-O5'-P).

La estructura de un polinucleótido tiene dos finales diferentes, llamados extremos 5' y 3', los cuales tienen implicaciones químicas y biológicas. Frecuentemente, en el ADN existe un grupo hidróxilo en el extremo 3' (3' OH) y un único grupo fosfato en el extremo 5' (5' PO_4). El proceso de replicación y transcripción ocurre por la adición de un nucleósido trifosfato 5' sobre el grupo hidroxilo 3' del nucleótido final del polinucleótido [24], Figura (1.8).

Las bases de los nucleósidos son planas, y son libres de rotar alrededor del enlace glucosídico, pero existen dos orientaciones predominantes de la base en relación al azúcar.

- **Syn:** Si la base está sobre el plano de la pentosa. En las purinas encontramos el H8 sobre el azúcar; y en las pirimidinas, el H6 se encuentra sobre el anillo de la pentosa. *El ángulo de torsión entre la base y el azúcar χ se encuentra entre los 0° y los 72° .*
- **Anti:** Si el plano de la base está alejado del plano de la pentosa. Para las purinas aleja el H8 del azúcar; y el conjunto de la purina rota sobre él. *El ángulo χ se encuentra cerca de los 210° .*

El empaquetamiento del azúcar es influenciado por modificaciones en la base. En especial si el C8 de la purina o C6 de la pirimidina es sustituido por grupos que cambian entre orientaciones *syn-anti*. En general, la conformación *anti* se correlaciona con el empaquetamiento del azúcar C2'-endo, sobre todo para purinas [27].

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)[23].

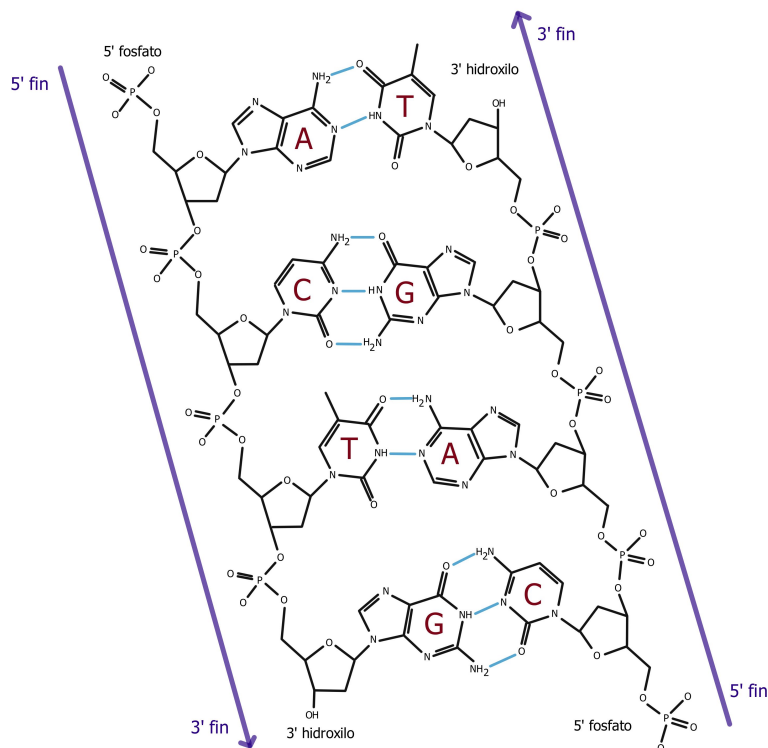


Figura 1.8: Cadena del ADN.

Las cadenas antiparalelas de azúcar-fosfato están unidas mediante emparejamiento de Watson y Crick.

1.4. Niveles de Organización de Estructuras del ADN

La estructura del ADN, representada según el modelo de Watson y Crick, es una hélice dextrógira (se enrolla a la derecha) de doble cadena antiparalela. El esqueleto de azúcar-fosfato de las cadenas constituye la parte exterior de la hélice, mientras que las bases nitrogenadas se encuentran apiladas cerca del centro de la hélice cilíndrica y forman pares unidos por puentes de hidrógeno que mantienen juntas a las cadenas. Existen alrededor de diez pares de bases por vuelta de la doble hélice. El apilamiento de las bases le da una estabilidad considerable a la doble hélice.

El ADN de doble cadena es una molécula antiparalela, lo que significa que se compone de dos cadenas que corren una junto a la otra pero en direcciones opuestas. En una molécula, el extremo 5' (el que termina con un grupo fosfato) de una cadena se alinea con el extremo 3' (el que termina con un grupo hidroxilo) de su pareja, y viceversa, Figura (1.9).

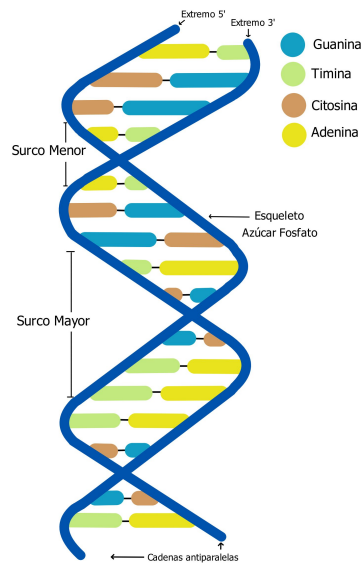


Figura 1.9: Hélice doble del ADN.

Los esqueletos azúcar-fosfato se enrollan alrededor de las bases unidas, formando la estructura de doble hélice. Se pueden apreciar los surcos que forman al enrollarse.

Todas las hélices tienen una direccionalidad, que es una propiedad que describe cómo se orientan los surcos en el espacio. La torsión de la doble hélice y la geometría de las bases crea un hueco más amplio (llamado surco mayor) y uno más estrecho (surco menor) que corren a lo largo de la molécula, Figura (1.9). Estos surcos son importantes sitios de unión para las proteínas que mantienen el ADN y regulan la actividad de los genes.

En este modelo, las cadenas se mantienen unidas por puentes de hidrógeno entre las bases nitrogenadas en cadenas opuestas, cada par de bases forma un peldaño en la escalera de la molécula. Como vimos anteriormente, estos pares no se forman por cualquier combinación, sino que se forman con correspondencias entre A-T y G-C (se conocen como pares de bases complementarias).

El emparejamiento de bases explica las reglas de Chargaff. Y puesto que una purina grande se empareja siempre con una pirimidina pequeña, el diámetro de la hélice es uniforme, de aproximadamente 2 nanómetros.

El primer nivel de organización del ADN consta de secuencias de nucleótidos unidos mediante un enlace fosfodiéster entre el hidroxilo C3' del primer nucleótido y el fosfato C5', esto da una cadena polinucleotídica 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.

El segundo nivel de organización viene representado por la hélice doble. En el ADN estas cadenas son complementarias, antiparalelas y continuas, mientras que el ARN solo cuenta con una cadena, que va de 5'fin a 3' fin en una región cercana.

El tercer y cuarto nivel de organización están relacionados con la forma en la que la hélice doble se enrolla sobre sí misma, para dar lugar a los nucleosomas, que a su vez se condensan y se enrollan nuevamente, dando como resultado la cromatina y después los cromosomas, Figura (1.10).

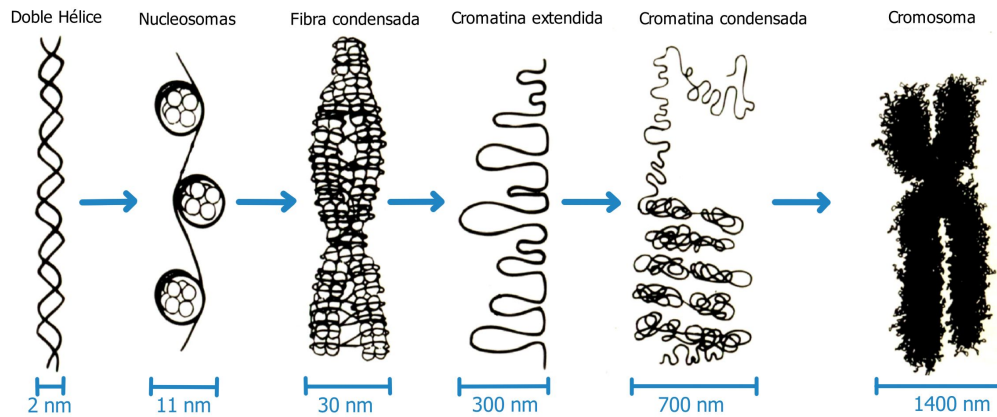


Figura 1.10: Niveles de organización del ADN.

El ADN se compacta alrededor de un conjunto de 8 proteínas llamadas histonas y forma estructuras globulares llamados nucleosomas, este complejo generado por la combinación de proteínas y ADN se denomina cromatina. La cromatina se condensa hasta alcanzar su máximo grado de compactación, el cual forma los cromosomas.

1.5. Familias de ADN

Las estructuras del ADN presentan algunas diferencias notorias que nos permiten catalogarlas en familias conformacionales; se conocen 3 tipos principales **A-ADN**, **B-ADN**, **Z-ADN**, Figura (1.11).

El B-ADN tiene mayor interés biológico, pues esta forma es la que tiene mayor interacción con las proteínas. En esta estructura no hay mucha diferencia entre el surco mayor y el surco menor, las bases son en general coplanarias y tenemos el eje helicoidal prácticamente perpendicular a las bases y que pasa por el centro de ellos. Las conformaciones del azúcar C2'endo y C3'exo son características de esta familia.

El A-ADN se presenta cuando se deshidrata levemente el B-ADN. Es una hélice dextrógira, la elevación de la hélice es de aproximadamente 2.5mm y tiene un surco mayor mucho más estrecho y profundo que la forma B. Y para esta familia, las conformaciones del azúcar C3'endo y C2'exo son características.

El Z-ADN se diferencia por enrollarse hacia la izquierda (levógiro), esta familia presenta más restricciones para obtenerla: sólo se logra obtener en soluciones salinas y en secuencias específicas como G-C-G-C-G. Esta hélice es en general más delgada y alargada con respecto a las dos familias previas para el mismo número de bases.

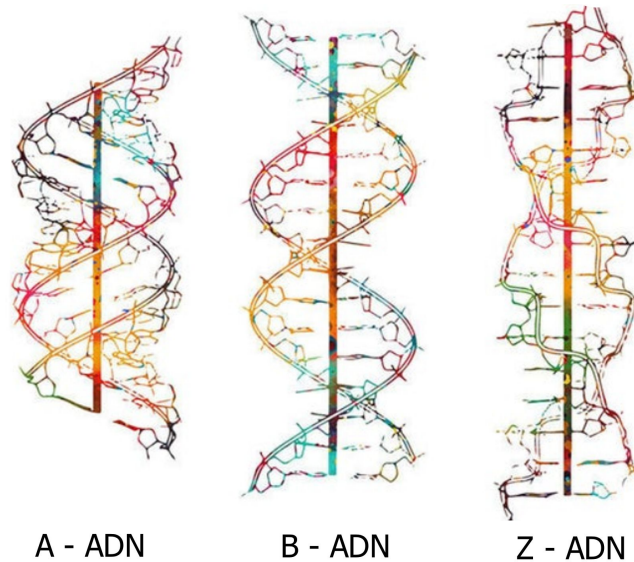


Figura 1.11: Familias de conformaciones de ADN A,B y Z. Podemos observar las diferencias estructurales entre las familias. La familia A tiene los surcos muy marcados, mientras que la familia B no tiene mucha diferencia entre el surco mayor, ambas son dextrógiras. Por otro lado la familia Z es levógira y es más delgada y alargada.

1.6. Parámetros Estructurales

La geometría molecular puede ser especificada en términos de **longitud de enlace**, **ángulo de valencia** y **ángulo de torsión**. La longitud de enlace es la distancia media entre los centros de dos átomos enlazados. Un ángulo de enlace es el ángulo formado por tres átomos enlazados consecutivamente. Para cuatro átomos unidos consecutivamente en una cadena, el ángulo de torsión es el ángulo entre el plano formado por los tres primeros átomos y el plano de los tres últimos átomos. La información estructural se encuentra presente en los ángulos de torsión (ver la Figura (1.12) para una representación de los ángulos de torsión en un dDMP) y en el empaquetamiento del anillo de azúcar. Pero podemos valernos de distintos parámetros para el estudio estructural del ADN, como la superposición de bases (*Stacking*), o los parámetros *h-Twist*, *h-Rise*, *Buckle* y *Propeller*, Figura (1.13).

Se definen 6 ángulos de torsión principales para describir la configuración del esqueleto azúcar-fosfato (δ , ϵ , ζ , α , β , γ), Figura (1.12). En particular, el ángulo δ contiene información sobre el empaquetamiento del anillo de desoxirribosa [27].

El ángulo glucosídico χ es otro ángulo de torsión importante. Se encuentra alrededor del enlace entre el átomo C1' del anillo de azúcar y el nitrógeno de su base adjunta.

En los nucleósidos, todos los ángulos de torsión excepto γ están correlacionados, pero en polinucleótidos incluso γ está involucrado. Esto significa que los cambios estructurales siguen un movimiento establecido.

Para el ADN existen los parámetros helicoidales n y h . Siendo n el número de nucleótidos por vuelta, y h la altura axial por nucleótido.

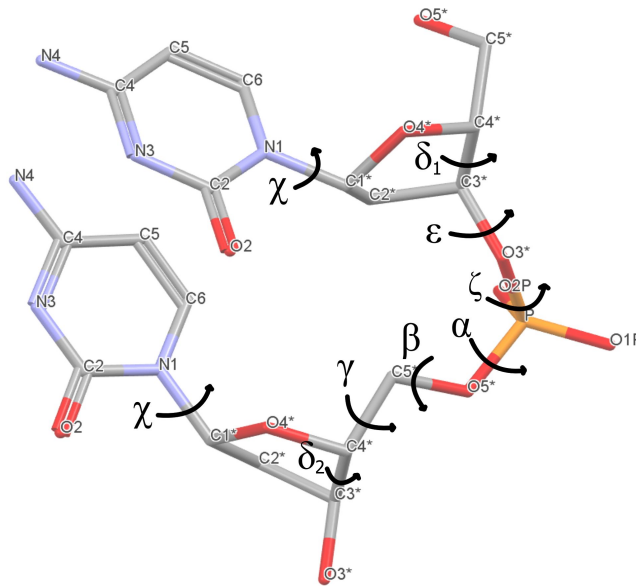


Figura 1.12: Ángulos de torsión para un desoxidinucléosido mono-fosfato (dDMP). Los ángulos de torsión principales se encuentran a lo largo de la cadena azúcar-fosfato δ , ϵ , ζ , α , β , γ ; además del ángulo entre la base nitrogenada y el azúcar llamado ángulo glucosídico χ .

Hay otros parámetros que definen características sobre la organización de las bases y que permiten una clasificación rápida de las estructuras, Figura (1.13): El antes mencionado *Stacking*, que cuantifica el área de traslape entre bases apiladas, este parámetro se mide en $[\text{\AA}^2]$. *Propeller* define el ángulo que hay entre los planos de las bases complementarias, cuando una gira respecto a la otra alrededor del eje vertical. *Buckle* define el ángulo que se produce por el doblaje de los planos de las bases, alrededor del eje horizontal. *Helical Twist* (o *h-Twist*) es el ángulo de rotación de pares de bases sucesivos alrededor del eje helicoidal. Y por último *Helical Rise* (o *h-Rise*) es la longitud que separa a pares de bases sucesivos, el promedio de estas separaciones es de 3.4\AA [27].

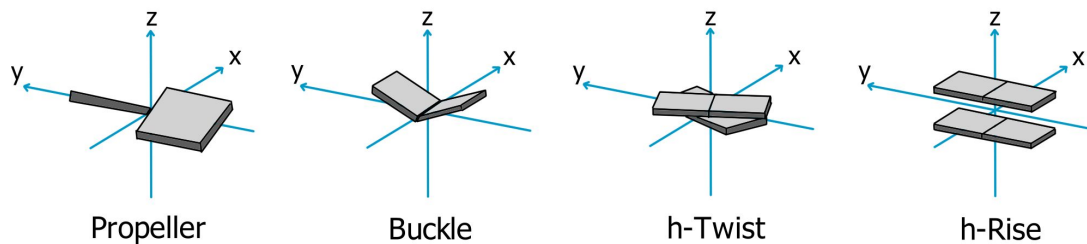


Figura 1.13: Parámetros sobre la organización de las bases. Se ilustran los parámetros, Propeller, Buckle, h-Twist y h-Rise en el sistema de coordenadas rectangulares. Cada cuadrilátero representa al plano de una base nitrogenada, para los parámetros helicoidales se usan pares de bases.

Capítulo 2

Optimización Geométrica del ADN

Actualmente existen dos modelos de estructuras moleculares que nos permiten estudiar sistemas dentro de la biofísica molecular, estos son el modelo mecánico-molecular, modelo con un bajo consumo de recursos que nos permite modelar estructuras geométricas, interacciones y diferencias energéticas. Y el modelo mecánico-cuántico, que se basa en resolver la ecuación de Schrödinger recurriendo a aproximaciones y a herramientas computacionales más potentes. Este modelo permite modelar reacciones químicas, pero sólo es óptimo para moléculas pequeñas por su alto consumo de recursos [13].

Para este trabajo usamos el método de Mecánica Molecular (MM) y los campos de fuerzas AMBER para hacer la optimización geométrica de las moléculas implicadas y así encontrar los mínimos energéticos del sistema. También usamos otros tipos de software para realizar mediciones y cálculos después de las optimizaciones, como lo es el software Mercury.

2.1. Mecánica Molecular

En el método de Mecánica Molecular se utiliza la mecánica clásica para emular sistemas moleculares que varían en tamaño y complejidad, desde moléculas orgánicas de poco peso (como los hidrocarburos), hasta sistemas biológicos complejos como ácidos nucleicos y proteínas, o materiales que contengan miles de millones de átomos. Podemos considerar al método de MM como una extensión de un simple átomo y un enlace a la representación de moléculas y sus complejos en química clásica [13].

Se considera a los átomos como esferas y además se toma como válida la *aproximación de Born-Oppenheimer* [28]. Esta aproximación acopla los movimientos electrónico y nuclear. Como la masa del núcleo del átomo es mucho mayor que la de los electrones, su velocidad es correspondientemente pequeña, así el núcleo experimenta a los electrones como si fueran una nube de carga, mientras que los electrones sienten a los núcleos como si estuvieran estáticos. De esta forma, los electrones se adaptan instantáneamente a cualquier posición de los núcleos. Este acoplamiento hace posible el trabajo en física molecular, por no tener solución analítica los problemas de más de dos cuerpos.

En el método de MM se tienen las siguientes consideraciones:

- Cada átomo es simulado como una partícula individual.
- A cada partícula se le asigna un radio, una polarizabilidad y una carga neta constante.

- Las interacciones enlazantes son tratadas como resortes, usando la *ley de Hooke*.

El radio de cada partícula usualmente es el radio de van der Waals, el cual representa la mitad de la distancia más cercana posible de aproximación entre dos átomos o moléculas en una interacción no enlazante.

La carga neta de los átomos proviene de cálculos cuánticos o experimentales, en el presente trabajo se usaron las cargas para nucleósidos de ADN correspondientes al artículo *A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules*, publicado por Wendy D. Cornell *et al.*, en 1994 [22]. Dichas cargas son usadas con los campos de fuerza *BSC1*, *FF99* y *OL15* en el presente trabajo para realizar las optimizaciones geométricas y encontrar los mínimos energéticos de cada sistema utilizando el software AMBER.

En este método la energía potencial del sistema es calculada como una función de las coordenadas nucleares utilizando **campos de fuerza**, Ecuación (2.1).

Para calcular la energía del sistema, se hace una sumatoria de energía dada por diferentes niveles de cálculos. Es decir, la sumatoria de términos de energía que corresponden a las contribuciones de diferentes características físicas y la sumatoria de contribuciones de los electrones. La forma básica de la función de energía potencial en MM incluye términos (enlazantes¹ y no enlazantes²) que describen la interacción electrostática de grupos cargados y la fuerza de van der Waals.

La energía potencial del sistema se representa como una función de las coordenadas de los átomos (r) que es la suma de los componentes *enlazantes* y *no enlazantes*, donde los componentes de cada contribución se desglosan en las sumas de la siguiente ecuación, cada uno siendo a su vez una suma de muchas contribuciones:

$$\Delta E(r) = \sum E_l + \sum E_\theta + \sum E_\phi + \sum E_e + \sum E_{vdW}, \quad (2.1)$$

donde E_l representa los enlaces químicos, E_θ los ángulos de valencia, E_ϕ los ángulos de torsión, E_e las interacciones electrostáticas y E_{vdW} las interacciones de van der Waals.

Los términos de energía dependen de las posiciones de los átomos y de los parámetros (constantes ajustables). Además, como la función depende de las coordenadas nucleares de los átomos, una variación en la posición implica un cambio en la energía potencial. Así, cuando una estructura representa un mínimo en la superficie de la energía potencial del sistema, decimos que la estructura está optimizada.

Como mencionamos antes, las interacciones de enlaces entre átomos son tratadas como resortes, siguiendo la ley de Hooke, con una distancia de equilibrio igual a la longitud de enlace calculada experimentalmente. La longitud de enlace es la distancia media entre los núcleos de dos átomos unidos mediante un enlace químico en una molécula.

Así, los dos **primeros términos** de la Ecuación (2.1), que son los cambios de energía debido a variaciones de longitud, estiramiento, ángulos de enlaces y desviaciones, se pueden escribir como:

$$\begin{aligned} E_{l_i} &= k_{l_i}(l_i - l_i^0)^2 \\ E_{\theta_i} &= k_{\theta_i}(\theta_i - \theta_i^0)^2, \end{aligned} \quad (2.2)$$

donde l_i es la longitud, θ_i el ángulo; l_i^0 y θ_i^0 son los valores de equilibrio para longitud y ángulo en el enlace; k_{l_i} y k_{θ_i} son el estiramiento y la desviación de las constantes de fuerza, respectivamente. Todos estos parámetros dependen de los tipos de átomos involucrados en el

¹Interacciones atómicas que están unidas por enlaces covalentes.

²Interacciones entre pares de átomos no covalentes.

enlace o en el ángulo de valencia.

El **tercer término** de la Ecuación (2.1) es el cambio de energía de torsión, que es responsable de interacciones de capas de electrones de dos átomos (a y d) y de dos enlaces ($a-b$ y $c-d$) los cuales están conectados a través de un enlace químico ($b-c$). Este término se puede expresar de la siguiente forma:

$$E_{\phi_i} = k_{\phi_i}[1 + \cos(n_i\phi_i - \delta_i)] \quad (2.3)$$

Aquí ϕ_i es el ángulo i de rotación alrededor del enlace químico ($b-c$), es decir el ángulo entre el plano definido por (a, b, c) y el plano (b, c, d). Esta función es periódica, y los parámetros k_{ϕ_i} , n_i (multiplicidad) y δ_i (fase del ángulo) pueden ser estimados con datos experimentales de la molécula considerada.

Los **últimos dos términos** representan las interacciones no enlazantes de energía potencial. Esta es la energía de no enlace entre todos los pares de átomos no relacionados el uno al otro o al mismo tercer átomo. Esta energía se descompone en energía electrostática y energía de van der Waals.

Los términos no enlazantes son más intensivos computacionalmente.

Las interacciones electrostáticas se pueden representar mediante una energía potencial de Coulomb:

$$E_e = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}, \quad (2.4)$$

donde q_i son las cargas atómicas usadas para representar el enlace químico, r_{ij} es la distancia entre los átomos i y j . La energía total es la suma de todas las combinaciones de pares de átomos, y generalmente excluye las interacciones que existen a más de un enlace de distancia.

Por último, las interacciones de van der Waals son las fuerzas atractivas o repulsivas entre moléculas distintas a las debidas a un enlace intermolecular o a la interacción electrostática, por lo que el término incluye:

- La fuerza entre dos dipolos permanentes³.
- La fuerza entre un dipolo permanente y un dipolo inducido⁴.
- La fuerza entre dos dipolos inducidos instantáneamente⁵.

Todos los tipos de fuerzas de van der Waals dependen fuertemente del entorno debido a que se originan de interacciones de dipolos instantáneos e inducidos (interacciones intermoleculares).

2.2. Campo de Fuerza

Un **campo de fuerza** es un método computacional utilizado para estimar las fuerzas entre los átomos dentro de las moléculas. La combinación de una función de energía potencial (como la del Sección (2.1)), con todos los parámetros que involucra, constituye un *campo de fuerza*.

³Si las interacciones son entre moléculas que están polarizadas de manera permanente se conocen como fuerzas de Keesom [29].

⁴Cuando un dipolo inducido interactúa con una molécula que tiene momento dipolar permanente se le conoce como fuerza de Debye [29].

⁵Si las interacciones son entre dos dipolos que están inducidos en los átomos o moléculas, se conocen como fuerzas de London [29].

En la práctica existe una conexión cercana entre los campos de fuerza y los códigos computacionales que los implementan, aunque esto se vuelve cada vez menos cierto de lo que solía ser [30].

Diferentes campos de fuerza están diseñados para diferentes propósitos, y todos están implementados en diferentes softwares computacionales.

El software computacional y campo de fuerza AMBER (Assisted Model Building with Energy Refinement) [3] es uno de los paquetes más usados para una amplia gama de moléculas orgánicas, incluyendo AN, proteínas y ADN. Este software es el que se ha seleccionado para el presente trabajo. El campo de fuerza contiene el arreglo común *minimalista* para los términos de energía, Ecuaciones (2.1) a (2.4); y los parámetros fueron ajustados mediante datos de mecánica cuántica y experimentales.

Los campos de fuerza son actualmente una herramienta común para la investigación estructural en moléculas biológicas, pues los campos de fuerza combinados con nuevos algoritmos computacionales permiten reproducir de manera más exacta las observables experimentales.

Las cargas eléctricas de los átomos en los AN fueron determinadas usando cálculos con el potencial electrostático restringido (RESP, por sus siglas en inglés) adecuado, y han logrado reproducir con un muy buen grado de precisión las energías de interacción y conformacionales de moléculas pequeñas. Estas cargas fueron publicadas en el artículo de Cornell et al.[22] (mencionado en la Sección (2.1)), y son las que usaremos en el presente trabajo para desoxinucleósidos.

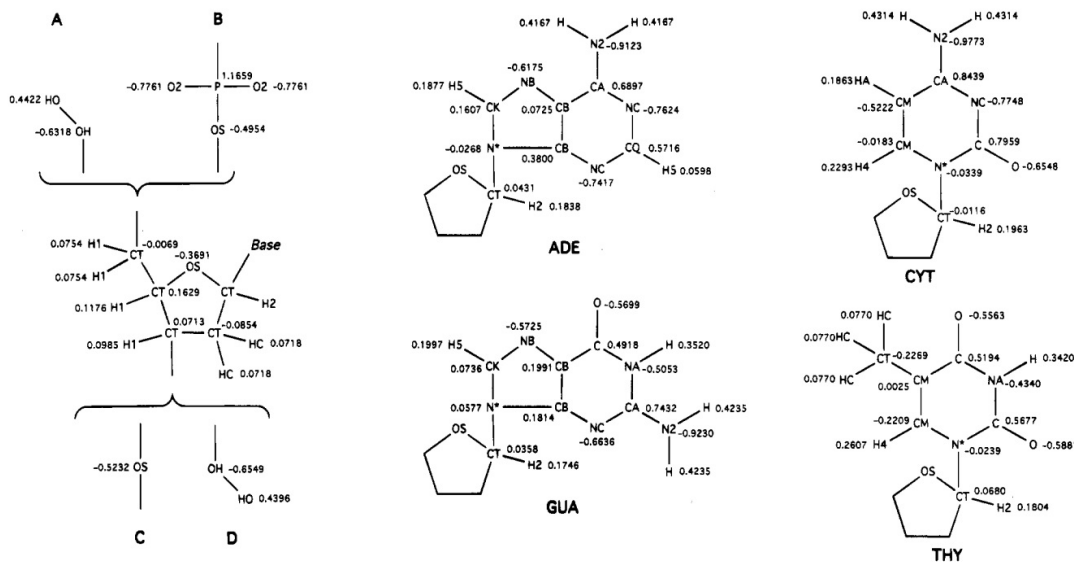


Figura 2.1: Cargas para ADN usadas en el paquete AMBER, artículo de Cornell et al.[22] Podemos observar las cargas para nucleósidos (fragmento A y D con el azúcar) y nucleótidos (fragmento B y D) de ADN. Las cargas para las bases y el C1' y H1' se muestran por separado para poder combinarlas con el azúcar. El extremo 5' de la cadena corresponde al fragmento A, C y el azúcar; y el extremo 3' corresponde a la combinación de B y D con el azúcar; mientras que una estructura central en la cadena corresponde a la combinación de B y C con el azúcar.

En los campos de fuerza AMBER los parámetros son ajustados para que reproduzcan propiedades estructurales. En particular, los parámetros de van der Waals fueron ajustados a propiedades de líquidos, y añadiendo términos extra de Fourier se consiguieron mejorar notablemente los

potenciales para los ángulos de torsión.

Los campos de fuerzas AMBER incorporan una función de energía potencial relativamente simple, que nos lleva de la Ecuación (2.1) a la siguiente expresión:

$$V(l, \theta, \phi) = \sum_{\text{enlaces}} k_l (l - l_0)^2 + \sum_{\text{angulos}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{torsiones}} k_\phi [1 + \cos n\phi] + \sum_{j=1}^{N-1} \sum_{i=j+1}^N \left[\frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \epsilon_{ij} \left(\frac{r_{0ij}}{r_{ij}}^{12} - 2 \frac{r_{0ij}}{r_{ij}}^6 \right) \right], \quad (2.5)$$

donde l es la longitud del enlace y l_0 es su valor en equilibrio; θ el ángulo de valencia y θ_0 su valor en equilibrio; ϕ el ángulo de torsión; y los parámetros $k_l, k_\theta, k_\phi, n, \epsilon_{ij}, r_{0ij}$ son constantes. En la Figura (2.2) podemos ver estos potenciales esquematizados.

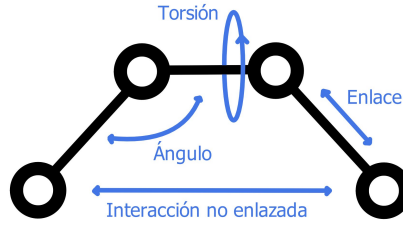


Figura 2.2: Potenciales usados en MM.

La figura muestra un esquema de los potenciales considerados en el método de MM, Ecuación (2.1). Los potenciales de enlace, ángulo y torsión son potenciales enlazantes (pares de átomos enlazados), mientras que para los pares de átomos no enlazados tenemos los potenciales eléctrico y de van der Waals.

Los primeros tres términos de la Ecuación (2.5) se derivan de las Ecuaciones (2.2) y (2.3). El último término, la doble sumatoria sobre i, j , representa la energía de los pares de átomos no enlazados, y se puede descomponer en dos términos, el de energía electrostática (primer término de la suma) y de van der Waals (segundo término de la suma).

Para la expresión de la energía electrostática de la Ecuación (2.4), se asume que las cargas debidas a los protones y electrones en un átomo pueden ser representadas por una carga puntual (o en el caso de conjuntos de parámetros, un pequeño número de cargas puntuales).

La energía de van der Waals en la Ecuación (2.5) se calcula usando la distancia de equilibrio r_{0ij} y la profundidad del pozo, ϵ . El factor 2 en el segundo término de la expresión de dicha energía, asegura que la distancia de equilibrio es r_{0ij} . Esta energía se puede escribir también en términos de sigma, con $r_{0ij} = 2^{1/6}\sigma$.

Estos dos últimos potenciales (electrostático y de van der Waals) son interacciones entre átomos no enlazados (V_{nb}). Donde k_l es la constante de fuerza de estiramiento, k_θ es la constante de fuerza de doblamiento y k_ϕ la constante del peso de la barrera de torsión.

2.3. Optimización Geométrica con AMBER

Nos referimos con optimización geométrica a aquel proceso que, mediante una serie de pasos, intenta encontrar la configuración espacial de los núcleos atómicos que resulte en un mínimo de energía de la molécula.

El campo de fuerza AMBER ha sido desarrollado para la dinámica molecular de macromoléculas y es utilizado para la minimización de energía. Por lo tanto, las coordenadas de todos los átomos se consideran como variables independientes.

De esta forma, la optimización geométrica depende de la geometría de partida, de una función de energía potencial y de la condición impuesta para obtener un gradiente aceptable en un determinado número de pasos [2].

El software computacional AMBER usado en el presente trabajo, sigue una serie de algoritmos computacionales [20] que podemos resumir en los siguientes pasos:

- Calcula la energía de un punto de la distribución espacial de átomos inicial.
- Se modifican las coordenadas de un subconjunto de átomos y se calcula un nuevo valor de energía para la nueva configuración.
- Las derivadas de primer y segundo orden de la energía con respecto a las posiciones indican la dirección y la magnitud de la siguiente conformación geométrica.
- Se realiza el incremento en la distribución y se determina el nuevo valor de energía.

Este proceso se repite hasta que se alcanza un valor de convergencia previamente fijado.

2.3.1. Procedimiento de Optimización Usando el Software AMBER

En el paquete computacional AMBER existen varias herramientas, programas y comandos que podemos usar para hacer la optimización geométrica de nuestras estructuras. Para empezar debemos preparar la estructura que vamos a utilizar (este se lleva a cabo en detalle en el Apéndice (A), realizando una búsqueda y selección en el banco de datos NDB (Nucleic acid DataBase, por sus siglas en inglés), y posteriormente *recortando* el fragmento que sea de nuestro interés. En el presente trabajo usamos fragmentos cortos de ADN, más específicamente desoxinucleósidos.

Se utilizaron tres campos de fuerza distintos implementados en el software, estos son *BSC1*, *FF99* y *OL15*. Buscaremos con cuál de ellos podemos representar mejores resultados para cada tipo de estructura.

Dentro del paquete AMBER existe un editor llamado **XLeap: Universe Editor** con el cuál se ejecutará primero el campo de fuerza que debemos llamar desde la dirección que usemos como repositorio de AMBER (*AMBERHOME*) usando una terminal (o shell) en nuestro sistema operativo. Con los siguientes comandos llamamos a ejecutar el editor XLeap con los campos de fuerza *FF99*, *BSC1* y *OL15*, respectivamente.

```
xleap -s -f $AMBERHOME/dat/leap/cmd/oldff/leaprc.ff99
xleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.DNA.bsc1 -s -f
           $AMBERHOME/dat/leap/cmd/leaprc.water.tip3p
xleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.DNA.OL15 -s -f
           $AMBERHOME/dat/leap/cmd/leaprc.water.tip3p
```

Una vez ejecutado el editor, usaremos su consola para dar las instrucciones siguientes. Primero necesitamos definir una variable (llamada *a* por ejemplo) que cargue el archivo con el fragmento que preparamos previamente en formato .pdb que vamos a optimizar.

```
a=loadpdb "archivo".pdb
edit a
```

Con la segunda línea del comando ejecutamos el editor **Unit Editor** de XLeap para nuestra variable *a*. Este editor nos permite visualizar la molécula que tenemos en nuestro archivo de coordenadas, y tiene muchas herramientas para modificar la distribución espacial de nuestra estructura, Figura (2.3).

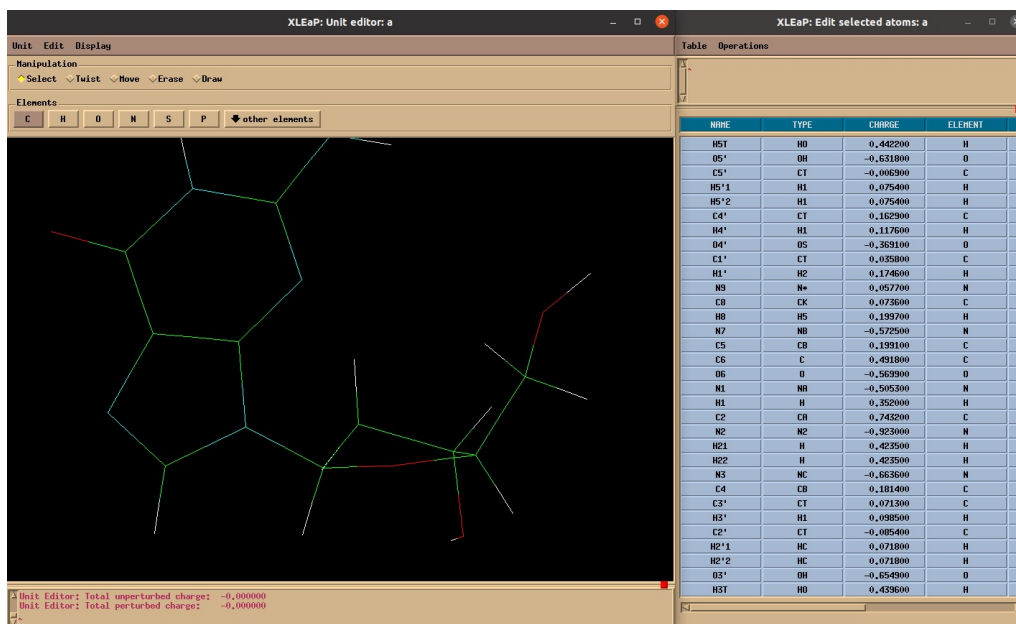


Figura 2.3: Interfaz de Unit Editor de XLeap, software AMBER.

En la parte izquierda podemos visualizar la estructura de la molécula, donde tenemos también herramientas para su edición. En la derecha de la figura se encuentra la tabla con las cargas para cada átomo seleccionado.

En el Unit Editor podemos modificar los átomos de la molécula, añadir nuevos átomos, borrar otros, etc. Pero como ya preparamos las coordenadas específicas que vamos a usar para la optimización entonces pasamos a la herramienta *Unit >Add H & Build* que agrega los hidrógenos faltantes (habrá hidrógenos faltantes en los nucleósidos con recortes o con cambio de base).

Siempre tendremos que revisar que la carga neta de la molécula sea neutra, de lo contrario podría haber un error con las cargas individuales de los átomos, esto se hace con la herramienta *Unit >Calculate net charge*, en la pestaña inferior de salida nos indicará la carga neta de la molécula.

Algo muy importante es revisar que la carga de los átomos, sus nombres y sus tipos correspondan a los acordados en el artículo de Cornell et al. [22], seleccionando toda la molécula y con la herramienta *Edit >Edit selected atoms* se desplegará una tabla para editar los átomos, donde podemos modificar los valores de las cargas, nombres y tipos.

Al terminar de hacer los ajustes necesarios, debemos revisar que la molécula no tenga errores estructurales (por ejemplo átomos repetidos, errores en enlaces o en ángulos de torsión), para esto se usa la herramienta *Unit >Check Unit*.

Cuando todo esté correcto cerramos el Unit Editor: *Unit >Quit*.

Una de las modificaciones estructurales muy frecuente en el presente trabajo, usando el editor *Unit Editor* de XLeap, es el *Cambio de Base*. Consiste en usar las coordenadas de la desoxirribosa de una estructura, que tenga parámetros que nos interesen, pero cambiando la base nitrogenada a la que está conectada. Este cambio, detallado en el Apéndice (B), nos permite explorar diferentes conformaciones usando la misma distribución espacial de una estructura.

Regresando a la consola principal de XLeap, usaremos las líneas:

```
saveamberparm a "nombre".tipo "archivo".coor  
savepdb a "nombre-optimizado".pdb
```

La primera genera y guarda dos archivos, uno con la información de los parámetros que se usarán para optimizar la estructura (con formato *.tipo*) y el otro con la información de las coordenadas (con formato *.coor*).

La segunda línea guarda el nuevo archivo inicial (con formato *.pdb*) con las modificaciones realizadas.

Para realizar la optimización geométrica de una molécula, necesitamos los dos archivos generados con el programa XLeap (archivos *.tipo* y *.coor*), y otros dos programas del paquete AMBER que ejecutaremos directamente desde la terminal de nuestro sistema.

Primero la instrucción **SANDER**, que realiza la optimización geométrica mediante los pasos descritos en la Sección (2.3) y calcula la energía de la estructura:

```
sander -i 999.in -o "energía".out -p "archivo".tipo -c "archivo".coor -r "nombre".outcoor -0
```

Para hacer la optimización, el programa utiliza un archivo de inicio con parámetros establecidos por nosotros (como el número máximo de ciclos que llevará, el radio de corte, etc.) al que llamamos *999.in*, ya que 999Å es el radio de corte⁶ que seleccionamos para estos fragmentos pequeños, además de los dos archivos generados por XLeap (*.tipo* y *.coor*).

Al realizar el proceso computacional, el programa nos genera dos archivos, el primero (formato *.out*) es donde se guardan los valores de la energía en cada paso, dándonos un resultado final de la energía minimizada. El otro archivo (formato *.outcoor*) es donde se guardan las nuevas coordenadas de la molécula optimizada.

En este paso también podemos hacer uso de otro archivo inicial (formato *.in*) para aplicar una restricción a uno o varios ángulos de torsión de nuestra molécula al momento de optimizar. Este procedimiento es detallado en el Apéndice (C). En el presente trabajo lo usamos para evitar *errores* estructurales al momento de trabajar con nucleósidos, ya que al no tener un grupo fosfato en la molécula inicial algunos ángulos pueden tomar valores *antimaturales* que representan un mínimo de energía artificial, por lo que no son relevantes para nuestra investigación.

Por último la instrucción **AMPDB** ejecuta un programa de análisis que sirve para extraer la información resultante de las modificaciones y de la optimización de la molécula.

```
ampdb -p "archivo".tipo -c "archivo".outcoor > "nombre-optimizado".pdb
```

Ejecutando esa línea, el programa usa los archivos *.tipo* y *.outcoor* y genera un archivo (formato *.pdb*) que corresponde a la estructura optimizada.

⁶El radio de corte es la distancia máxima en la que el programa deja interactuar a los átomos entre ellos, para este tipo de estructuras se debe elegir un radio de corte que englobe a todos los átomos de la molécula.

2.4. Mediciones y Cálculos Finales

Después de hacer la optimización geométrica de las estructuras con el software AMBER, es necesario realizar mediciones y cálculos para poder expresarlos en tablas (Capítulo 4. Resultados) para su respectivo análisis.

Nos ayudaremos de otro software llamado **Mercury** para la visualización y mediciones, y de algún programa de tipo *hoja de cálculo* para algunos cálculos y presentación.

2.4.1. Mediciones con Mercury

El software Mercury⁷ (Crystal Structure Visualisation, Exploration and Analysis software)[31] ayuda a la visualización en tres dimensiones de estructuras moleculares, y sirve entre otras cosas, para el análisis del empaquetamiento de cristales y de interacciones intermoleculares. Este software puede leer archivos con formato .pdb (entre muchos otros) y este es el que usaremos para realizar las mediciones que analizaremos más adelante, mediciones de los ángulos de valencia y de torsión del anillo de desoxirribosa, ángulo de torsión glucosídico y de los ángulos de torsión del fragmento de esqueleto azúcar-fosfato del desoxinucleósido.

Primero abrimos nuestro archivo .pdb de la estructura optimizada geoméricamente. Seleccionamos este software ya que cuenta con una amplia gama de herramientas, y es sumamente amigable con las opciones de visualización y la simplicidad de su interfaz.

Para el presente trabajo, usamos las herramientas que nos permiten medir ángulos de valencia, ángulos de torsión y distancia entre átomos, Figura (2.4).

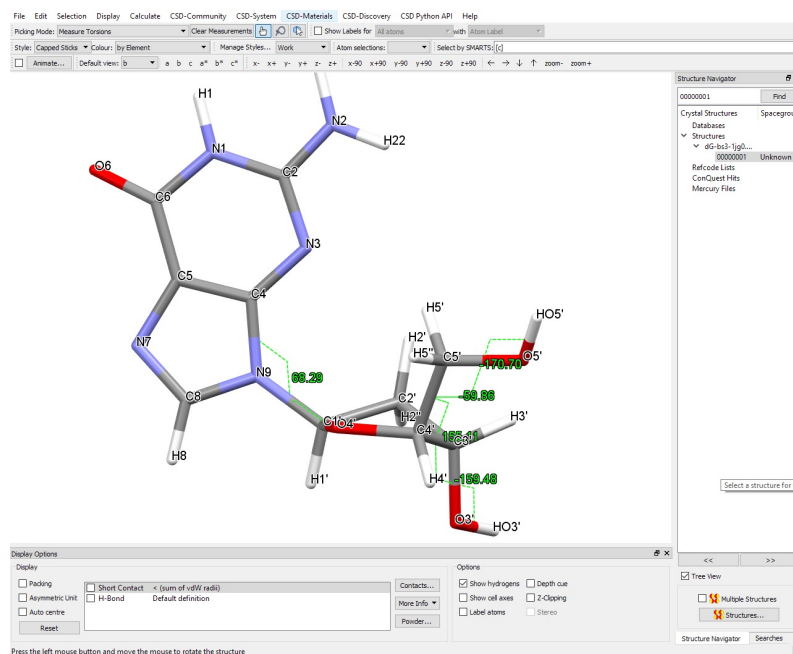


Figura 2.4: Interfaz del software Mercury.

En la figura se están midiendo ángulos de torsión con la herramienta *Measure Torsions*. Tenemos diferentes herramientas de para medición y visualización en este software.

⁷Desarrollado por Cambridge Crystallographic Data Centre, 2001 [31].

Para cada estructura optimizada realizamos las siguientes mediciones y verificamos que no existiera ningún error, en los casos con errores se revisó el proceso de optimización y se propusieron soluciones, que incluyen por ejemplo optimizar con restricciones de ángulos de torsión o realizar cambios de base nitrogenada.

- Revisamos que no existiera ningún posible enlace de hidrógeno artificial. Con ayuda de la herramienta *Measure Distances* medimos los pares de átomos que sean muy cercanos sin enlaces covalentes. Esto debido a que las distancias con las que se pueden formar puentes de hidrógeno están alrededor de los 2.5Å.
- Usando la herramienta *Measure Angles* medimos los ángulos de valencia del anillo de desoxirribosa alrededor de los átomos O4', C1', C2', C3' y C4'.
- Con la herramienta *Measure Torsions* medimos los ángulos de torsión del anillo de desoxirribosa ν_0 (C4'-O4'-C1'-C2'), ν_1 (O4'-C1'-C2'-C3'), ν_2 (C1'-C2'-C3'-C4'), ν_3 (C2'-C3'-C4'-O4') y ν_4 (C3'-C4'-O4'-C1').
- Finalmente con la última herramienta *Measure Torsions* medimos los ángulos de torsión que corresponden al fragmento del esqueleto azúcar-fosfato presentes en nuestro nucleósido γ , δ , β y ϵ ; y el ángulo de torsión glucosídico χ , Figura (1.12).

Después de pasar la revisión y de realizar las mediciones correspondientes, capturamos estos últimos datos en un software de hoja de cálculo para hacer otros cálculos pertinentes.

2.4.2. Cálculos de Parámetros de Pseudorrotación y Distancias C2', C3'

En nuestra hoja de cálculo (por ejemplo Microsoft Excel) haremos algunos cálculos con las mediciones obtenidas en la sección anterior. Como lo son la fase del ángulo de pseudorrotación P , la amplitud de pseudorrotación τ_M y las distancias de $C2'$ y $C3'$ al plano formado por C1'-O4'-C4'.

Para calcular la fase del ángulo P y la amplitud de pseudorrotación τ_M , se utilizan los ángulos de torsión del anillo de desoxirribosa ν_0 , ν_1 , ν_2 , ν_3 y ν_4 . Haciendo una serie de cálculos sencillos derivados de la Ecuación (1.1) en la hoja de cálculo, obtenemos los valores de P y τ_M para cada estructura.

El cálculo de las distancias de $C2'$ y $C3'$ al plano del anillo de desoxirribosa es un poco más laborioso, pues debemos hacer uso del archivo .pdb de la estructura optimizada para obtener las coordenadas de los átomos involucrados (C4', O4', C1', C2', C5' y C3'). Ya con estas coordenadas simplemente hacemos una serie de operaciones para encontrar la ecuación del plano C4'-O4'-C1' y posteriormente la distancia de $C2'$ y $C3'$ a este plano.

Después de presentar estas mediciones y cálculos en una tabla organizada, podemos analizar y catalogar de mejor manera nuestras estructuras optimizadas.

Capítulo 3

Resultados

A continuación, se mostrarán las tablas obtenidas mediante la optimización con campos de fuerza AMBER (BSC1, FF99 y OL15) de distintas estructuras de desoxirribosa y desoxinucleósidos, así como las figuras correspondientes a mínimos de energía locales y mínimos de energía globales.

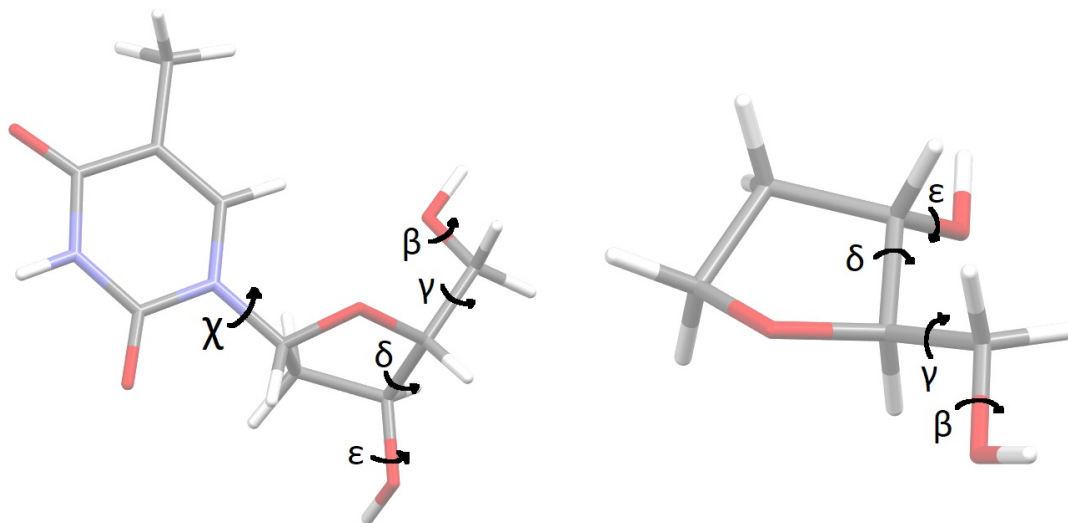


Figura 3.1: Ángulos de torsión para desoxinucleósidos y desoxirribosa. Se observan los ángulos de torsión principales que corresponden a la cadena azúcar-fosfato β , γ , δ y ϵ , además del ángulo glucosídico χ .

Denominamos mínimos de energía locales a las estructuras optimizadas encontradas con los diferentes campos de fuerza, estos fueron ordenados de menor energía a mayor energía en las tablas presentadas y se les asignó un número correspondiente a su posición en la tabla. Los mínimos de energía globales son las estructuras optimizadas con energía más baja para cada campo de fuerza, es decir los que se encuentran en la primera fila de datos de cada tabla, representados con el número 1 en cada tabla.

En la **primera columna** de las tablas tenemos el número asignado a la estructura minimizada, se ordenaron los mínimos dependiendo del valor su energía final, de menor a mayor.

En la **segunda columna** se expresa la energía final obtenida del proceso de optimización

geométrica, esta energía engloba todas las interacciones que tiene campo de fuerza con la molécula (ver Sección (2.2)). La energía está expresada en kilocalorías por mol (kCal/mol).

Las **siguientes 5 columnas** son los valores de los ángulos de torsión χ , β , γ , δ y ϵ , respectivamente, Figura (1.12). Dichos ángulos fueron medidos usando el software Mercury [31], y se encuentran expresados en grados.

En las columnas **P** y τ_M se expresan la fase del ángulo de pseudorrotación y la amplitud de pseudorrotación, respectivamente, Sec. (2.4.2).

Las columnas **C2'** y **C3'** indican la distancia que tienen los carbonos C2' y C3' del azúcar, respectivamente, con respecto al plano formado por los otros 3 átomos del anillo de la ribosa (es decir, el plano formado por los átomos C1'-O4'-C4'). Esta distancia está expresada en Angstroms en las tablas, se denota como negativa si la dirección es hacia C5' (endo) y positiva si es en la dirección contraria. Sec. (2.4.2).

La **última columna** indica el empaquetamiento o *puckering* del anillo de azúcar. Este empaquetamiento está designado siguiendo el círculo de pseudorrotación, Fig. (1.6).

En las imágenes se mantiene horizontal el plano formado por los átomos **C1'-O4'-C4'**, para tener una mejor visualización de las distancias de **C2'** y **C3'** con respecto a dicho plano, y para poder observar los diferentes ángulos de torsión.

3.1. Mínimos de energía de desoxicitidina

Veamos primero los resultados obtenidos con las pirimidinas. Para desoxicitidina (dCyd) los resultados usando el campo BSC1 se encuentran en la tabla 3.1.

Tabla 3.1: Mínimos locales encontrados con el campo BSC1 para desoxicitidina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^\circ$)	β ($^\circ$)	γ ($^\circ$)	δ ($^\circ$)	ϵ ($^\circ$)	τ_M	P ($^\circ$)	C2' (\AA)	C3' (\AA)	Puckering
1	-120.55	215.4	179.7	55.8	143.1	156.9	38.2	158.9	-0.62	-0.04	C2'-endo
2	-118.82	210.7	276.2	55.7	145.9	149.5	36.7	165.7	-0.52	0.08	C2'-endo
3	-118.76	205.4	181.0	55.9	82.7	174.6	37.7	15.4	0.06	-0.55	C3'-endo
4	-118.63	204.3	295.1	300.5	146.7	163.1	42.5	155.6	-0.73	-0.11	C2'-endo
5	-118.44	201.5	187.3	300.3	150.0	168.2	42.7	159.3	-0.69	-0.04	C2'-endo
6	-117.62	197.5	271.1	48.0	81.8	199.6	39.3	7.5	0.20	-0.45	C3'-endo
7	-115.39	197.3	178.0	180.1	147.6	151.9	41.9	159.8	-0.67	-0.03	C2'-endo
8	-112.41	197.3	176.5	179.6	80.0	160.3	40.0	19.3	-0.01	-0.64	C3'-endo

Con base en la tabla 3.1, podemos agrupar los mínimos locales encontrados con el campo BSC1 con respecto a valores de los ángulos de torsión γ y δ , y también con respecto al empaquetamiento del azúcar.

Con el campo BSC1 seleccionamos 8 estructuras con mínimos de energía locales que consideramos interesantes por la variedad en los parámetros que contienen. En cuanto al empaquetamiento del azúcar solamente tenemos dos puckerings diferentes con este campo de fuerza, C2'-endo (para valores del ángulo de pseudorrotación cercanos a 160°) y C3'-endo (para valores del ángulo de pseudorrotación entre 0° y 36°). Estas conformaciones son las más comunes en el ADN. También cabe mencionar que, para los mínimos locales con puckering C2'-endo, el átomo C2' se encuentra muy alejado del plano C1'-O4'-C4' (con distancias alrededor de 0.6\AA en dirección hacia C5'),

mientras que el átomo C3' se encuentra prácticamente alineado a dicho plano. Por otra parte, para los mínimos locales con puckering C3'-endo, pasa lo contrario, el átomo C2' se encuentra casi alineado con el plano y el átomo C3' se aleja mucho del plano (con distancias cercanas a 0.5\AA en dirección a C5').

En la tabla 3.1 vemos tres grupos respecto al ángulo γ , los mínimos 1, 2, 3 y 6, tienen un valor del ángulo γ alrededor de los 60° que corresponden a la región **+gauche**. Los mínimos no.7 y no.8, tienen el ángulo γ cercano a 180° , a esta región se le conoce como **trans**. Mientras que los mínimos 4 y 5, tienen un valor de γ alrededor de los 300° (-gauche). Los valores de cada una de estas tres regiones corresponden a clases conformacionales específicas de dDMP en la cadena del ADN.

Con respecto al ángulo δ , vemos que éste cambia de región dependiendo del puckering. Para las estructuras que tienen puckering C2'-endo, el ángulo δ toma valores alrededor de los 145° , mientras que para las estructuras con puckering C3'-endo, δ tiene valores cercanos a 80° .

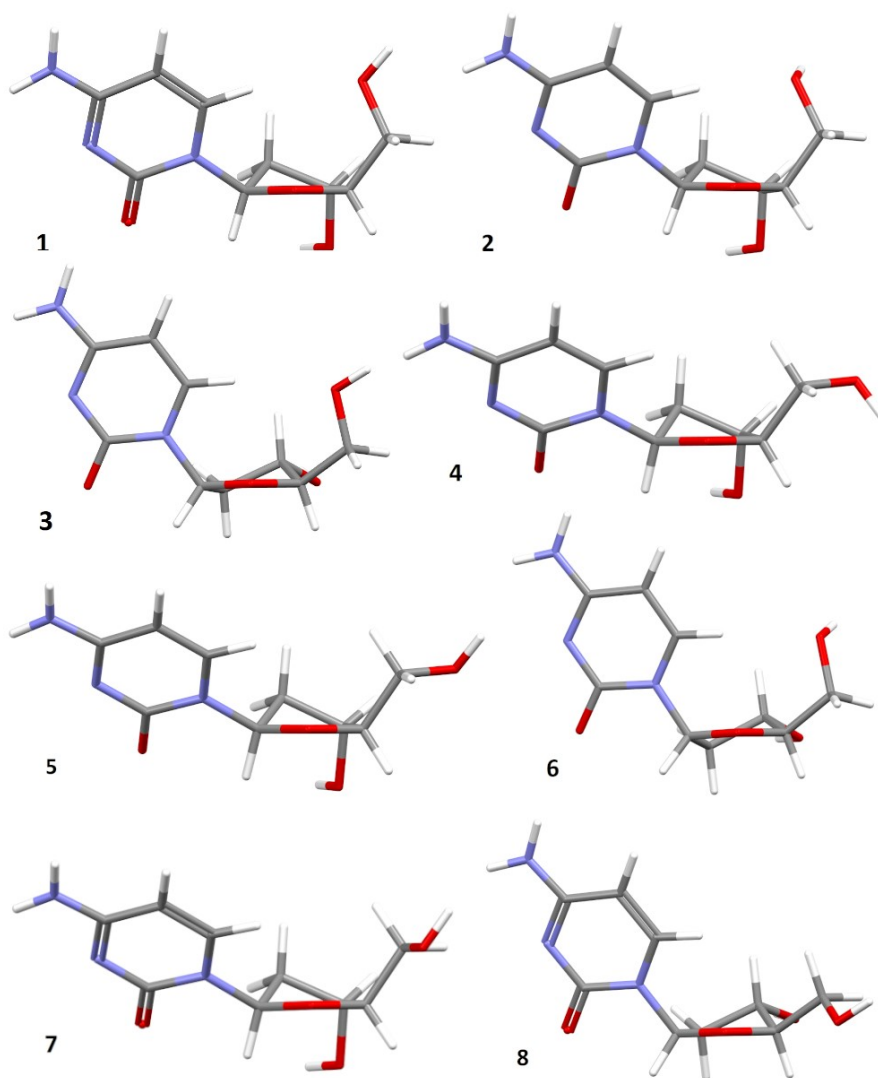


Figura 3.2: Figura comparativa de los mínimos locales encontrados con el campo BSC1 para desoxicitidina. Los datos de cada estructura se muestran en la tabla 3.1.

Resultados

3.1 Mínimos de energía de desoxicitidina

En las tablas 3.2 y 3.3, tenemos los resultados de los mínimos locales selectos para desoxicitidina obtenidos con los campos de fuerza FF99 y OL15, respectivamente.

Tabla 3.2: Mínimos locales encontrados con el campo FF99 para desoxicitidina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-122.16	216.9	178.9	52.5	138.2	170.5	34.7	152.3	-0.64	-0.14	C2'-endo
2	-121.49	208.6	180.7	53.9	85.4	179.9	33.2	17.7	0.02	-0.51	C3'-endo
3	-119.03	196.4	182.3	191.9	132.9	165.8	42.3	42.7	-0.92	-0.45	C4'-exo

Tabla 3.3: Mínimos locales encontrados con el campo OL15 para desoxicitidina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-120.86	217.1	178.2	55.1	138.3	170.3	34.9	151.8	-0.65	-0.15	C2'-endo
2	-119.54	191.7	271.5	50.4	85.2	169.0	37.2	1.8	0.28	-0.34	C3'-endo
3	-116.35	187.2	182.0	181.8	132.0	165.0	42.9	136.5	-0.94	-0.47	C1'-exo
4	-114.80	182.0	162.8	170.3	99.6	170.1	36.3	331.9	0.67	0.16	C2'-exo

En la Fig. (3.3) podemos observar las conformaciones de los mínimos globales de desoxicitidina para los tres campos de fuerza usados. Las estructuras correspondientes a los mínimos globales obtenidos por los tres campos de fuerza tienen azúcar C2'-endo y los valores del ángulo gamma se encuentran cerca de los 60° , como en la mayoría de dDMPs en cristales de fragmentos del ADN.

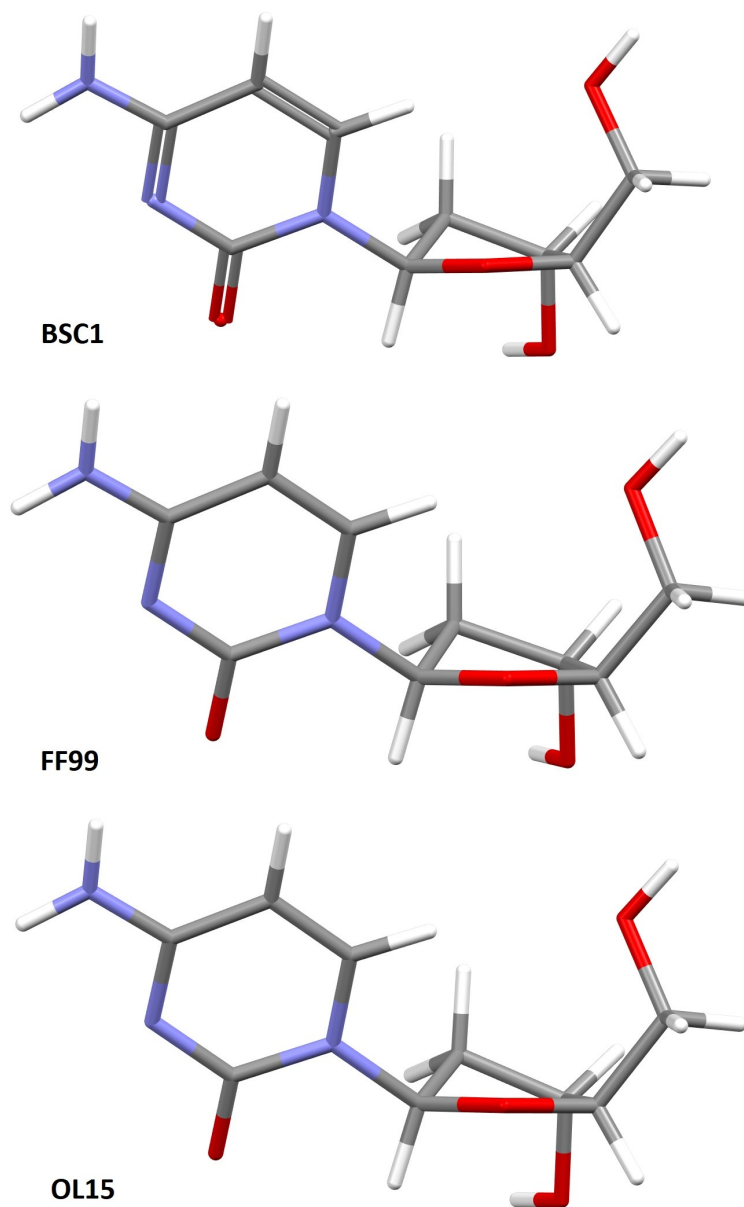


Figura 3.3: Figura comparativa de las estructuras optimizadas de desoxicitidina con energía más baja, obtenidos con los campos de fuerza BSC1, FF99 y OL15.

Para esta pirimidina podemos observar que todos los mínimos locales encontrados tienen una orientación de la base nitrogenada **anti** en los tres campos de fuerza (todos los mínimos con valores del ángulo χ alrededor de los 200°). No obtuvimos mínimos con orientación **syn**.

3.2. Mínimos de energía de desoxitimidina

Para finalizar con las pirimidinas, tenemos para desoxitimidina (dThd) la tabla 3.4 y la Fig. (3.4), que corresponden a los mínimos locales obtenidos con el campo de fuerza BSC1.

Tabla 3.4: Mínimos locales encontrados con el campo BSC1 para desoxitimidina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-1.4061	221.1	178.9	55.3	143.9	170.5	37.7	160.3	-0.60	-0.01	C2'-endo
2	0.8786	210.5	178.9	55.6	86.0	181.9	36.9	16.6	0.04	-0.55	C3'-endo
3	1.1413	209.9	188.8	300.8	150.1	184.9	42.4	159.3	-0.68	-0.03	C2'-endo
4	1.3919	225.6	69.4	86.2	71.1	185.4	36.0	23.7	-0.08	-0.63	C3'-endo
5	4.5894	200.9	169.1	173.8	155.8	164.3	38.0	186.0	-0.22	0.42	C3'-exo
6	4.5952	201.5	175.9	178.3	150.8	166.0	40.4	167.3	-0.54	0.11	C2'-endo
7	7.5926	202.4	178.2	180.0	79.8	167.0	40.4	23.2	-0.08	-0.69	C3'-endo

Analizando la tabla 3.4, podemos agrupar los mínimos locales encontrados con el campo BSC1 con respecto a valores de los ángulos de torsión γ y δ , y con respecto al puckering del anillo de azúcar.

Para este desoxinucleósido predominan nuevamente las conformaciones con puckering C2'-endo y C3'-endo. La única con un puckering diferente es la estructura no.5, pero el valor del ángulo de pseudorrotación es bastante cercano al de las conformaciones C2'-endo, por lo que vemos que tienen características algo similares. En este mínimo la distancia de C2' también es un poco alejada del plano, solo que también tenemos a C3' alejado en dirección contraria.

En la tabla 3.4 vemos los mismos tres grupos respecto al ángulo γ , los mínimos 1 y 2, tienen un valor del ángulo γ en la región **+gauche**. Los mínimos 5, 6 y 7, tienen el ángulo γ en la región **trans**. Mientras que el mínimo no.3, tiene un valor de γ alrededor de la región **-gauche**. Los valores de cada una de estas tres regiones corresponden a conformaciones específicas de dDMP en la cadena del ADN.

Nuevamente tenemos dos regiones predominantes para el ángulo δ , para las estructuras con puckering C2'-endo (y en este caso también para la C3'-exo) tenemos un ángulo δ cercano a los **150 $^{\circ}$** . Mientras que para las estructuras con C3'-endo está alrededor de los **80 $^{\circ}$** .

También podemos observar en la Fig. (3.4) que el mínimo local de BSC1 no.4 tiene una distancia corta entre los átomos HO5' y O4', esta orientación del grupo OH implicado (ángulo $\beta = 69.4^{\circ}$) no puede generarse en el ADN (ya que en lugar de este hidrógeno tenemos el enlace con un grupo fosfato). Y podemos ver que esto también hace que los valores de los ángulos γ y δ no correspondan a los grupos comunes que revisamos en este capítulo.

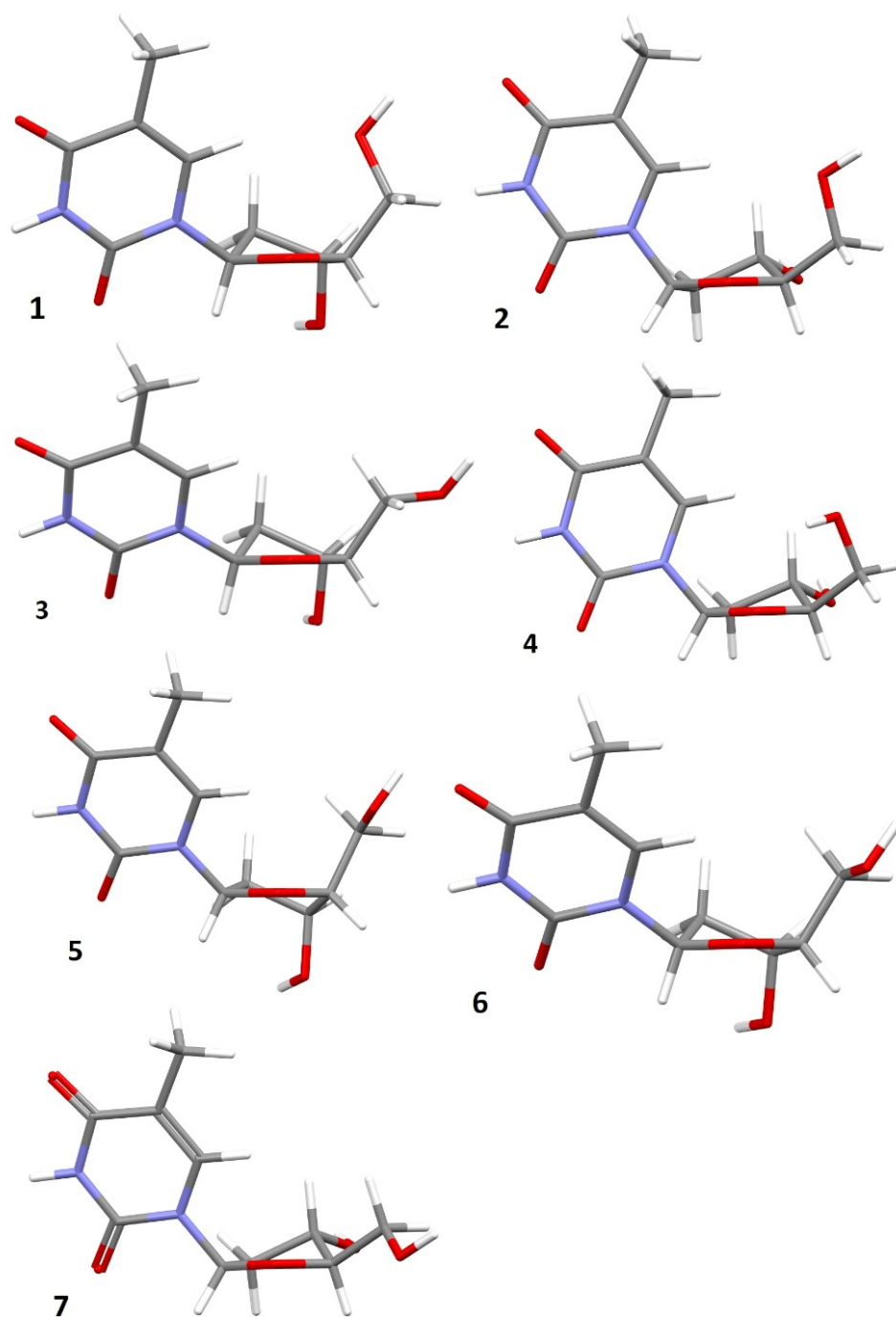


Figura 3.4: Figura comparativa de los mínimos locales encontrados con el campo BSC1 para desoxitimidina. Los datos de cada estructura se muestran en la tabla 3.4.

En las tablas 3.5 y 3.6, tenemos los resultados de los mínimos locales para desoxitimidina obtenidos con los campos de fuerza FF99 y OL15, respectivamente.

Resultados

3.2 Mínimos de energía de desoxitimidina

Tabla 3.5: Mínimos locales encontrados con el campo FF99 para desoxitimidina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-3.0530	220.5	179.0	52.3	139.8	177.6	34.0	155.6	-0.59	-0.08	C2'-endo
2	-1.9951	211.2	179.7	53.7	86.0	184.7	32.6	18.1	0.01	-0.51	C3'-endo
3	-1.5171	213.8	48.6	84.6	172.4	272.3	34.5	18.4	0.01	-0.54	C3'-endo
4	0.6762	198.7	183.0	192.0	133.0	174.9	42.0	137.2	-0.91	-0.45	C1'-exo
5	1.5899	235.4	71.2	91.5	70.2	73.2	27.4	20.7	-0.02	-0.45	C3'-endo

Tabla 3.6: Mínimos locales encontrados con el campo OL15 para desoxitimidina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-3.0706	225.1	177.7	54.5	139.9	178.2	34.1	154.9	-0.60	-0.09	C2'-endo
2	-1.8271	143.8	176.6	55.8	86.9	187.7	32.0	22.1	-0.05	-0.54	C3'-endo
3	-1.3058	218.8	52.2	85.5	168.9	272.2	33.6	22.1	-0.05	-0.57	C3'-endo
4	0.6893	244.9	68.8	69.3	67.1	70.7	24.6	13.3	0.06	-0.34	C3'-endo
5	3.2263	195.9	181.4	181.4	133.7	174.0	41.8	138.8	-0.89	-0.42	C1'-exo

En la Fig. (3.5) podemos observar las características geométricas de los mínimos globales de desoxitimidina para los tres campos de fuerza usados.

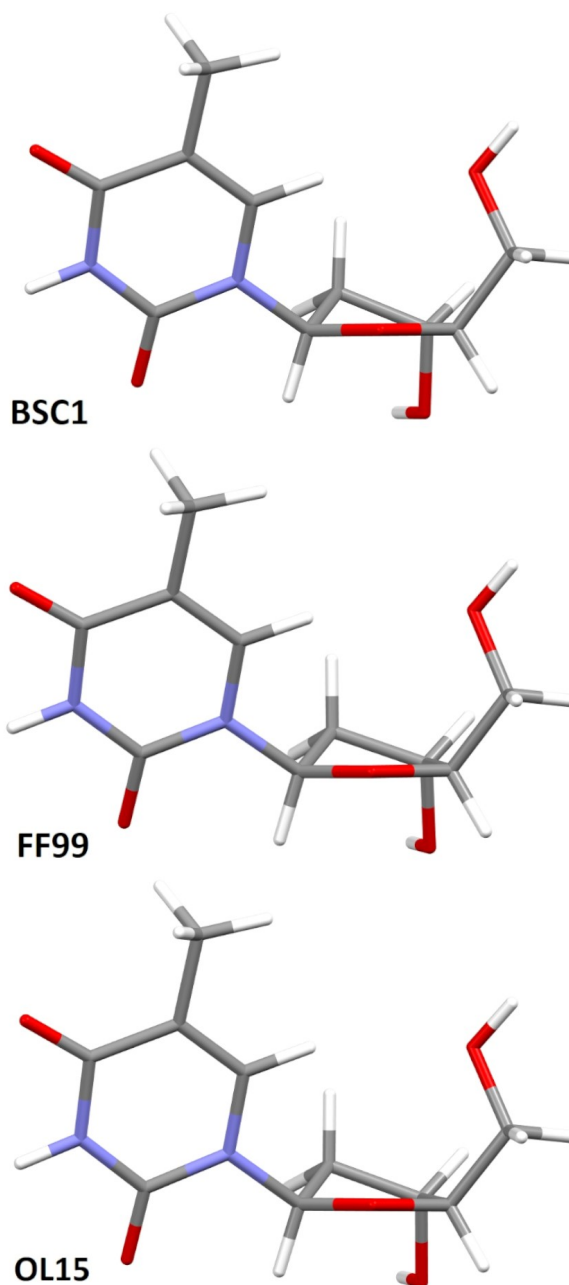


Figura 3.5: Figura comparativa de las estructuras optimizadas de desoxitimidina con energía más baja, obtenidos con los campos de fuerza BSC1, FF99 y OL15.

Nuevamente para esta pirimidina podemos observar que todos los mínimos locales encontrados tienen una orientación de la base nitrogenada **anti** en los tres campos de fuerza (todos los mínimos con valores del ángulo χ alrededor de los 200°). No obtuvimos mínimos con orientación **syn** para ambas pirimidinas.

3.3. Mínimos de energía de desoxiadenosina

La siguiente tabla, 3.7, contiene los resultados para las estructuras optimizadas de desoxiadenosina (dAdo) usando el campo BSC1. Y en la Fig. (3.6) se muestran las estructuras presentadas en la tabla 3.7, las cuales son los mínimos locales encontrados con el campo BSC1.

Tabla 3.7: Mínimos locales encontrados con el campo BSC1 para desoxiadenosina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-41.531	220.0	174.1	54.7	144.6	169.5	36.9	162.5	-0.56	0.03	C2'-endo
2	-40.097	204.9	178.1	55.5	84.4	182.6	37.5	9.3	0.16	-0.46	C3'-endo
3	-39.465	69.7	181.5	299.3	156.5	204.5	39.6	174.4	-0.43	0.24	C2'-endo
4	-36.343	191.9	185.2	182.6	147.3	166.7	41.8	158.8	-0.68	-0.04	C2'-endo
5	-36.198	190.5	176.3	176.4	158.1	162.2	37.2	200.8	0.04	0.61	C3'-exo
6	-34.463	178.4	174.1	174.0	94.5	166.9	39.2	340.3	0.62	0.02	C2'-exo

De la tabla 3.7 vemos que solamente el mínimo local no.3 tiene una orientación **syn** (con χ cercano a 70°), mientras que los demás tienen una orientación **anti**, y esto lo podemos visualizar mejor en la Fig. (3.6).

Aquí tenemos más variedades en cuanto al pucker de las estructuras, predominan las conformaciones C2'-endo (con ángulo de pseudorrotación alrededor de 160°), para estas estructuras tenemos distancias de C2' y C3' al plano bastante parecidas, C2' se encuentra alrededor de 0.5\AA alejado del plano y C3' casi coplanar. También tenemos una conformación C3'-endo (ángulo P = 9.3°). Recordemos que estos dos empaquetamientos del azúcar son los más comunes en fragmentos de ADN, con C2' muy cercana al plano y C3' a casi 0.5\AA del plano. Los últimos dos mínimos tienen puckerings diferentes, sin embargo no están tan alejados de los dos grupos predominantes. Para la conformación C3'-exo tenemos a C2' muy cerca del plano y a C3' muy alejado de éste pero en dirección contraria a la conformación C3'-endo. De forma similar, con C2'-exo tenemos a C3' casi coplanar y C2' bastante alejado en dirección contraria a C2'-endo.

En cuanto a los demás ángulos de torsión, podemos agrupar los mínimos locales encontrados con el campo BSC1 con respecto al valor del ángulo γ .

En la tabla 3.7 vemos dos grupos claros, los mínimos no.1 y no.2 tienen un valor del ángulo γ alrededor de los 55° (+gauche). Mientras que los mínimos no.4, no.5 y no.6, tienen el ángulo γ cercano a 180° (trans). Los valores del ángulo en estas regiones corresponden a conformaciones específicas de dDMP en la cadena del ADN.

Para los valores del ángulo δ tenemos dos regiones, una cercana a los 150° , que corresponde a las conformaciones C2'-endo y C3'-exo (podemos observar esta similitud entre las estructuras 1, 3, 4 y 5 en la Fig. (3.6)). Y otra región con el ángulo δ alrededor de 85° , que corresponde a las conformaciones C3'-endo y C2'-exo (estructuras 2 y 6 de la Fig. (3.6)).

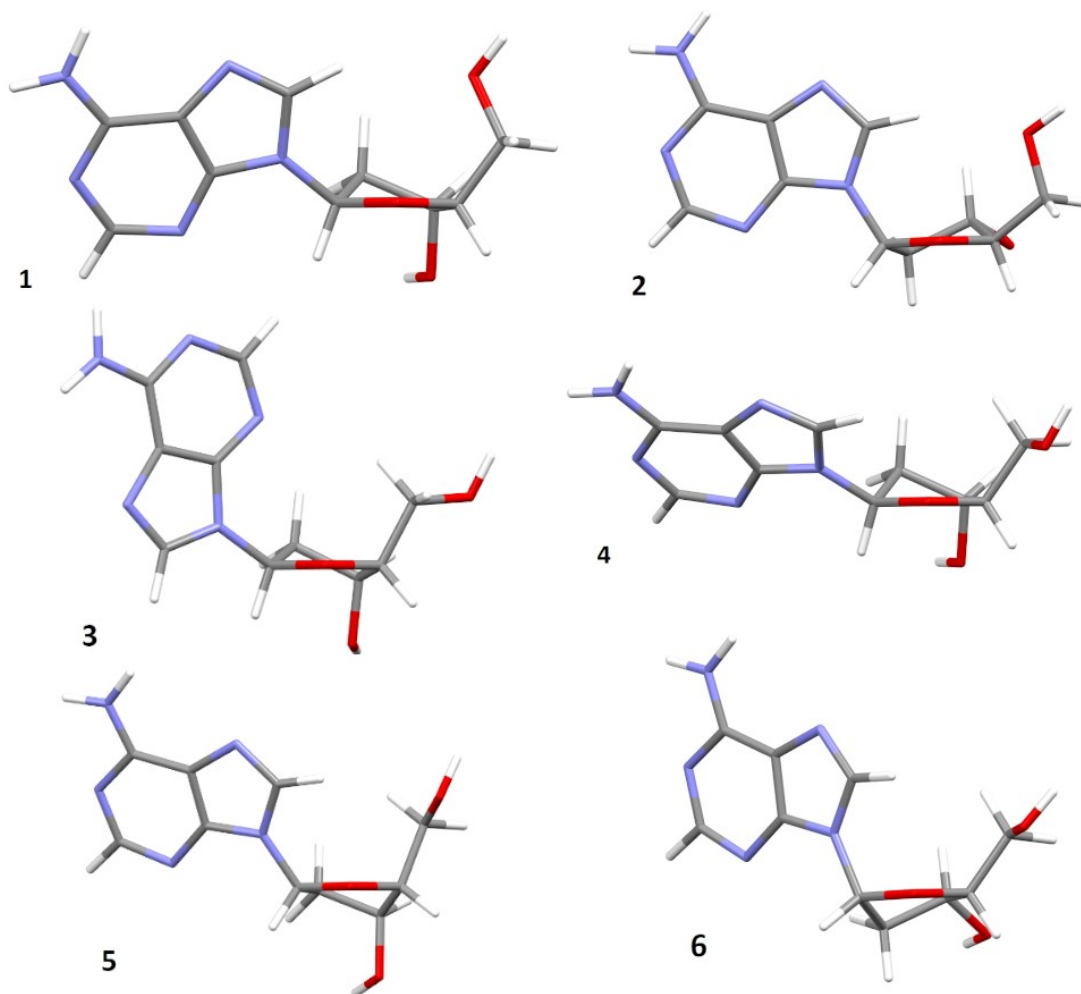


Figura 3.6: Figura comparativa de los mínimos locales encontrados con el campo BSC1 para desoxiadenosina. Los datos de cada estructura se muestran en la tabla 3.7.

A continuación tenemos las tablas 3.8 y 3.9, las cuales contienen los resultados para las estructuras optimizadas de desoxiadenosina (dAdo) usando los campos FF99 y OL15, respectivamente.

Tabla 3.8: Mínimos locales encontrados con el campo FF99 para desoxiadenosina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-44.453	220.6	174.0	50.9	140.2	177.6	33.7	156.4	-0.58	-0.07	C2'-endo
2	-43.847	229.1	182.7	48.4	145.5	58.5	29.8	183.0	-0.21	0.29	C3'-exo
3	-42.213	53.4	292.4	150.1	181.3	284.1	36.7	166.1	-0.51	0.09	C2'-endo
4	-40.620	185.1	182.8	187.2	159.2	297.5	35.5	211.3	0.21	0.69	C3'-exo
5	-40.422	189.5	180.0	190.7	85.7	175.2	36.8	9.7	0.15	-0.46	C3'-endo
6	-39.128	197.2	184.4	195.0	84.1	65.3	35.7	15.2	0.06	-0.52	C3'-endo

Tabla 3.9: Mínimos locales encontrados con el campo OL15 para desoxiadenosina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-43.876	62.6	296.5	151.4	180.6	284.1	36.1	169.7	-0.46	0.14	C2'-endo
2	-43.012	228.7	173.1	53.8	141.6	177.7	32.4	160.8	-0.51	0.00	C2'-endo
3	-42.602	245.3	177.8	51.7	146.4	55.8	30.0	185.3	-0.18	0.32	C3'-exo
4	-42.476	202.1	177.9	55.8	84.7	184.5	34.2	9.3	0.15	-0.42	C3'-endo
5	-39.571	179.3	178.2	176.2	158.3	49.3	37.7	226.0	0.46	0.86	C4'-endo
6	-37.351	174.0	172.8	171.7	104.9	174.4	36.7	323.6	0.76	0.30	C1'-endo
7	-37.119	180.3	182.7	178.2	158.2	297.5	35.6	219.1	0.33	0.76	C4'-endo

En la Fig. (3.7) podemos observar las características geométricas de los mínimos globales de desoxiadenosina para los tres campos de fuerza que usamos, esta figura nos sirve para comparar el mínimo global de cada campo de fuerza. Recordemos que dichos mínimos globales son las estructuras número 1 de las tablas anteriores, es decir, la estructura número 1 de la tabla 3.7, el número 1 de la tabla 3.8 y el de la tabla 3.9.

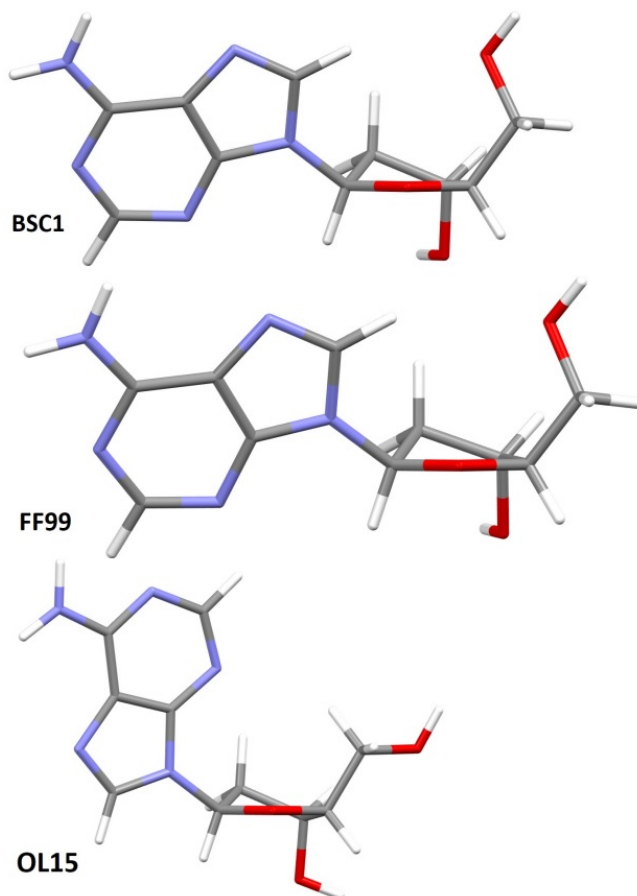


Figura 3.7: Figura comparativa de los mínimos globales de energía para desoxiadenosina, obtenidos con los campos de fuerza BSC1, FF99 y OL15.

Para esta purina podemos observar que predomina la orientación de la base nitrogenada **anti**, con valores de ángulo glucosídico χ cercanos a **200°-210°** en los tres campos de fuerza.

3.4. Mínimos de energía de desoxiguanosina

La tabla 3.10 contiene los resultados para desoxiguanosina (dGuo), usando el campo BSC1. Y la Fig. (3.8) nos permite observar y comparar dichos resultados.

Tabla 3.10: Mínimos locales encontrados con el campo BSC1 para desoxiguanosina.

No.	Energía ($\frac{kcal}{mol}$)	χ (°)	β (°)	γ (°)	δ (°)	ϵ (°)	τ_M	P (°)	C2' (Å)	C3' (Å)	Puckering
1	-88.013	57.3	47.8	55.1	145.5	189.3	36.6	162.4	-0.56	0.02	C2'-endo
2	-81.204	224.2	170.9	53.7	144.6	174.8	36.8	162.7	-0.56	0.03	C2'-endo
3	-79.693	68.3	189.3	300.1	155.1	200.5	39.8	170.9	-0.48	0.18	C2'-endo
4	-79.680	205.1	176.5	55.1	84.5	188.6	37.6	8.6	0.17	-0.45	C3'-endo
5	-75.797	190.8	180.6	177.6	158.4	167.6	37.1	202.5	0.07	0.63	C3'-exo
6	-73.549	78.1	192.1	184.9	80.4	173.7	40.3	36.8	-0.33	-0.83	C4'-exo

Ahora contamos con tres mínimos locales usando BSC1 con conformación **syn**. En la tabla 3.10 están el mínimo no.1, el no.3 y el no.6 (con valores de χ entre 50° y 80°), mientras que los demás tienen una orientación **anti** (con χ alrededor de 200°). Esto lo podemos visualizar mejor en la Fig. (3.8).

De igual forma que con la otra purina, podemos agrupar los mínimos locales encontrados con el campo BSC1 con respecto a valores de los ángulos de torsión γ y δ , y con respecto a su empaquetamiento del azúcar.

Analizando los mínimos locales obtenidos con el campo BSC1, podemos ver que predominan las estructuras con puckering **C2'-endo**, en dichas estructuras el átomo C2' está alejado del plano (con una distancia cercana a los **0.5Å** en dirección a C5'), mientras que C3' está muy cerca del plano. También tenemos un mínimo con puckering C3'-endo como en los casos anteriores, donde C2' está cerca del plano y C3' alejado en dirección a C5' 0.45Å. Por último tenemos dos puckering diferentes más, C3'-exo (con C3' muy alejado del plano y C2' casi coplanar) y C4'-exo, en esta última estructura tenemos tanto a C2' como a C3' alejadas del plano (C3' más alejada) en dirección a C5'.

En la tabla 3.10 vemos tres grupos respecto al ángulo γ , los mínimos 1, 2 y 4, tienen un valor del ángulo γ en la región **+gauche**.

Los mínimos no.5 y no.6, tienen el ángulo γ en la región **trans**.

Mientras que el mínimo no.3 tiene un valor de γ de casi **300°** (-gauche).

Estos diferentes grupo de valores para el ángulo γ corresponden a conformaciones específicas de dDMP en la cadena del ADN.

Para los valores del ángulo δ en estos mínimos locales, tenemos nuevamente dos regiones. Para las estructuras con puckering C2'-endo y C3'-exo tenemos un ángulo δ alrededor de los **150°** (Estructuras 1, 2, 3 y 5 de la Fig. (3.8)). Mientras que para las estructuras con puckering C3'-endo y C4'-exo el valor del ángulo δ está cerca de los **80°** (Estructuras 4 y 6).

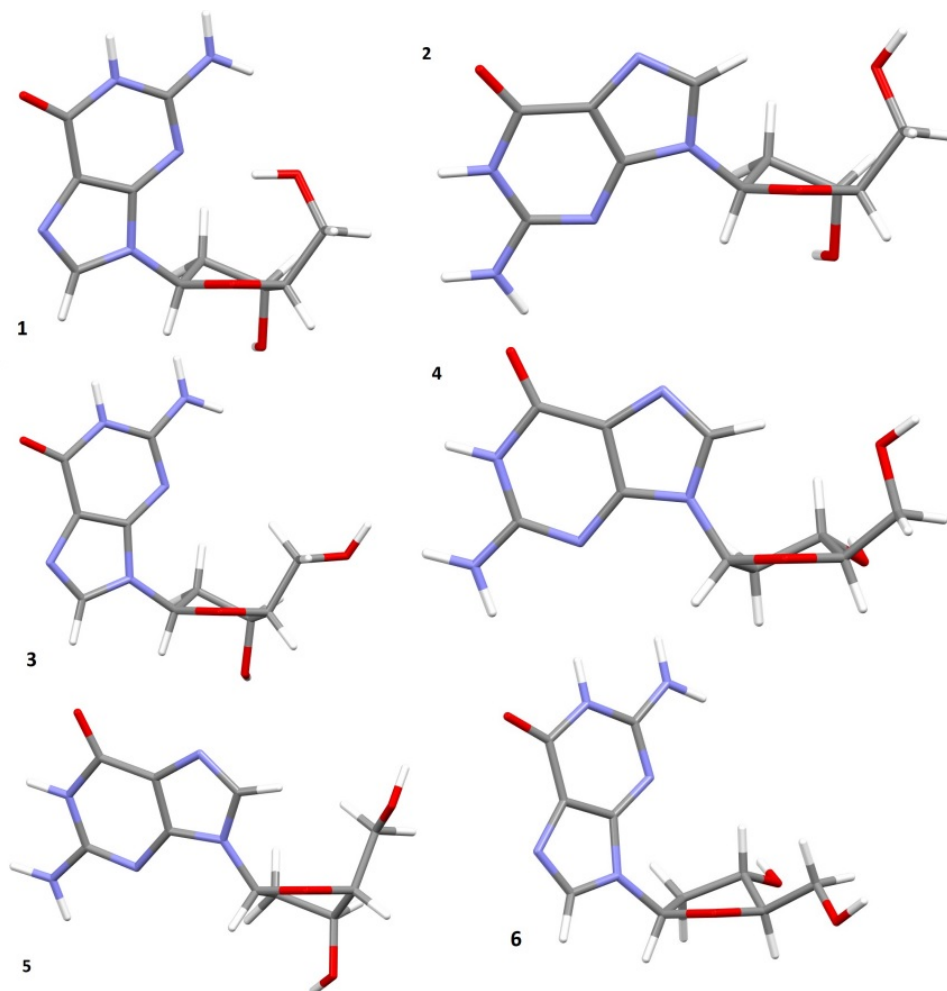


Figura 3.8: Figura comparativa de los mínimos locales encontrados con el campo BSC1 para desoxiguanosina. Los datos de cada estructura se muestran en la tabla 3.10.

En las tablas siguientes, 3.11 y 3.12, tenemos los resultados de los mínimos locales para desoxiguanosina obtenidos con los campos de fuerza FF99 y OL15, respectivamente.

Tabla 3.11: Mínimos locales encontrados con el campo FF99 para desoxiguanosina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-91.668	49.3	48.4	55.7	140.1	190.5	33.8	153.5	-0.61	-0.11	C2'-endo
2	-84.450	226.2	171.7	49.9	141.4	180.6	32.6	160.0	-0.52	-0.01	C2'-endo
3	-83.860	211.3	175.5	52.7	87.3	189.2	32.9	9.0	0.15	-0.40	C3'-endo
4	-81.056	186.1	186.8	188.7	159.3	300.2	35.5	210.4	0.19	0.69	C3'-exo
5	-80.960	53.9	156.9	43.8	128.9	189.5	42.3	134.9	-0.93	-0.48	C1'-exo
6	-80.907	185.4	182.3	184.1	155.8	176.8	33.7	215.3	0.25	0.70	C3'-exo

Tabla 3.12: Mínimos locales encontrados con el campo OL15 para desoxiguanosina.

No.	Energía ($\frac{kcal}{mol}$)	χ ($^{\circ}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	-92.193	54.9	47.1	55.3	141.2	190.3	33.5	155.9	-0.58	-0.08	C2'-endo
2	-83.762	233.9	169.8	52.9	142.3	180.6	31.9	163.0	-0.48	0.03	C2'-endo
3	-83.058	63.0	189.5	299.4	150.6	201.5	37.0	164.7	-0.54	0.06	C2'-endo
4	-83.001	214.2	173.4	55.2	87.4	188.8	32.9	11.2	0.11	-0.43	C3'-endo
5	-78.178	182.8	181.9	177.0	154.6	176.0	33.4	220.4	0.32	0.73	C4'-endo

En la Fig. (3.9) tenemos una comparación geométrica de los mínimos globales (No.1) de desoxiguanosina para los tres campos de fuerza que usamos. Estos mínimos corresponden a una formación con enlace artificial de hidrógeno y no pueden existir en la cadena polinucleotídica.

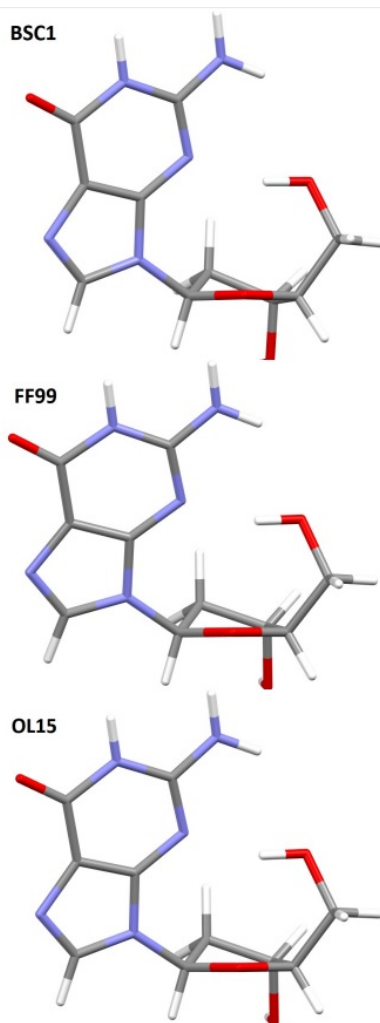


Figura 3.9: Figura comparativa de las estructuras optimizadas de desoxiguanosina con energía más baja, obtenidos con los campos de fuerza BSC1, FF99 y OL15.

Observando los resultados que obtuvimos de este desoxinucleósido, tenemos que existen mínimos locales tanto con la orientación **syn**, como con la orientación **anti** en los tres campos de fuerza.

En todos los desoxinucleósidos predominaron los empaquetamientos del azúcar (o puckering) C2'-endo y C3'-endo. Esta es la forma más común de encontrarlos en el ADN.

3.5. Mínimos de energía de desoxirribosa

Por último veremos los resultados obtenidos para moléculas de desoxirribosa, sin base nitrogenada. Estos resultados no pueden ser comparados con datos experimentales, ya que no existen cristales de molécula de desoxirribosa ni de los fragmentos considerados a continuación.

Las estructuras iniciales fueron cortadas de algunos cristales obtenidos en el banco de datos, y también de algunos desoxinucleósidos con configuración geométrica interesante, buscamos ver cómo se comportaba la desoxirribosa bajo los efectos del campo de fuerza sin interacciones posibles de una base nitrogenada.

A continuación mostramos la tabla 3.13, que contiene los resultados obtenidos con el campo BSC1, seguida de la Fig. (3.10) para visualizar las estructuras de dicha tabla.

Tabla 3.13: Mínimos locales encontrados con el campo BSC1 para desoxirribosa.

No.	Energía ($\frac{kcal}{mol}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	43.334	188.9	184.3	77.6	171.8	42.7	33.3	-0.28	-0.84	C3'-endo
2	42.738	196.2	186.4	134.7	160.6	43.3	140.9	-0.90	-0.39	C1'-exo
3	40.685	188.4	300.8	128.7	190.2	41.7	300.4	0.99	0.69	C1'-endo
4	39.291	164.9	53.9	106.4	175.2	41.9	320.3	0.89	0.42	C1'-endo
5	38.615	183.9	300.6	76.5	315.9	41.9	29.61	-0.28	-0.85	C3'-endo
6	38.066	157.9	50.2	134.3	167.7	41.8	143.4	-0.85	-0.34	C1'-exo

En las tablas 3.14 y 3.15 están detallados los resultados obtenidos para desoxirribosa con los campos FF99 y OL15, respectivamente.

Tabla 3.14: Mínimos locales encontrados con el campo FF99 para desoxirribosa.

No.	Energía ($\frac{kcal}{mol}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	36.688	197.4	197.8	121.1	183.2	43.0	122.3	-1.01	-0.69	C1'-exo
2	36.498	188.4	298.2	134.3	188.4	38.8	286.2	0.94	0.80	O4'-exo
3	35.365	189.4	195.1	96.2	179.5	43.1	91.6	-0.99	-0.98	O4'-endo
4	34.211	165.8	48.7	115.8	179.2	39.2	306.8	0.94	0.61	C1'-endo
5	33.754	161.4	48.2	120.5	181.0	40.7	125.2	-1.00	-0.70	C1'-exo

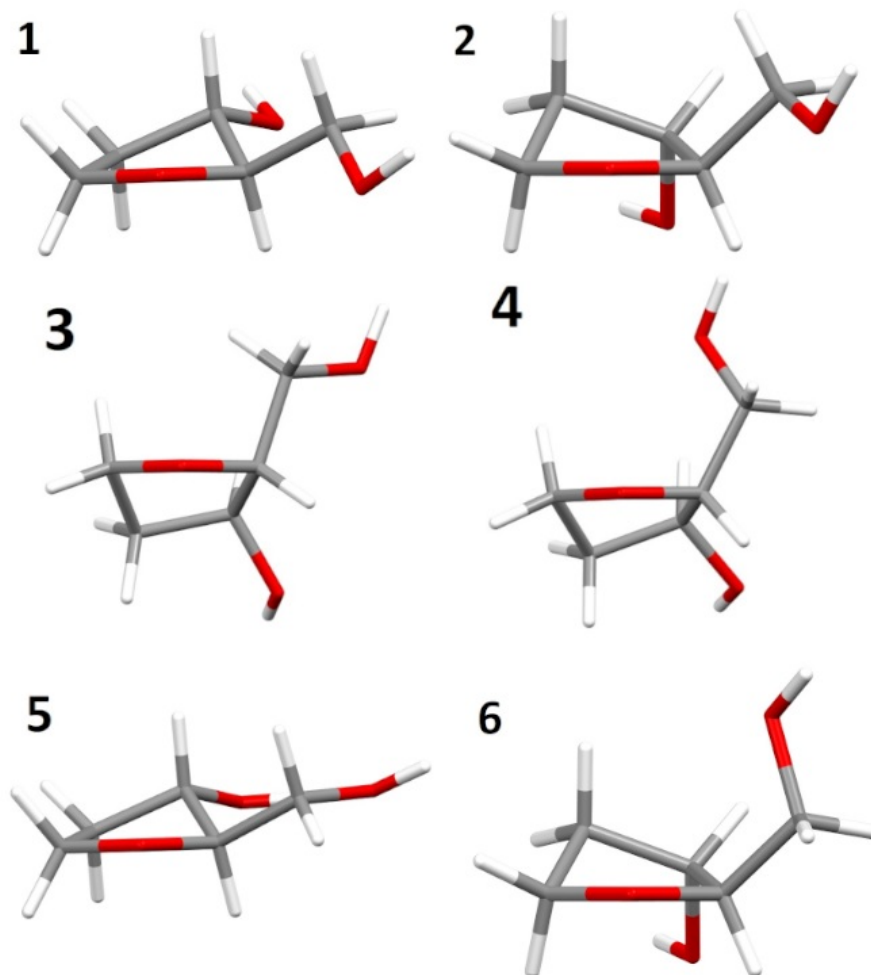


Figura 3.10: Figura comparativa de los mínimos locales encontrados con el campo BSC1 para desoxirribosa. Los datos de cada estructura se muestran en la tabla 3.13 y el número de cada imagen corresponde a su número en la tabla.

Tabla 3.15: Mínimos locales encontrados con el campo OL15 para desoxirribosa.

No.	Energía ($\frac{kcal}{mol}$)	β ($^{\circ}$)	γ ($^{\circ}$)	δ ($^{\circ}$)	ϵ ($^{\circ}$)	τ_M	P ($^{\circ}$)	C2' (\AA)	C3' (\AA)	Puckering
1	39.257	201.9	187.4	119.7	182.7	43.0	121.2	-1.01	-0.07	C1'-exo
2	37.844	190.6	185.1	95.1	179.3	43.1	90.5	-0.99	-0.99	O4'-endo
3	35.111	188.5	299.9	133.9	188.6	38.9	286.7	0.94	0.79	O4'-exo
4	34.251	163.7	52.3	114.2	179.2	39.2	309.2	0.92	0.55	C1'-endo

En la Fig. (3.11) se muestran las estructuras optimizadas de desoxirribosa con energía más baja, los mínimos globales, obtenidos con los tres campos de fuerza, BSC1, FF99 y OL15. Las características de estas tres estructuras son el número 1 de cada tabla para desoxirribosa, es decir las tablas 3.13, 3.14 y 3.15.

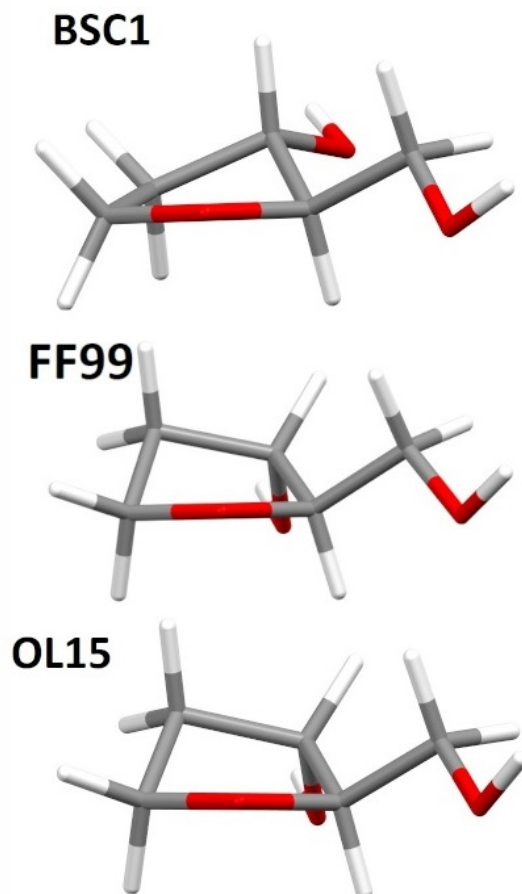


Figura 3.11: Figura comparativa de las estructuras optimizadas de desoxirribosa con energía más baja, obtenidos con los campos de fuerza BSC1, FF99 y OL15.

Podemos observar en las tablas 3.13, 3.14 y 3.15, que con los campos de fuerza AMBER (BSC1, FF99 y OL15) utilizando moléculas de desoxirribosa separadas de las bases nitrogenadas, no se pudo obtener una estructura con puckering C2'-endo. Aún cuando, en la naturaleza, esta configuración para el azúcar del ADN es la más común.

Además, para la molécula de desoxirribosa separada de la base nitrogenada y del esqueleto azúcar-fosfato, tenemos tres nuevos átomos de hidrógeno que no podemos encontrar en el estado natural del ADN (los átomos HO3', HO5' y H1'), ya que en lugar de estos hidrógenos tenemos los enlaces con la base nitrogenada y con los grupos de fosfato que forman el esqueleto. Por lo tanto en nuestra optimización están presentes algunas interacciones que no pueden existir en el ADN debidas a nuevas orientaciones para los dos grupos OH. Y a lo largo de este trabajo encontramos frecuentemente algunos *enlaces artificiales* (el más común fue el enlace formado por HO5'-O4'). Como estas orientaciones de los ángulos β y γ no son interesantes para el presente trabajo, generalmente ajustamos estos ángulos para no tener dichos enlaces artificiales.

Concentraremos ahora el análisis en los mínimos locales obtenidos con el campo de fuerza BSC1 (tabla 3.13), ya que con este campo pudimos obtener de forma más concisa ciertos grupos comparando algunos parámetros.

Primero podemos comparar el puckering y las distancias de C2' y C3' (recordemos que dichas distancias son respecto al plano formado por C1'-O4'-C4', que se mantiene horizontal en la Figura (3.10)).

Tenemos las estructuras no.1 y no.5 con puckering **C3'-endo**, a este puckering corresponde una distancia de C2' cercana al plano (-0.28\AA) y una distancia de C3' mucho más alejada del plano (alrededor de -0.85\AA).

Las estructuras no.3 y no.4 están en otro grupo, con puckering **C1'-endo**, en ellas tanto la distancia de C2' como la de C3' es bastante alejada del plano (hacia abajo en las imágenes), mayor que 0.85\AA para C2' y mayor que 0.40\AA para C3'.

El último grupo de esta comparativa estaría formado por las estructuras no.2 y no.6, con un puckering **C1'-exo**, aquí tenemos el caso contrario al del grupo anterior, con distancias de C2' y C3' alejadas del plano pero en dirección a C5' (alrededor de -0.85\AA para C2' y alrededor de -0.35\AA para C3').

Podemos agrupar estos mínimos locales con respecto a otro parámetro, observemos el ángulo γ . Aquí tenemos nuevamente tres grupos.

El primero formado con las estructuras no.1 y no.2, con valores del ángulo γ cercanos a **180°** (trans).

Después tenemos los mínimos no.3 y no.5, con valores del ángulo γ cercanos a **300°** (-gauche).

Por último tenemos las estructuras no.4 y no.6, que tienen valores del ángulo γ cercanos a **60°** (+gauche).

Otra comparación es respecto al ángulo δ , y podemos ver claramente solo dos grupos. Primero los que tienen un valor del ángulo δ cercano a **77°** (no.1 y no.5), los cuales tienen puckering C3'-endo. Y los que tienen un valor de δ cercano a los **134°** (no.2 y no.6), con puckering C1'-exo.

En la Fig. (3.10) podemos observar que en el mínimo local de BSC1 no.5 existe distancia algo corta entre los átomos HO3' y O5', que corresponde a un enlace artificial de hidrógeno. Esta orientación del grupo OH implicado constituye un mínimo local para la desoxirribosa separada, sin embargo no se puede encontrar en el ADN.

Conclusión

La optimización geométrica de desoxirribonucleósidos y de fragmentos de desoxirribosa separada usando los tres campos de fuerza AMBER (BSC1, FF99 y OL15) mostró algunos mínimos locales para cada subunidad de ADN. Estos mínimos tienen diferentes conjuntos para las tres características conformacionales más importantes de parámetros de los desoxinucleósidos: el empaquetamiento del anillo de azúcar y los ángulos de torsión alrededor del enlace glucosídico y alrededor del enlace C4'-C5'.

Las estructuras de los mínimos globales de energía para ambas pirimidinas obtenidas con los tres campos de fuerza tienen una orientación **anti** de la base con respecto al azúcar, puckering **C2'-endo** y ángulo de torsión alrededor del enlace C4'-C5' en la región **+gauche** (cerca de 60°). Esta conformación corresponde a estructuras de fragmentos mínimos del ADN en las clases conformacionales más frecuentes del Banco de Datos de Ácidos Nucléicos, incluyendo las familias BI y BII.

Las conformaciones con puckering C3'-endo y con regiones de ambos ángulos de torsión parecidas al caso anterior tienen energías menos favorables (de algunas Kcal/mol más). Otros mínimos corresponden a diferentes conjuntos de características conformacionales y a clases conformacionales menos frecuentes de fragmentos mínimos de ADN.

No detectamos mínimos locales con el ángulo glucosídico en la región **syn** para ambas pirimidinas usando los tres campos de fuerza.

Los mínimos de energía para las purinas corresponden a orientaciones **anti** y **syn** de la base con respecto al azúcar. Los mínimos globales de dAdo obtenidos con los campos BSC1 y FF99 corresponden a la región **anti** del ángulo glucosídico, puckering **C2'-endo** y a la región **+gauche** del ángulo alrededor del enlace C4'-C5'. La optimización de dAdo con el campo de fuerza OL15 mostró un mínimo global con ángulo glucosídico **syn** y con el ángulo alrededor del enlace C4'-C5' en la región **trans** (cerca de 180°).

El mínimo con características conformacionales más frecuentes en fragmentos de ADN experimentales tiene energía 0.8 Kcal/mol menos favorable que el mínimo global obtenido por este campo de fuerzas.

Los mínimos de energía para dAdo con orientación **syn** obtenidos usando los campos de fuerza BSC1 y FF99 corresponden a energías menos favorables que las energías de los mínimos globales.

Los mínimos globales para dGuo obtenidos usando los tres campos de fuerza corresponden a una orientación **syn** de la base respecto al anillo de azúcar, puckering **C2'-endo** y a un ángulo de torsión alrededor del enlace C4'-C5' en la región **+gauche**. Estas conformaciones tienen un enlace de hidrógeno entre HO5' y N3, por lo tanto no pueden existir en ácidos nucleicos, pero sí existen en cristales de desoxiguanosina. Las otras conformaciones de dGuo tienen valores de energía menos favorables y corresponden a diferentes clases conformacionales de fragmentos del ADN.

La comparación de las conformaciones de nucleósidos en mínimos de energía obtenidos usando

los tres campos de fuerza con sus estructuras en cristales nos permite concluir que el campo de fuerza **BSC1** da una representación de estas moléculas mejor que los otros dos campos de fuerza AMBER.

A partir de la optimización de moléculas de desoxirribosa en ADN obtenidas de varias conformaciones en cristales de nucleósidos obtuvimos conjuntos de mínimos con diferentes empaquetamientos del anillo de azúcar y diferentes regiones para el ángulo de torsión alrededor del enlace C4'-C5'.

Ningún mínimo obtenido con los tres campos de fuerzas corresponde a la conformación más frecuente en el ADN, en desoxinucleósidos y en desoxinucleótidos (puckering C2'-endo).

No podemos comparar directamente estos resultados con resultados experimentales, ya que no existen datos sobre la estructura de esta molécula.

Apéndice A

Preparación de Estructuras

En éste apéndice explicaré el proceso previo a los cálculos de la Sección (2.3.1) preparando las estructuras, desde la búsqueda y selección de la estructura en el banco de datos, pasando por *recortes* o la construcción que va a tener nuestra estructura, por ejemplo si hay que hacerle algún cambio de base, cambio en coordenadas o en el grupo fosfato, etc.

A.1. Búsqueda y Selección

Para hacer los cálculos correspondientes con una estructura, primero tenemos que tener dicha estructura inicial, y para eso, tenemos que encontrarla en el banco de datos. Para estos pasos usé tres sitios web, el primero es **DNATCO** (<https://dnatco.datmos.org> [32]). Esta página cuenta con muchas funciones, pero yo la usé únicamente por su buscador, el cual permite filtrar por familias las estructuras disponibles, Figura (A.1).



Figura A.1: Interfaz del sitio web DNATCO.

A continuación explicaré cómo hacer la búsqueda de una estructura usando este sitio.

Primero se ingresa a la página, usaremos la pestaña *Browse Conformers*. En el segundo recuadro seleccionable está el filtro de familia de ADN que podemos usar para nuestra búsqueda. Seleccionamos la familia con las características estructurales que nos interesen. Para este ejemplo usaré la familia BB1S. En seguida damos click en el botón *Submit*.

Ahora aparecerá otra interfaz de la página, en la cual nos mostrará del lado izquierdo la simulación gráfica de nuestra estructura, y en el lado derecho una lista de los fragmentos de dicha

Preparación de Estructuras

A.1 Búsqueda y Selección

estructura, fijaremos nuestra atención aquí, Figura (A.2).

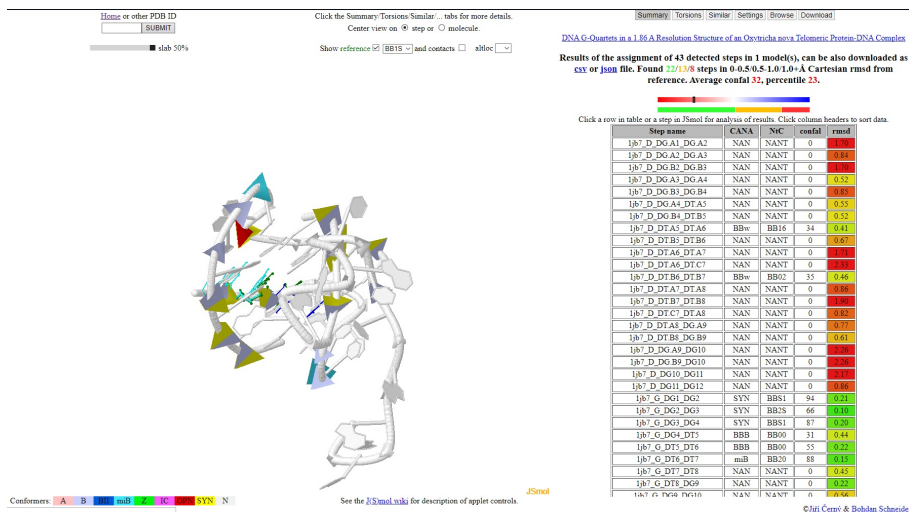


Figura A.2: Estructura en DNATCO.

En la lista, el STEP NAME nos indicará el nombre del fragmento, aquí podemos buscar la secuencia que nos interesa, por ejemplo si queremos preparar una estructura para desoxiguanosina, buscaremos la secuencia dGpdG. Para nuestro ejemplo encontramos una estructura con secuencia dGpdG de la familia BB1S con el número CONFAL grande y RMSD cercano a cero, estos números son medidas que nos dicen qué tan cercano es el fragmento a esa clase.

Al hacer click en el nombre del fragmento en la lista, nos dará una mejor visualización del fragmento en el lado izquierdo de la interfaz, y podemos seleccionar la pestaña Torsions para revisar algunos ángulos de torsión y otros datos del fragmento seleccionado, Figura (A.3).

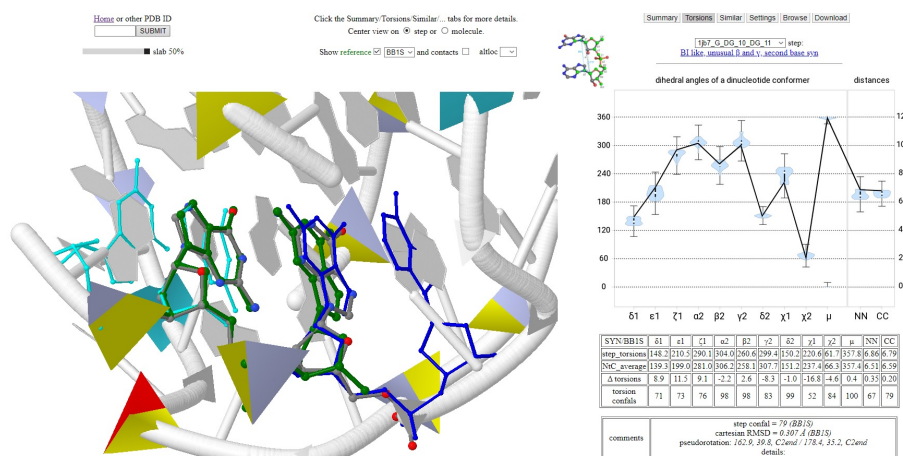


Figura A.3: Fragmento en DNATCO.

Cuando estemos satisfechos con los parámetros de nuestra estructura inicial, revisamos el nombre del fragmento, (en nuestro ejemplo es 1jb7_G_DG_10_DG_11) éste tendrá el siguiente formato: *NombreDeLaEstructura_Hélice_Nucleótido1_#Secuencia1_Nucleótido2_#Secuencia2*.

Este nombre es importante, con él iremos al banco de datos para descargar la estructura en formato .pdb (por sus siglas en inglés, Protein Data Bank) y después encontraremos el fragmento deseado. Ahora usaremos el sitio web **NDB** (Nucleic Acid DataBase, por sus siglas en inglés) (<http://ndbserver.rutgers.edu>), Figura (A.4).

The screenshot shows the NDB website interface. At the top, there is a navigation menu with icons for home and email, and buttons for 'About NDB', 'Standards', 'Education', 'Tools', 'Software', and 'Download'. Below this is the NDB logo and the text 'A Portal for Three-dimensional Structural Information about Nucleic Acids' and 'As of 14-Jul-2021 number of released structures: 11603'. A search bar is located on the right with the text 'Enter an NDB ID or PDB ID' and a search icon. Below the search bar are three buttons: 'Search DNA', 'Search RNA', and 'Advanced Search'. The main content area is divided into two columns. The left column is titled 'Search Structures' and contains three links: 'Search DNA' (Search DNA and its complexes), 'Search RNA' (Search for RNA structures in the NDB archive or in the Non-Redundant list), and 'Advanced Search' (Search for structures based on structural features, chemical features, binding modes, citation and experimental information). The right column is titled 'Featured Tools' and contains five links: 'RNA 3D Motif Atlas', 'Non-redundant Lists of RNA-containing 3D structures', 'RNA Base Triple Atlas', 'WebFR3D', and 'R3D Align'. The footer contains the email 'ndbadmin@ndbserver.rutgers.edu' and the copyright notice '©1995-2021 The Nucleic Acid Database Project | Rutgers, The State University of New Jersey'.

Figura A.4: Sitio web NDB.

Aquí simplemente introduciremos el nombre de nuestra estructura que encontramos en DNATCO en el buscador de NDB. Nos saldrá una página con la descripción y algunos datos de dicha estructura. Esta página solo la usaremos para revisar que sea la estructura correcta y para usarla de puente hacia el último sitio web usado, mediante el enlace resaltado en el nombre de la estructura. Al dar click en dicho enlace, nos redireccionará a una página externa a NDB, **RCSB PDB** (por sus siglas en inglés, Research Collaboratory for Structural Bioinformatics, Protein Data Bank), Figura (A.5).

En ésta página nos darán también datos de nuestra estructura, pero lo importante es que podemos dar click en el recuadro *Download Files* y seleccionar un formato para descargarla. Yo usé únicamente estructuras en formato .pdb, así que para nuestro ejemplo también descargaremos la estructura en dicho formato.

Preparación de Estructuras


A.1 Búsqueda y Selección

RCSB PDB Deposit Search Visualize Analyze Download Learn More Documentation MyPDB

PDB-101 PDB EMDataResource Worldwide Protein Data Bank Foundation Celebrating 40 YEARS OF Protein Data Bank

Structure Summary 3D View Annotations Experiment Sequence Genome Versions

Biological Assembly 1



1JB7

DNA G-Quartets in a 1.86 Å Resolution Structure of an *Oxytricha nova* Telomeric Protein-DNA Complex

DOI: 10.2210/pdb1JB7/pdb NDB: PD0218

Classification: DNA-BINDING PROTEIN/DNA
Organism(s): *Sterkiella nova*
Expression System: *Escherichia coli*
Mutation(s): No

Deposited: 2001-06-02 **Released:** 2001-06-11
Deposition Author(s): Horvath, M.P., Schultz, S.C.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 1.86 Å
R-Value Free: 0.246
R-Value Work: 0.230
R-Value Observed: 0.232

wwPDB Validation

Metric	Percentile Ranks	Value
Rfree		0.255
Clashscore		7
Ramachandran outliers		0.6%
Sidechain outliers		2.8%
RSRZ outliers		8.0%

This is version 1.3 of the entry. See complete history.

Figura A.5: Sitio web RCSB PDB.

A.2. Recorte del Fragmento

Ya con nuestra estructura descargada en formato .pdb, la abriremos en cualquier cualquier software procesador de texto.

Ahora tenemos que *cortar* el fragmento que vamos a usar, recordemos que en la búsqueda de nuestro fragmento encontramos en el nombre la secuencia a la que pertenecía el nucleótido en cuestión, por ejemplo en nuestro caso dicha secuencia para nuestro dinucleótido dGpdG era la dG_10_dG_11, como se ve en la Figura (A.6).

ATOM	625	N3	DG	G	9	18.112	42.226	-10.782	1.00	38.03	N
ATOM	626	C4	DG	G	9	17.012	41.451	-10.928	1.00	41.11	C
ATOM	627	P	DG	G	10	16.606	43.856	-16.479	1.00	52.22	P
ATOM	628	OP1	DG	G	10	16.701	43.842	-17.963	1.00	53.93	O
ATOM	629	OP2	DG	G	10	16.532	45.154	-15.760	1.00	51.60	O
ATOM	630	O5'	DG	G	10	15.380	42.925	-16.069	1.00	50.55	O
ATOM	631	C5'	DG	G	10	14.041	43.399	-16.125	1.00	49.73	C
ATOM	632	C4'	DG	G	10	13.109	42.391	-15.494	1.00	49.02	C
ATOM	633	O4'	DG	G	10	13.541	42.087	-14.143	1.00	47.88	O
ATOM	634	C3'	DG	G	10	11.659	42.860	-15.366	1.00	48.60	C
ATOM	635	O3'	DG	G	10	10.804	41.721	-15.488	1.00	49.78	O
ATOM	636	C2'	DG	G	10	11.600	43.383	-13.943	1.00	47.64	C
ATOM	637	C1'	DG	G	10	12.477	42.365	-13.242	1.00	45.76	C
ATOM	638	N9	DG	G	10	13.051	42.786	-11.968	1.00	43.56	N
ATOM	639	C8	DG	G	10	14.149	43.586	-11.759	1.00	41.85	C
ATOM	640	N7	DG	G	10	14.438	43.735	-10.494	1.00	40.94	N
ATOM	641	C5	DG	G	10	13.465	42.996	-9.832	1.00	40.98	C
ATOM	642	C6	DG	G	10	13.268	42.766	-8.444	1.00	38.48	C
ATOM	643	O6	DG	G	10	13.940	43.176	-7.500	1.00	36.48	O
ATOM	644	N1	DG	G	10	12.156	41.962	-8.211	1.00	37.66	N
ATOM	645	C2	DG	G	10	11.340	41.441	-9.187	1.00	39.66	C
ATOM	646	N2	DG	G	10	10.311	40.699	-8.772	1.00	38.14	N
ATOM	647	N3	DG	G	10	11.522	41.639	-10.483	1.00	41.05	N
ATOM	648	C4	DG	G	10	12.595	42.419	-10.729	1.00	41.22	C
ATOM	649	P	DG	G	11	9.336	41.882	-16.115	1.00	52.75	P
ATOM	650	OP1	DG	G	11	8.758	40.515	-16.192	1.00	52.58	O

Figura A.6: Recorte en el Archivo.pdb con el procesador de texto.

Entonces buscaremos en nuestro archivo .pdb dicha secuencia, una vez encontrada simplemente hacemos el recorte, en este ejemplo solo usaremos el nucleótido 10 desoxiguanosina.

De esta forma tenemos las coordenadas de los átomos que forman el fragmento de interés y podemos trabajar con ese archivo .pdb inicial.

Apéndice B

Cambio de Base de un Nucleósido

En este apéndice se detallará el procedimiento para cambiar la base nitrogenada de nuestro fragmento de desoxirribosa que queremos optimizar. Este cambio de base fue muy usado en el presente trabajo, y nos sirvió para explorar todas las opciones que nos brindaba una configuración específica de las coordenadas del azúcar en un pequeño fragmento de ADN.

Al realizar el cambio de base tenemos que cuidar que la posición de la base respecto al azúcar sea exáctamente igual en la estructura inicial y en la resultante. Recordemos que la geometría de las bases nitrogenadas es planar, por lo tanto basta con mantener las coordenadas del anillo de la base, de esta forma el ángulo glucosídico χ , además de los ángulos de torsión de la desoxirribosa, se mantienen sin cambios.

Primero veamos las características del formato .pdb con un procesador de texto, Figura (B.1).

Este cambio solo lo haremos entre bases con el mismo anillo, es decir, Adenina a Guanina, Citosina a Timina, y viceversa.

Para realizar el cambio seguiremos las siguientes instrucciones:

- Abrimos el archivo de coordenadas de nuestra estructura inicial en formato .pdb con un procesador de texto y borramos los átomos de la base nitrogenada que no correspondan a los anillos. Para purinas dejamos los carbonos 2, 4, 5, 8 y los nitrógenos 1, 3, 7, 9. Para pirimidinas dejamos los carbonos 2, 4, 5, 6 y los nitrógenos 1, 3. *Por ejemplo borramos los átomos de color rojo en la Figura (B.1).*
- En el mismo archivo cambiamos la descripción de la molécula que forma cada átomo de la lista, usamos desoxinucleósidos con notación **DTN** (para desoxitimidina), **DCN** (desoxicitidina), **DAN** (desoxiadenosina) y **DGN** (desoxiguanosina). *Por ejemplo, cambiamos la columna DAN por DGN en el texto del archivo.*
- Guardamos el nuevo formato como un archivo inicial .pdb con la base nueva.
- Antes de optimizar la molécula, ejecutamos XLeap con el campo de fuerza *ff99*, Sección (2.3.1). El programa completará automáticamente los átomos faltantes (los que se encuentran fuera del anillo de la base) según el tipo de nucleósido lea en la columna 4.
- Revisamos con Unit Editor que nuestra nueva estructura haya respetado la posición del anillo de la base y las coordenadas de los átomos del azúcar.
- Guardamos el archivo .pdb completado con la instrucción *savepdb* de XLeap.

Cambio de Base de un Nucleósido

```

ATOM  1  H5T  DAN   1   10.853  35.206 -11.909  1.00  0.00
ATOM  2  O5'  DAN   1   10.399  35.319 -11.026  1.00  0.00
ATOM  3  C5'  DAN   1   10.993  34.446 -10.071  1.00  0.00
ATOM  4  H5'1  DAN   1   11.918  34.942  -9.777  1.00  0.00
ATOM  5  H5'2  DAN   1   11.227  33.488 -10.535  1.00  0.00
ATOM  6  C4'  DAN   1   10.143  34.208  -8.843  1.00  0.00
ATOM  7  H4'  DAN   1   10.036  33.129  -8.731  1.00  0.00
ATOM  8  O4'  DAN   1   10.846  34.630  -7.645  1.00  0.00
ATOM  9  C1'  DAN   1   10.046  35.559  -6.918  1.00  0.00
ATOM 10  H1'  DAN   1    9.574  35.048  -6.079  1.00  0.00
ATOM 11  N9   DAN   1   10.940  36.521  -6.279  1.00  0.00
ATOM 12  C8   DAN   1   11.084  36.739  -4.926  1.00  0.00
ATOM 13  H8   DAN   1   10.492  36.187  -4.211  1.00  0.00
ATOM 14  N7   DAN   1   11.977  37.649  -4.644  1.00  0.00
ATOM 15  C5   DAN   1   12.452  38.061  -5.883  1.00  0.00
ATOM 16  C6   DAN   1   13.450  39.017  -6.215  1.00  0.00
ATOM 17  N6   DAN   1   14.207  39.774  -5.409  1.00  0.00
ATOM 18  H61  DAN   1   14.881  40.414  -5.805  1.00  0.00
ATOM 19  H62  DAN   1   14.105  39.704  -4.407  1.00  0.00
ATOM 20  N1   DAN   1   13.608  39.129  -7.595  1.00  0.00
ATOM 21  C2   DAN   1   12.907  38.410  -8.531  1.00  0.00
ATOM 22  H2   DAN   1   13.025  38.501  -9.601  1.00  0.00
ATOM 23  N3   DAN   1   11.989  37.506  -8.234  1.00  0.00
ATOM 24  C4   DAN   1   11.811  37.385  -6.902  1.00  0.00
ATOM 25  C3'  DAN   1    8.795  34.933  -8.818  1.00  0.00
ATOM 26  H3'  DAN   1    8.426  35.276  -9.785  1.00  0.00
ATOM 27  C2'  DAN   1    9.058  36.132  -7.922  1.00  0.00
ATOM 28  H2'1  DAN   1    8.998  37.047  -8.512  1.00  0.00
ATOM 29  H2'2  DAN   1    8.312  36.163  -7.128  1.00  0.00
ATOM 30  O3'  DAN   1    7.824  34.057  -8.240  1.00  0.00
ATOM 31  H3T  DAN   1    6.937  34.420  -8.178  1.00  0.00
TER
END

```

Figura B.1: Archivo .pdb de un desoxinucleósido.

En la *primera columna* tenemos el tipo de átomo; la *segunda columna* muestra el número del átomo en la estructura; la *tercera columna* el nombre del átomo de acuerdo a su posición; en la *cuarta columna* vemos el tipo de desoxinucleósido en el que está cada átomo; y en las *columnas 6, 7 y 8* tenemos las coordenadas del átomo. El color verde indica los átomos del azúcar, el color azul indica los átomos del anillo de la base nitrogenada, mientras que el color rojo indica los átomos de la base nitrogenada que están fuera del anillo.

Ahora tendremos una nueva estructura inicial con la misma geometría que la anterior pero con una base nitrogenada diferente, y podemos hacer las optimizaciones correspondientes.

Apéndice C

Restricción de Ángulos de Torsión

Muchas veces nos encontramos con casos en los que la optimización geométrica de ciertas estructuras, Sección (2.3), resulta en mínimos locales con enlaces de hidrógeno artificiales, estos mínimos energéticos no pueden ser encontrados en la naturaleza, por lo que se desvían del objetivo de este trabajo. La restricción geométrica de uno o varios ángulos de torsión de la molécula nos ayuda a que dichos enlaces artificiales no presenten problemas al momento de optimizar la estructura con el software.

Para realizar la optimización con restricción necesitamos un archivo llamado **RST**, Apéndice (C.0.2), el cual podremos editar usando cualquier programa de procesamiento de texto.

Este archivo tendrá escritas las siguientes líneas:

```
&rst      iat =  A,  B,  C,  D,
          r1 = N-1, r2 = N, r3 = M, r4 = M+1
          rk2 = x, rk3 = y,                /
```

En la primera línea *A*, *B*, *C*, *D*, son los números de los átomos implicados en el ángulo según el orden en el archivo .pdb de la estructura a optimizar, debemos modificar cada número de los átomos que formen el ángulo diedro (ángulo entre los dos planos) que queremos restringir.

N y *M* son los ángulos entre los cuales restringiremos la torsión, es decir, el ángulo de torsión α estará restringido al intervalo $N \geq \alpha \geq M$.

En la última línea tenemos las constantes de fuerza, *rk*, en el presente trabajo usamos valores de $rk2 = rk3 = 50.0$ ya que en el manual de AMBER [33] se utiliza este valor para la constante de fuerza, Figura (C.1).

Lo que se restringe realmente es la parte de la energía potencial que modela los ángulos de torsión, por lo tanto necesitamos los valores *r1*, *r2*, *r3* y *r4*, que nos indicarán la posición de las barreras de potencial. Al definir dichas barreras de potencial, debemos escoger las constantes de fuerza tales que nuestro sistema se encuentre alrededor de dichas barreras.

Una vez tengamos los parámetros escritos en el archivo RST debemos crear otro archivo .in que llame a RST al momento de optimizar. A nuestro archivo 999.in le añadimos tres líneas que ejecutará a RST en la instrucción **SANDER**, Sección (2.3).

```
&wt type='END' /
LISTOUT=POUT
DISANG=RST
```

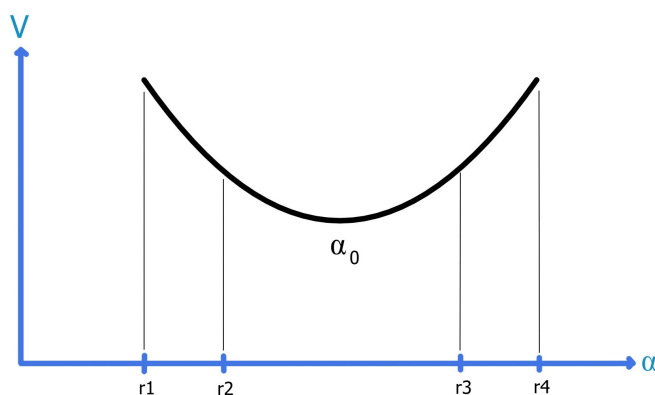


Figura C.1: Potencial del ángulo de torsión con restricción.

Se crean barreras de potencial usando las constantes de fuerza r_1 , r_2 , r_3 y r_4 . Dependiendo de estas constantes el ángulo de torsión podrá salir o no de dichas barreras.

Entonces al momento de ejecutar SANDER el nuevo archivo `.in` (`999-rst.in`) aplicará las instrucciones detalladas en RST para restringir el ángulo.

Debemos tener cuidado con la energía de las barreras de potencial, ya que muchas veces el valor de las constantes rk no es suficiente para forzar la restricción y debemos aumentarlo, pero si la contribución de energía se ve demasiado afectada puede ser una *falsa* optimización geométrica de la molécula.

C.0.1. Cambiar Ángulos de Torsión Antes de Optimizar

Si aún restringiendo no obtenemos una optimización geométrica sin enlaces artificiales o si para hacerlo debemos usar valores de rk muy altos, podemos optar por una modificación en el ángulo de torsión donde se genera el enlace artificial.

Esta modificación la haremos a la estructura inicial en formato `.pdb` usando el software **UCSF Chimera**¹ [34], que entre muchas otras cosas, nos ofrece herramientas para la visualización y edición de estructuras moleculares en formato `.pdb`.

Después de abrir nuestro archivo `.pdb` con este software, usaremos la herramienta `Tools>Structure Analysis>Angles/Torsions`. Se abrirá la ventana `Structure Measurements`, en la pestaña `Adjust Torsions` podemos acceder a la herramienta `Build Structure`.

En el visualizador de la molécula seleccionamos el enlace sobre el cual modificaremos el ángulo de torsión. Después en la ventana de la herramienta `Build Structure` activamos el enlace seleccionado como el ángulo de torsión, Figura (C.2). Así el programa nos mostrará los átomos implicados en el ángulo de torsión y podremos ir ajustando dicho ángulo a nuestra conveniencia.

Cuando tengamos el ángulo de torsión ajustado al valor inicial que deseemos, guardamos la estructura nueva en formato `.pdb` y seguimos a hacer la optimización de forma normal tomándola como estructura inicial, Sección (2.3.1).

¹Es un software diseñado para visualización interactiva y análisis de estructuras moleculares y ensamblajes conformacionales. Desarrollado por la RBVI (*Resource for Biocomputing, Visualization, and Informatics* en la Universidad de California, San Francisco, 2019) [34].

Este método también fue utilizado en el presente trabajo para explorar las posibilidades estructurales de algunas moléculas, así como también para evitar los contactos cortos artificiales entre átomos.

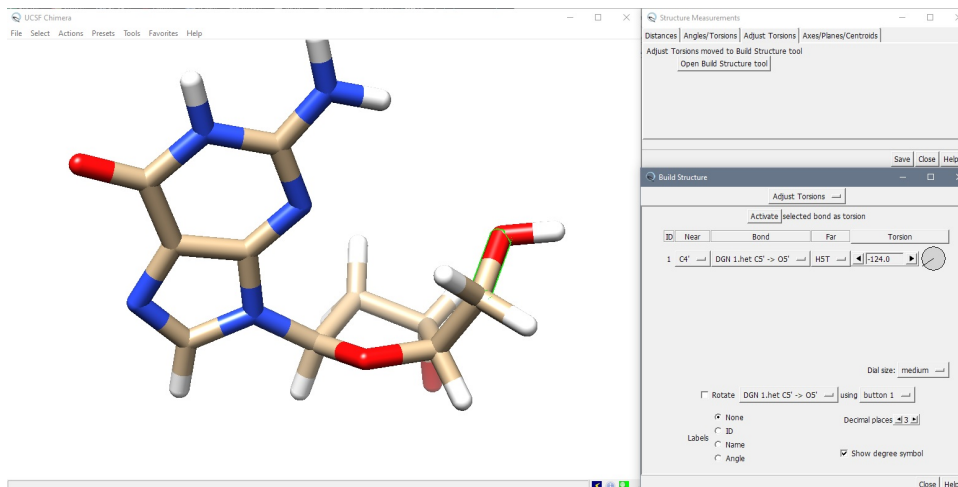


Figura C.2: Interfaz del software Chimera.

Del lado izquierdo podemos observar la estructura de la molécula y tenemos diferentes herramientas para su edición o visualización; del lado derecho tenemos la herramienta *Build Structure* con la que podemos ajustar ángulos de torsión.

C.0.2. Crear Archivo RTS

- Creamos un archivo de texto de 5 columnas.

```
res# res_name torsion_name lower_bound upper_bound
(ejemplo) 1 DT BETA 170.0 190.0
```

Lo llamamos (archivo.bb).

- Convertimos el archivo .bb en un formato AMBER, con las siguientes líneas en la terminal.

```
makeANG_RST -pdb pdb_file -con constraint -lib library_file [-les lesfile]
makeANG_RST -pdb 1dou-tcNa.pdb -con "archivo".bb -lib
$AMBERHOME/AmberTools/src/nmr_aux/prepare_imput/tordef.lib > RST.bb
cat RST.bb > RST
```


Bibliografía

- [1] Watson, J. D. *The double helix; a personal account of the discovery of the structure of DNA*. New York, Atheneum (1968).
- [2] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. y Case, D. A. “Development and testing of a general amber force field”. *Journal of Computational Chemistry*, 26: 1157–1174 (2004).
- [3] Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBolt, S., Ferguson, D., Seibel, G. y Kollman, P. “AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules”. *Computer Physics Communications*, 91: 1–41 (1995).
- [4] Poltev, V., Anisimov, V. M., Danilov, V. I., García, D., Sánchez, C., Deriabina, A., Gonzáles, E., Rivas, F. y Polteva, N. “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) (2014).
- [5] Poltev, V., González, E., Mourik, T. V., Deriabina, A., M.Padua, García, D., Rivas, F., Polteva, N., Anisimov, V. M. y Danilov, V. I. “DFT study of polymorphism of the DNA double helix at the level of dinucleosid monophosphates.” *Wiley InterScience*, 110 (13): 2548–2559 (2010).
- [6] Watson, J. D. y Crick, F. H. C. “Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid”. *Nature*, 171: 737–738 (1953).
- [7] Miescher, F. “Über die chemische Zusammensetzung der Eiterzellen”. *Medicinsich-chemische Untersuchungen*, 4: 441–460 (1871).
- [8] Pauling, L. y Niemann, C. “The Structure of Proteins”. *Journal of the American Chemical Society*, págs. 1860–1867 (1939).
- [9] Ruth, H. *The politics of women’s biology*. New Brunswick, [N.J.]: Rutgers University Press (1990).
- [10] Elson, D. y Chargaff, E. “On the deoxyribonucleic acid content of sea urchin gametes”. *Experientia*, 8 (4): 143–145 (1952).

- [11] Franklin, R. y Gosling, R. G. "Molecular Configuration in Sodium Thymonucleate". *Nature*, 171: 740–741 (1953).
- [12] Nobel Media AB 2020. "The Nobel Prize in Physiology or Medicine". *NobelPrize.org* (1962).
- [13] Poltev, V. "Molecular Mechanics: Principles, History, and Current Status". *Handbook of Computational Chemistry*, págs. 21–67 (2017).
- [14] Hendrickson, J. B. "Molecular geometry. I. Machine computation of the common rings". *Journal of the American Chemical Society*, 83 (22): 4537–4547 (1961).
- [15] Hendrickson, J. B. "Molecular geometry. II. Methyl-cyclohexanes and cycloheptanes." *Journal of the American Chemical Society*, 84: 3355–3359 (1962).
- [16] Hendrickson, J. B., Boeckman, R. K., Glickson, J. D. y Grunwald, E. "Molecular geometry. VIII. Proton magnetic resonance studies of cycloheptane conformations". *Journal of the American Chemical Society*, 95: 494–505 (1973).
- [17] Allinger, N. L., Tribble, M. T., Miller, M. A. y Wertz, D. W. "Conformational analysis. LXIX. Improved force field for the calculation of the structures and energies of hydrocarbons." *Journal of the American Chemical Society*, 93: 1637 (1971).
- [18] Allinger, N. L. y Sprage, J. T. "Calculation of the structures of hydrocarbons containing delocalized electronic systems by the molecular mechanics method." *Journal of the American Chemical Society*, 95: 3893 (1973).
- [19] Engler, E. M., Andose, J. D. y Schleyer, P. V. R. "Critical evaluation of molecular mechanics". *Journal of the American Chemical Society*, 95 (24): 8005–8025 (1973).
- [20] Weiner, P. y Kollman, P. "AMBER: Assisted model building with energy refinement. a general program for modeling molecules and their interactions". *Journal of Computational Chemistry*, 2: 287 (1981).
- [21] Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Jr., S. P. y Weiner, P. "A new force field for molecular mechanical simulation of nucleic acids and proteins". *Journal of the American Chemical Society*, 106: 765 (1984).
- [22] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. y Kollman, P. A. "A second generation force field for the simulation of proteins, nucleic acids and organic molecules". *Journal of the American Chemical Society*, 117 (19): 5179–5197 (1995).
- [23] Alberts, B., Jhonson, A., Lewis, J., Raff, M., Roberts, K. y Walter, P. *Molecular Biology of the Cell*. Garland Science (2002).
- [24] Sinden, R. *DNA Structure and Function*. Academic Press, 1^a ed^{ón}. (1994).
- [25] Lagerkvist, U. *DNA pioneers and their legacy*. New Haven, CT: Yale University Press (1998).

- [26] Altona, C. y Sundaralingam, M. “Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation”. *Journal of the American Chemistry Society*, 94 (23): 8205–8212 (1972).
- [27] Saenger, W. *Principles of Nucleic Acid Structure*. Springer-Verlag New York, 1^a ed^{ón}. (1984).
- [28] Teller, E. y Sahlin, H. *Physical Chemistry An Advanced Treatise*. Academic Press, New York (1971).
- [29] Dzyaloshinskii, I. E., Lifshitz, E. M. y Pitaevskii, L. P. “General theory of van der Waals forces”. *Soviet Physics Uspekhi*, 4 (2): 153–176 (1961).
- [30] Ponder, J. W. y Case, D. A. “Force fields for proteins simulations”. *Advances in Protein Chemistry*, págs. 27–85 (2003).
- [31] Macrae, C. F., Edgington, P. R., McCabe, P., Pidcock, E., Shields, G. P., Taylor, R., Towler, M. y van de Streek, J. “Mercury: visualization and analysis of crystal structures”. *Journal of Applied Crystallography*, 39: 453–457 (2006).
- [32] Cerny, J., Bozikova, P., Svoboda, J. y Schneider, B. “A unified dinucleotid alphabet describing both RNA and DNA strustures”. *Nucleic Acids Research*, 48 (11): 6367–6381 (2020).
- [33] Case, D. A., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Giambasu, G., Gilson, M. K., Gohlke, H., Goetz, A. W., Harris, R., Izadi, S., Izmailov, S. A., Kasavajhala, K., Kovalenko, A., Krasny, R., Kurtzman, T., Lee, T. S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Man, V., Merz, K. M., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., Onufriev, A., Pan, F., Pantano, S., Qio, R., Roe, D. R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C. L., Skrynnikov, N. R., Smith, J., Swails, J., Walker, R. C., Wang, J., Wilson, L., Wolf, R. M., Wu, X., Xiong, Y., Xue, Y., York, D. M. y Kollman, P. A. *Amber 2020*. University of California, San Francisco (2020).
- [34] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. y Ferrin, T. E. “UCSF chimera: A visualization system for exploratory research and analysis”. *Journal of Computational Chemistry*, 25 (13): 1605–1612 (2004).