



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

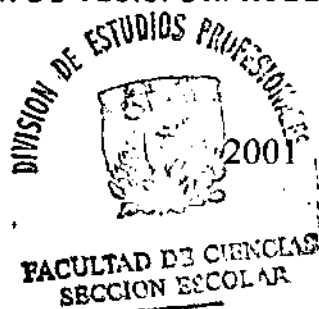
FACULTAD DE CIENCIAS

INVESTIGACIÓN DE LA CATÁLISIS DE LA  
RIBOZIMA "HAMMERHEAD" EN PRESENCIA  
DE IONES DE MAGNESIO MEDIANTE EL  
MÉTODO DE DINÁMICA MOLECULAR

292826

T E S I S  
QUE PARA OBTENER EL TÍTULO DE  
F Í S I C A  
P R E S E N T A :  
ANA PATRICIA L GÁMIZ HERNÁNDEZ

DIRECTOR DE TESIS: DR. RUBÉN SANTAMARÍA ORTÍZ





Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**MAT. MARGARITA ELVIRA CHÁVEZ CANO**  
Jefa de la División de Estudios Profesionales de la  
Facultad de Ciencias  
Presente


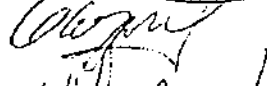

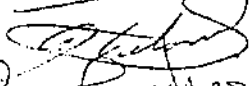

Comunicamos a usted que hemos revisado el trabajo de Tesis: INVESTIGACION DE LA CATALISIS DE LA RIBOZIMA "HAMMERHEAD" EN PRESENCIA DE IONES DE MAGNESIO MEDIANTE EL METODO DE DINAMICA MOLECULAR

realizado por ANA PATRICIA L. GAMIZ HERNANDEZ

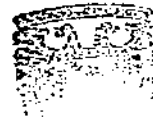
con número de cuenta 9450342-7, pasante de la carrera de FÍSICA.

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis	DR. RUBEN SANTAMARIA ORTIZ	
Propietario	DR. IGNACIO LUIS GARZON SOSA	
Propietario	DRA. MARCELA BELTRAN SANCHEZ	
Suplente	DRA. ANGELICA LEONOR GELOVER SANTIAGO	
Suplente	DR. GERMINAL COCHO GIL	

Consejo Departamental de Física



  
DRA. PATRICIA GOLDSTEIN MENACHE FACULTAD DE CIENCIAS  
DEPARTAMENTO DE FÍSICA

*Agradecimientos:*

*A mi padre que con su cariño, dedicación y confianza siempre me ha motivado a seguir adelante.*

*A mi madre por su cariño y comprensión.*

*A Eduardo, Jorge, Norma, Héctor, Oscar, Leticia, Leslie, Fernanda, Eugenio, Luis, Ester, Susana y a toda mi familia por su apoyo.*

*A Juan Valdez que en el desarrollo de este trabajo siempre estuvo presente con sus ideas y consejos.*

*A todos mis amigos.*

# Indice

Introducción	1
<b>1 Ribozimas: moléculas de ARN catalítico</b>	<b>5</b>
1.1 Ácidos nucleicos	7
1.1.1 Bases	7
1.1.2 Nucleósidos	8
1.1.3 Nucleótidos	10
1.1.4 Formación de polinucleótidos	11
1.2 Tipos de ARN	12
1.3 Conformación terciaria y plegamiento del ARN	13
1.3.1 Puentes de hidrógeno	14
1.3.2 Pares de Watson-Crick	14
1.3.3 Otro tipo de interacciones	16
1.4 El ARN catalítico	16
1.4.1 Aplicaciones	19
1.5 Papel de los iones metálicos en la catálisis y plegamiento del ARN	20
1.5.1 Propiedades de los iones metálicos	20
1.5.2 Unión de iones divalentes al ARN	22
1.6 Ribozima "Hammerhead"	23
1.6.1 Estructura primaria y secundaria	23
1.6.2 Estructura terciaria	24
1.6.3 El sitio de rompimiento y la catálisis del "Hammerhead"	28
1.7 Origen de los datos estructurales de las moléculas	30
1.7.1 Bases de datos de estructuras moleculares	32
1.7.2 Ribozima "Hammerhead" cristalográfico	32
1.8 Planteamiento del problema	34

<b>2</b>	<b>Dinámica Molecular</b>	<b>37</b>
2.1	Ecuaciones de movimiento y teoremas de conservación	39
2.2	Métodos de integración	43
2.2.1	Integrador de Verlet	44
2.2.2	Integrador de velocidad de Verlet	46
2.2.3	Integrador "Leap-Frog"	46
2.2.4	Métodos de predecir-correr	47
2.3	Condiciones de frontera	48
2.4	Ensamblajes	48
2.4.1	Dinámica molecular a temperatura constante	49
2.4.2	Dinámica molecular a presión constante	50
2.5	Variables estadísticas	52
2.6	Métodos de minimización de la energía	56
2.6.1	Método del descenso más profundo	56
2.6.2	Técnica del gradiente conjugado	56
2.6.3	Método Newton-Raphson	56
2.6.4	Método de base adaptada Newton-Raphson	57
2.6.5	Dinámica y enfriamiento	57
2.7	Dinámica molecular cuántica	58
2.7.1	Aproximación de Born-Oppenheimer	59
2.7.2	Teoría de funcional de la densidad	61
2.7.3	Orbitales moleculares y atómicos	63
<b>3</b>	<b>Campos de Fuerza</b>	<b>65</b>
3.1	Factores para la formulación de un campo de fuerza	66
3.2	Formas funcionales	68
3.2.1	Energía de enlace	70
3.2.2	Energía de ángulo de enlace	71
3.2.3	Energía de ángulo de torsión	72
3.2.4	Energía de ángulo de torsión impropia	73
3.2.5	Potencial de Lennard-Jones	73
3.2.6	Energía de interacciones de no enlace	75
3.2.7	Puentes de hidrógeno	76
3.2.8	Interacciones entre las moléculas de agua	76
3.2.9	Constricciones	78
3.3	Programa computacional CHARMM	79
3.3.1	Estructurar la molécula	80

3.3.2	Ambiente . . . . .	83
3.3.3	Minimización . . . . .	83
3.3.4	Cristalizar . . . . .	84
3.3.5	Dinámica . . . . .	84
<b>4</b>	<b>Resultados de la dinámica molecular clásica</b>	<b>87</b>
4.1	Dinámica de la molécula "Hammerhead" completa sin átomos de magnesio . . . . .	89
4.2	Dinámicas de la molécula completa con un átomo de magnesio . . . . .	91
4.2.1	Posición VI . . . . .	92
4.2.2	Posición VI, dinámica corta sin cristal . . . . .	96
4.2.3	Posición VI, dinámica corta con cristal . . . . .	100
4.2.4	Posición III . . . . .	103
4.3	Dinámicas de la molécula completa con dos átomos de magnesio . . . . .	106
4.3.1	Posición III y VI . . . . .	107
4.4	Dinámica de la molécula completa con cinco átomos de magnesio . . . . .	111
4.4.1	Posición I, II, III, IV y VI . . . . .	111
4.5	Análisis de puentes de hidrógeno en las dinámicas . . . . .	115
4.6	Dinámicas ab initio de la zona catalítica . . . . .	117
4.6.1	Dinámica con un ión de magnesio . . . . .	117
4.6.2	Dinámica con dos iones de magnesio . . . . .	119
<b>5</b>	<b>Conclusiones</b>	<b>123</b>
	Apéndice I. Archivo del programa de entrada de CHARMM	131
	Apéndice II. Energía de las dinámicas	133
	Referencias	135

# Lista de Figuras

1.1	Estructuras básicas de la pirimidina y la purina. . . . .	7
1.2	Bases nucleicas: (C) citosina, (U) uracilo, (T) timina, (A) adenina y (G) guanina. . . . .	8
1.3	Las pentosas de los ácidos nucleicos: (A) ribosa, (B) 2'-desoxirribosa. . . . .	9
1.4	Enlace glicosídico. . . . .	9
1.5	Ácido fosfórico. . . . .	11
1.6	Enlace del ácido fosfórico a una pentosa. . . . .	11
1.7	Enlace entre nucleótidos. . . . .	12
1.8	Estructura primaria de un polinucleótido. . . . .	13
1.9	Estructura secundaria y terciaria del ARN de transferencia. . . . .	14
1.10	Apareamiento de bases Watson-Crick. . . . .	16
1.11	Enlace antiparalelo entre cadenas de polinucleótidos. . . . .	17
1.12	Pasos de acción de una ribozima: (A) la complementariedad de la ribozima permite que interaccione con un ARN blanco, (B) se une la ribozima con el blanco formando pares W-C, (C) un ión metálico divalente activa la reacción de catálisis, (D) el blanco se rompe y disocia. La ribozima puede repetir el ciclo. . . . .	19
1.13	Estructura secundaria de la ribozima "Hammerhead". . . . .	24
1.14	(A) Estructura terciaria de la ribozima "Hammerhead", (B) acercamiento a la zona catalítica. . . . .	26
1.15	Dominio I, puentes de hidrógeno. . . . .	27
1.16	Puentes de hidrógeno entre $U_7 - G_8 - C_{12}$ , $G_{12} - A_{13}$ (arriba), $A_{15.1} - U_{16.1}$ y $C_3 - C_{17}$ (abajo). . . . .	28
1.17	Mecanismo de reacción con uno y dos metales . . . . .	30
1.18	Ribozima "Hammerhead" cristalográfico . . . . .	34
1.19	Esquema de los nucleótidos involucrados en la catálisis. . . . .	35
3.1	Potencial de enlace. . . . .	70
3.2	(A) Potencial de ángulo de enlace, (B) diferentes valores del ángulo de enlace. . . . .	71
3.3	Potencial de torsión. . . . .	72



3.4	Definición de ángulo de torsión impropia. . . . .	73
3.5	$u(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right]$ , $\sigma = 1.5$ , $\epsilon = 0.03$ . . . . .	74
3.6	Modelo de una molécula de agua para el potencial $ST^2$ . . . . .	77
3.7	Ribozima "Hammerhead" solvatada . . . . .	83
4.1	Imagen de las posiciones de los cinco iones de magnesio. . . . .	88
4.2	Dinámica sin metales: (A) distancia relativa entre $P - O2'C_{17}$ , (B) ángulo de la conformación $CCA\ O5'A_{1,1} - P.A_{1,1} - O2'C_{17}$ . . . . .	90
4.3	(A) Imagen de la posición inicial del $Mg_{VI}$ , (B) acercamiento a la zona catalítica con $Mg_{VI}$ . . . . .	92
4.4	Zona catalítica después de 200 ps con $Mg_{VI}$ . . . . .	93
4.5	Dinámica con un metal, $Mg_{VI}$ : distancias relativas . . . . .	94
4.6	Dinámica con un metal, $Mg_{VI}$ : ángulos relativos . . . . .	95
4.7	Imagen de las posiciones iniciales del modelo corto del "Hammerhead". . . . .	96
4.8	Imagen de la zona catalítica en el modelo corto sin cristal: (A) después de una minimización de 30 pasos, (B) después de la dinámica de 200 ps. . . . .	97
4.9	Dinámica del modelo corto sin cristal, $Mg_{VI}$ : distancias relativas. . . . .	98
4.10	Dinámica del modelo corto sin cristal, $Mg_{VI}$ : ángulos relativos. . . . .	99
4.11	Imagen de la zona catalítica en el modelo corto con cristal: (A) después de una minimización de 30 pasos, (B) después de la dinámica de 200 ps. . . . .	100
4.12	Dinámica del modelo corto con cristal, $Mg_{VI}$ : distancias relativas. . . . .	101
4.13	Dinámica del modelo corto con cristal, $Mg_{VI}$ : ángulos relativos. . . . .	102
4.14	(A) Imagen de la posición inicial del $Mg_{III}$ , (B) acercamiento a la zona catalítica con $Mg_{III}$ . . . . .	103
4.15	Dinámica con un metal, $Mg_{III}$ : distancias relativas . . . . .	104
4.16	Dinámica con un metal, $Mg_{III}$ : ángulos relativos . . . . .	105
4.17	(A) Imagen de las posiciones iniciales con dos magnesios, $Mg_{VI}$ y $Mg_{III}$ , (B) Acercamiento a la zona catalítica con $Mg_{VI}$ y $Mg_{III}$ . . . . .	107
4.18	Dinámica con dos metales, $Mg_{VI}$ y $Mg_{III}$ : distancias relativas. . . . .	108
4.19	Dinámica con dos metales, $Mg_{VI}$ y $Mg_{III}$ : distancias relativas. . . . .	109
4.20	Dinámica con dos metales, $Mg_{VI}$ y $Mg_{III}$ : ángulos relativos. . . . .	110
4.21	Dinámica con cinco magnesios: distancias relativas. . . . .	113
4.22	Dinámica con cinco magnesios: ángulos relativos. . . . .	114
4.23	Gráficas de distancias interatómicas en la dinámica <i>ab initio</i> con un átomo de magnesio . . . . .	117
4.24	Dinámica cuántica con un ion de magnesio . . . . .	118
4.25	Gráficas de las distancias interatómicas en la Dinámica <i>ab initio</i> con dos metales . . . . .	119
4.26	Dinámica cuántica con dos iones de magnesio . . . . .	120

1	(A) Energía de la dinámica sin $Mg$ , (B) energía de la dinámica con un metal $Mg_{VI}$ , (C) energía de la dinámica con un metal $Mg_{III}$ . . . . .	133
2	(D) Energía de la dinámica con dos metales $Mg_{III}$ y $Mg_{VI}$ , (E) energía de la dinámica con cinco metales $Mg_{(I,II,III,IV,VI)}$ , (F) energía de la dinámica del modelo corto sin cristal, (G) energía de la dinámica del modelo corto con cristal. . . . .	134

# Introducción

La teoría científica tiene como último fin el entendimiento y la predicción de los fenómenos naturales. Hasta ahora, tanto la Física como la Química han podido conjuntar la teoría con los experimentos para interpretar los resultados de las mediciones. La Biología se ha mantenido más distante a implementar los logros teóricos ya que los sistemas biológicos (desde una molécula hasta una población) son sistemas con un gran número de variables cuya complejidad en muy raras ocasiones se puede explicar a partir de un modelo teórico que logre predecir de forma eficaz el comportamiento del sistema.

En los últimos años se ha tenido un considerable éxito al estudiar sistemas biológicos debido principalmente a dos causas, una es al desarrollo de nuevos conceptos en Física y Química que han permitido atacar el problema desde otro punto de vista como ha sido el desarrollo de teorías que facilitan los cálculos moleculares así como la implementación de teorías de dinámica no lineal a sistemas complejos como los fenómenos biológicos, entre otros nuevos conceptos. La otra razón del auge en el estudio de sistemas biológicos ha sido la implementación de cálculos computacionales menos costosos. Esta nueva forma de tratar a los sistemas biológicos utilizando principios físicos y químicos ha dado pie al desarrollo de lo que conocemos como Biofísica Teórica.

Las moléculas biológicas son sistemas inherentemente dinámicos. La actividad biológica de las moléculas depende de la temperatura y al modificar esta variable se afecta la movilidad térmica de las moléculas. Por ello, es necesario una mejor apreciación de la flexibilidad molecular y su dinámica para entender el comportamiento de las moléculas tanto de origen natural como para el diseño de nuevas moléculas con actividad específica.

La mayoría de los sistemas biológicos son sistemas grandes, aún las moléculas más pequeñas consisten de miles de átomos. Además, como estudiar a un

sistema en el vacío es insuficiente, el efecto de los alrededores tiene que ser considerado, lo que aumenta considerablemente el número de átomos<sup>1</sup>. Todo ello ha llevado a la creación de nuevas herramientas para el estudio de sistemas biomoleculares que no están necesariamente relacionadas con aquellos problemas resueltos en Física en los cuales se trabaja en el vacío o en fase gaseosa.

Un sistema molecular debería describirse en principio utilizando conceptos de mecánica cuántica, pero este planteamiento es complicado, los cálculos son largos y tediosos y debido al gran número de simplificaciones que se deben hacer, en pocas ocasiones permiten esclarecer el comportamiento de un sistema biológico. Es por ello que las simulaciones de sistemas biológicos están basadas en modelos físicos simples que permiten tener un mejor conocimiento sobre sus movimientos internos. La base de esto es que cualquier sistema, sea simple o complejo, puede ser visto como una colección de partículas elementales. Los métodos que se han desarrollado para este tipo de estudios son la simulación Montecarlo, la Teoría Cinética, la Mecánica Molecular, la Dinámica Molecular y la Mecánica Estadística.

El objetivo de este trabajo es estudiar los cambios conformacionales de una molécula de importancia biológica. El método que se ha elegido para el análisis es la Dinámica Molecular (DM). La DM ofrece ventajas sobre los otros métodos mencionados ya que permite tener un seguimiento temporal de la trayectoria del sistema y mostrar cambios conformacionales en tiempo real. La DM utiliza un potencial semiempírico o **Campo de Fuerza** cuyos parámetros (basados en cálculos *ab initio*) han sido ajustados explícitamente para el estudio de propiedades de sistemas biológicos y ofrece grandes ventajas al permitir modelar un gran número de átomos con un gran ahorro de tiempo computacional que los cálculos cuánticos no podrían permitir.

La molécula biológica que se quiere estudiar es un **ácido ribonucleico** con propiedades enzimáticas. Esta molécula recibe el nombre de **ribozima** y el trabajo se enfoca en la ribozima más pequeña conocida que es la ribozima "Hammerhead". La ribozima "Hammerhead" tiene la capacidad de unirse a otra molécula de ARN (blanco) a través de apareamiento **Watson-Crick** y

---

<sup>1</sup> Simplemente tomemos como ejemplo el presente trabajo, la molécula de ARN está formada por 1322 átomos y al colocarla en un cubo de agua el número de átomos aumenta a casi 9000.

romper esta molécula blanco. De aquí surge la importancia que a cobrado en los últimos días el estudio de las ribozimas ya que el conocimiento de su mecanismo de reacción permitirá el diseño de una nueva tecnología de fármacos genéticos que estarán dirigidos a combatir enfermedades y síndromes a partir de la alteración del material genético de organismo patológico.

Recientemente se ha podido resolver la estructura de la ribozima "Hammerhead" a través de cristalografía de rayos X por Scott et al [14], sin embargo, no se ha podido definir cuál es el mecanismo de reacción de la molécula, ni cuáles son las condiciones necesarias para que se efectúe la catálisis [15]. Se sabe que la presencia de iones metálicos divalentes aceleran la catálisis, pero tampoco se sabe cuál es el número ni las posiciones en las cuales se deban encontrar estos iones.

Existen dos problemas sobre la ribozima "Hammerhead" que se quieren atacar en el presente trabajo. El primero es la determinación de los cambios estructurales necesarios para que se efectúe la catálisis así como el número de iones metálicos divalentes que participan en esta transformación. Torres et al [6] han establecido un criterio para el análisis de este cambio estructural llamado Conformación Cercana a la línea de Ataque (CCA). Esta conformación se caracteriza porque los átomos están colocados en posiciones previas a la catálisis. El segundo problema es cómo se efectúa este rompimiento. La forma como se pretenden atacar estos problemas es a través de simulaciones moleculares. El primer problema se basa en un cambio conformacional, para lo cual se ha elegido Dinámica Molecular clásica. El segundo problema implica la observación del rompimiento enlaces atómicos, por lo que la DM basada en Campos de Fuerza ya no es útil, para este caso se utilizarán cálculos de DM cuántica.

Las simulaciones realizadas con DM clásica tienen como variables el número de iones de magnesio, así como un pequeño estudio que trata de justificar la posibilidad de reducir la simulación sobre la molécula completa en una simulación donde solo se incluyan los átomos cercanos a la zona de rompimiento. Posteriormente se efectuaron cálculos *ab initio* para observar el rompimiento de la molécula de ARN.

Los resultados obtenidos con DM clásica muestran que es posible simular con este método la ribozima "Hammerhead" completa en presencia de moléculas de agua y que el sistema muestra gran estabilidad (todos los puentes de

hidrógeno predichos por la estructura cristalográfica se mantienen constantes). También se pueden observar las zonas a las cuales se coordinan con mayor afinidad los iones de magnesio y esto también va de acuerdo con la evidencia experimental sobre la localización de los iones divalentes [17, 15, 24]. El cambio conformacional que indique una tendencia al rompimiento no se logra observar claramente, sin embargo, aún no se puede descartar la posibilidad que una implementación de mejores parámetros y mayor tiempo computacional permitan que estos se observen. Las dinámicas cuánticas indican que el rompimiento predicho es posible en la presencia de iones de magnesio, más este rompimiento se observa tanto con un metal como con dos metales, por lo que con los datos obtenidos no es posible discernir cuántos iones metálicos son necesarios en la catálisis. Estos resultados nos hacen sugerir que los mecanismos de reacción que hasta el momento se han propuesto necesitan ser revisados, ya que es probable que el rompimiento sea precursor al cambio conformacional, lo cual implicaría la necesidad de otro tipo de simulaciones para observar el rompimiento.

# Capítulo 1

## Ribozimas: moléculas de ARN catalítico

Las moléculas más importantes que conforman a un organismo vivo son las proteínas y los ácidos nucleicos; esta importancia es debida principalmente a la gran variedad de funciones básicas que pueden realizar. La forma de actuar de estas dos moléculas en la célula está altamente relacionada, de tal manera que no puede existir una sin la otra. Actualmente no se sabe de manera concluyente cuando ni como surgieron estas moléculas y asumieron estos papeles tan esenciales para la vida, pero lo que si se puede decir es que la vida no se daría sin la intervención de ellas. La estructura química de las proteínas y de los ácidos nucleicos es similar porque son polímeros; las proteínas son polímeros de aminoácidos y los ácidos nucleicos son polímeros de nucleótidos. Sin embargo, cada molécula realiza dentro de la célula funciones muy diferentes. Mientras que las proteínas forman parte de los elementos estructurales de los organismos y proveen de la energía necesaria para los procesos biológicos, los ácidos nucleicos participan en la coordinación de la secuencia de aminoácidos (síntesis de proteínas) y en el almacenamiento de la información básica para realizar todos los procesos que mantienen con vida a la célula.

Cuando las proteínas realizan la función de aportar energía lo hacen en forma de enzimas. Esta conformación las hace capaces de catalizar reacciones bioquímicas. Debido a que la secuencia de aminoácidos es capaz de plegarse de forma única, las enzimas son moléculas con una conformación específica que se unen a ciertos sitios de otras moléculas a los cuales no

acceder bajo otro tipo de conformación. Los ácidos nucleicos al ser portadores del código de esta secuencia de aminoácidos deben permanecer inalterados durante cualquier proceso con el fin de mantener la estructura de las proteínas de forma intacta. Un cambio en la secuencia de nucleótidos podría implicar un cambio en la secuencia de aminoácidos teniendo como consecuencia la síntesis de una proteína distinta, la cual podría ser inútil para la actividad enzimática o podría convertirse en un agente patógeno [1].

Cuando ocurre el proceso de reproducción celular, lo único que necesita transmitir la célula madre es este código inmerso en la secuencia de ácidos nucleicos y con esta información, la célula es capaz de generar su propia estructura. Eso hace que los ácidos nucleicos sean considerados como las moléculas portadoras del código genético, es decir, tienen la información necesaria para mantener la vida tal cual, generación tras generación [1].

Las dos funciones básicas que realizan los ácidos nucleicos resultan ser: la preservación del código genético y la coordinación de la síntesis de proteínas. Para esto es que existen dos tipos de ácidos nucleicos, los cuales realizarán una función específica y estos son: el Ácido Desoxirribonucleico (ADN) y el Ácido Ribonucleico (ARN), aunque hay que mencionar que no todas las células tienen los dos tipos de ácidos nucleicos, ver Elliot [1].

Recientemente Thomas Cech [12] y Sidney Altman [13] encontraron que además de estas funciones, el ARN es capaz de catalizar reacciones sin perder su estructura original, es decir, actuar como una enzima<sup>1</sup>. La idea que había prevalecido hasta antes de este descubrimiento había sido que las únicas moléculas que podían realizar la función de enzima eran las proteínas. Esto ha provocado que varias teorías sobre la evolución y origen de los ácidos nucleicos hayan cambiado poniendo como precursor del ADN y de las proteínas al ARN.

En las siguientes secciones se presenta de forma detallada como se encuentran conformados los ácidos nucleicos y en particular el ARN. Al final del capítulo se da una descripción amplia sobre una molécula de ARN que puede catalizar reacciones, a estas moléculas se les conoce como ribozimas y en particular se va a estudiar la reacción en que interviene la ribozima llamada "Hammerhead".

---

<sup>1</sup> Ambos ganadores del Premio Nobel de Química en 1989 por el descubrimiento de las propiedades catalíticas del ARN.



## 1.1 Ácidos nucleicos

Existen dos tipos básicos de ácidos nucleicos, el **Ácido Desoxirribonucleico (ADN)** y el **Ácido Ribonucleico (ARN)**, ambos tienen estructura similar pero sus funciones son diferentes. El ADN tiene el papel principal de guardar la información genética de la célula para sintetizar macromoléculas, mientras que el ARN participa en la transcripción, traducción e incluso transporte de esta información.

Las partes que conforman los ácidos nucleicos son: una base nucleica, una molécula de azúcar y un grupo fosfato. El azúcar y el grupo fosfato son el esqueleto que da forma a la cadena, en tanto que las bases son distintas y dan lugar a una secuencia.

A continuación se discutirán las bases nucleicas y las distintas formas como se encuentran en la naturaleza.

### 1.1.1 Bases

Las bases nucleicas son derivados tanto de la pirimidina como de la purina<sup>2</sup>, figura 1.1. La pirimidina es un anillo de 6 miembros con 2 átomos de nitrógeno y 4 de carbono. La purina es un sistema de 2 anillos unidos, uno de pirimidina más uno con 5 elementos llamado imidazol. El imidazol tiene 3 átomos de carbono y 2 de nitrógeno.

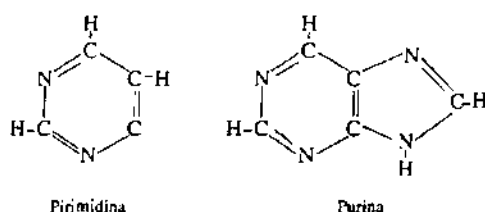


Figura 1.1: Estructuras básicas de la pirimidina y la purina.

La pirimidina es plana y la purina se desvía un poco del plano teniendo un ligero doblez en la unión de los anillos. Ambas moléculas son poco solubles en agua.

<sup>2</sup>La pirimidina y la purina son dos moléculas heterocíclicas caracterizadas por formar un anillo de uno o más tipos de átomos, en este caso, los anillos están formados básicamente de carbono y nitrógeno. De ahí el nombre de bases nitrogenadas.

Las pirimidinas que se encuentran en la naturaleza de forma más común son la citosina, uracilo y timina (5-metiluracilo), mientras que las purinas más comunes son adenina y guanina, figura 1.2.

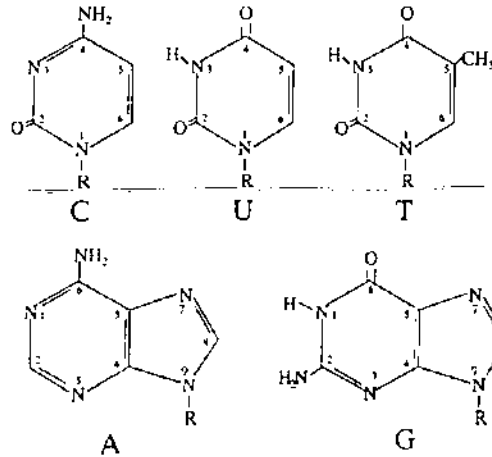


Figura 1.2: Bases nucleicas: (C) citosina, (U) uracilo, (T) timina, (A) adenina y (G) guanina.

La citosina y la timina son las pirimidinas encontradas típicamente en el ADN y la citosina y el uracilo en el ARN. Las bases púricas adenina y guanina se encuentran tanto en el ADN como en el ARN. Las bases nucleicas se encuentran a su vez unidas a un azúcar para conformar los nucleósidos, la forma como se enlazan se discute en la siguiente sección.

### 1.1.2 Nucleósidos

Un nucleósido esta constituido por una base y un azúcar. El azúcar es un anillo de 5 carbonos y un oxígeno llamada pentosa, en el caso del ARN es *ribosa* y para el ADN es 2' - *desoxirribosa*, figura 1.3.

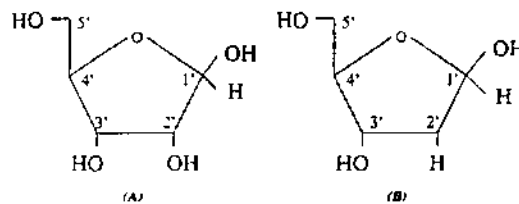


Figura 1.3: Las pentosas de los ácidos nucleicos: (A) ribosa, (B) 2'-desoxirribosa.

Nótese que en la figura anterior, la ribosa del ARN tiene un grupo  $2' - OH$  o  $2' - hidroxilo$  que está ausente en el ADN. Para el caso del ADN tal ausencia permite una mayor resistencia a la hidrólisis (rompimiento de enlaces con la liberación de una molécula de agua).

La forma como se une una base a el azúcar correspondiente es a través del **enlace glicosídico** [2, 3]. El enlace glicosídico se da entre el átomo 1 de carbono de la pentosa y los átomos de nitrógeno en la posición 9 para el anillo de la purina y 1 para el anillo de la pirimidina, figura 1.4. El enlace es *beta*, es decir, por encima del plano del anillo y se da a través de la pérdida de una molécula de agua.

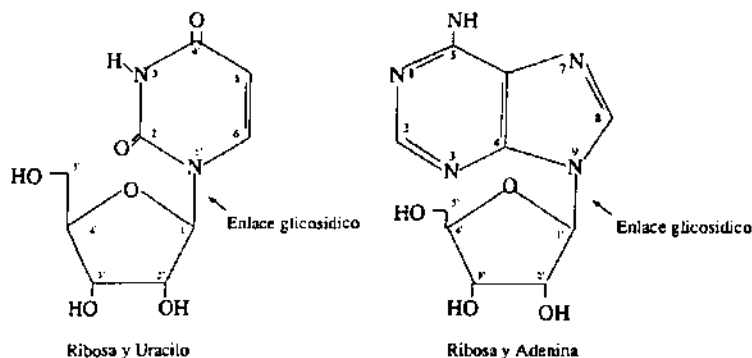


Figura 1.4: Enlace glicosídico.

Resumiendo hasta aquí, se observan dos diferencias fundamentales que distinguen al ADN del ARN [3]:

1. El ADN contiene  $2 - desoxirribosa$  en lugar de *ribosa*.
2. El ADN contiene timina en lugar de uracilo.

La razón de éstas diferencias posiblemente se basa en las funciones que el ADN y el ARN desempeñan en la naturaleza. El ADN tiene como función guardar la información genética viéndose obligada tal molécula a ser más estable que el ARN. El ARN, en cambio, se encuentra en muchas copias y en diferentes formas en la célula realizando un papel más dinámico.

A continuación discutiremos el último componente que conforma a las cadenas de ácidos nucleicos.

### 1.1.3 Nucleótidos

La estructura general de un nucleótido es: **base-azúcar-fosfato**. La unión del fosfato al nucleósido se da mediante la unión de un grupo *OH* o hidroxilo del ácido fosfórico, figura 1.5.

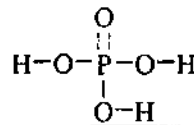


Figura 1.5: Ácido fosfórico.

Al enlace que une el grupo fosfato al azúcar se le denomina **enlace ester**. Este enlace es la unión del oxígeno del grupo *C5'OH* con el fósforo liberando una molécula de agua, figura 1.6.

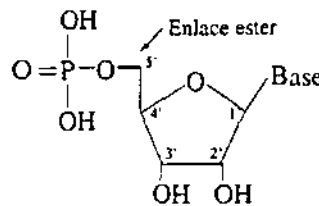


Figura 1.6: Enlace del ácido fosfórico a una pentosa.

Los fosfatos dan la característica ácida al ADN y ARN. Los grupos  $^-OH$  del ácido fosfórico se ionizan en un pH fisiológico, lo cual significa que son altamente hidrofílicos.<sup>3</sup> De ahí que los ácidos nucleicos estén cargados negativamente. Por otro lado el azúcar con su grupo hidroxilo es también altamente hidrofílico, en tanto que las bases nitrogenadas son poco solubles en agua debido a su naturaleza plana.

Si vemos los ácidos nucleicos como entes poliméricos, entonces podemos definir al nucleótido como la unidad constructiva básica (monómero) para formar al ADN y ARN. Para formar una cadena de ácido nucleico se tienen que unir los monómeros y esto se hace al enlazar el grupo fosfato de un nucleótido con el azúcar de otro nucleótido. A continuación se hace una descripción de estas ligaduras.

<sup>3</sup>Son capaces de formar puentes de hidrógeno con una molécula de agua, lo que significa que son solubles.

### 1.1.4 Formación de polinucleótidos

La formación de un dinucleótido se da entre el oxígeno del carbono 3' que pertenece al azúcar y el grupo *OH* del fosfato del nucleótido contiguo. Esta unión libera una molécula de agua, figura 1.7. En un nucleótido el grupo fosfato solo tiene un enlace ester, pero en el dinucleótido se forma un **enlace fosfodiéster**, donde el fosfato esta unido a dos grupos. Debido a la direccionabilidad del enlace, se le clasifica como un enlace *3' → 5' fosfodiéster* [1]. Donde el 3' indica el carbono que pertenece al azúcar del primer nucleótido y el 5' se refiere al carbono que pertenece al azúcar del segundo nucleótido.

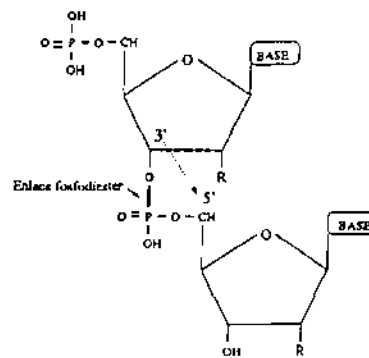


Figura 1.7: Enlace entre nucleótidos.

Cuando se enlazan varios nucleótidos se produce un polinucleótido y su estructura primaria es la siguiente [1]:

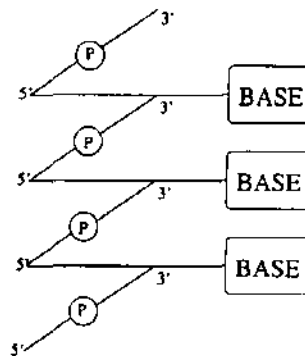


Figura 1.8: Estructura primaria de un polinucleótido.

El ARN existe comúnmente como una sola cadena de bases nucleicas y el ADN como una cadena doble. El tema de este trabajo es el ARN, así que veamos en la siguiente sección cuáles son las estructuras principales como se presenta el ARN en la naturaleza.

## 1.2 Tipos de ARN

El ARN tiene la mayor cantidad de funciones biológicas importantes. Las células contienen más de 8 veces la cantidad de ARN que de ADN. Las moléculas de ARN son catalogadas en varios tipos: ARN mensajero, ARN ribosomal y ARN de transferencia [1, 3].

El ARN mensajero (ARNm) es una transcripción del ADN. Tiene la función de llevar la información del ADN a los sitios en donde ocurre la síntesis de proteínas en la célula. Estos sitios se conocen como Ribosomas. El ARNm tiene diversas secuencias y tamaños, pero es posible distinguir una estructura general, el ARNm siempre tiene en un extremo un grupo trifosfato entre dos nucleótidos: una guanosina y una purina. Esta estructura le permite ser identificado como ARNm para el Ribosoma. En el otro extremo también se distingue un tallo de poliadenina. Esta secuencia en el inicio y fin del ARNm le proporciona estabilidad dentro de la célula.

Los ribosomas son organelos que se encuentran en el interior de las células y es en este lugar en donde ocurre la síntesis de proteínas. Estos organelos están formados en un 65 % por ARN. Las moléculas de ARN ribosomal (ARNr) se pliegan en estructuras secundarias características como consecuencia de interacciones entre sus fragmentos. En el ribosoma existen tres tipos principales de ARNr que se clasifican generalmente de acuerdo a su coeficiente de sedimentación que es una medida burda de su tamaño.

El ARN de transferencia (ARNt) funciona como portador de los aminoácidos residuales para la síntesis de proteínas. El tipo de aminoácido que el ARNt lleve al ribosoma depende de la secuencia de aminoácidos del ARNm. El conjunto de tres bases consecutivas del ARNm definen la clave para un código que identifica a cada aminoácido y recibe el nombre de **codón** [1]. El ARNt es el intérprete entre la instrucción indicada por un codón del ARNm y el aminoácido correspondiente. Por lo mismo tiene una forma dual, en un extremo presenta una secuencia de tres bases que son complementarias

a un codón (que recibe el nombre de anticodón) que le permiten identificar la secuencia indicada por el ARNm y en el otro extremo tiene una estructura que identifica al aminoácido correspondiente a la instrucción del codón. Por lo que el ARNt se pliega en estructuras secundarias características, figura 1.9.

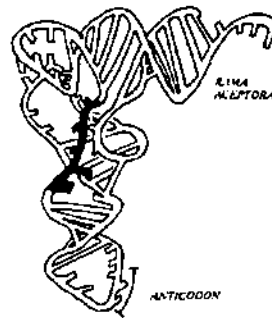


Figura 1.9: Estructura secundaria y terciaria del ARN de transferencia.

Ni las cadenas de ADN, ni las de ARN se encuentran estiradas dentro de la célula, sino que se pliegan en estructuras características. El ADN por ser mucho más largo que el ARN se pliega formando hélices dobles; el ARN que es más corto se pliega dando origen a conformaciones específicas de su función como veremos en la siguiente sección.

### 1.3 Conformación terciaria y plegamiento del ARN

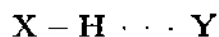
El análisis de la cadena de ARN muestra que esta molécula utiliza muchos tipos de interacciones no covalentes. En orden de estabilidad tenemos:

- Interacciones tipo van der Waals cuyo origen son las fluctuaciones correlacionadas de las nubes electrónicas de los átomos. Estas interacciones se dan entre todos los átomos, y en general son del orden de la energía térmica a temperatura ambiente (0.6 Kcal/mol) [29].
- Aquéllas que se dan en el plano debido a **puentes de hidrógeno**. Este tipo de interacciones es predominantemente (dipolar) electrostática, pero también incluye transferencia de carga [26].

- Aquéllas perpendiculares a los planos de las bases (empalme de bases), estabilizadas principalmente por fuerzas de dispersión y efectos hidrofóbicos [2].

### 1.3.1 Puentes de hidrógeno

Este tipo de enlace es de carácter electrostático. Tienen un papel principal en la estabilización de la estructura de los ácidos nucleicos. Estas fuerzas son de corto alcance, es decir, a distancias menores a 2 Å y dan lugar a enlaces débiles como los pares de Watson-Crick. La forma general de un "puente de hidrógeno" es [2]:



en la cual un átomo de hidrógeno conecta a dos átomos  $X$  e  $Y$  de mayor electronegatividad (como carbono, oxígeno y nitrógeno). El átomo  $X$  está unido covalentemente al Hidrógeno. La longitud de estos enlaces depende de las cargas localizadas sobre los átomos  $X$ ,  $H$ ,  $Y$ . Las interacciones más comunes son las de Watson-Crick que se explican en la siguiente sección y tienen la forma:



### 1.3.2 Pares de Watson-Crick

Los puentes de hidrógeno son enlaces débiles pero el apareamiento de átomos es muy preciso. Estos enlaces por lo mismo también son fáciles de romper, por ejemplo con energía térmica.

Debido a la forma complementaria entre las bases púricas y pirimídicas, éstas se unen a través de puentes de hidrógeno de acuerdo al modelo establecido por Watson y Crick (**W-C**) para la doble cadena del ADN. El enlace favorecido es el de *adenina - timina* y el de *guanina - citosina* [1, 2, 3], figura 1.10.

La geometría del apareamiento muestra que el tamaño del par de bases es aproximadamente el mismo ya que siempre incluye una base púrica (molécula grande) y una pirimidina (molécula chica). También se puede notar que la



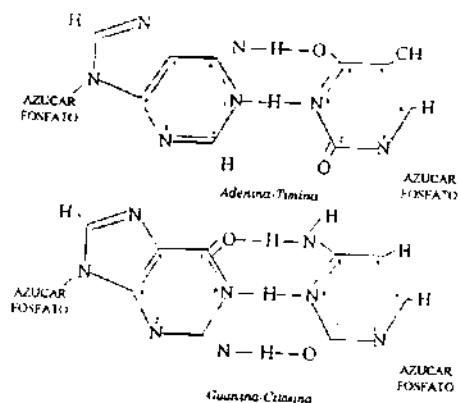


Figura 1.10: Apareamiento de bases Watson-Crick.

unión  $G \equiv C$  es más fuerte ya que se tienen tres puentes de hidrógeno, mientras que en la unión  $A = T$  sólo se tienen dos.

Tanto el ADN como el ARN pueden formar cadenas dobles en donde dos hileras se unen en sentido antiparalelo, es decir, una cadena tiene un extremo con el grupo  $5' - OH$  del azúcar de un nucleótido que no esta conectado a ningún nucleótido, y el otro extremo tiene el grupo  $3' - OH$  que no esta conectado tampoco a ningún otro nucleótido. Que las cadenas sean antiparalelas significa que nunca se encuentran dos grupos  $5'$  o  $3'$  juntos en un extremo, teniéndose entonces la forma de la figura 1.11.

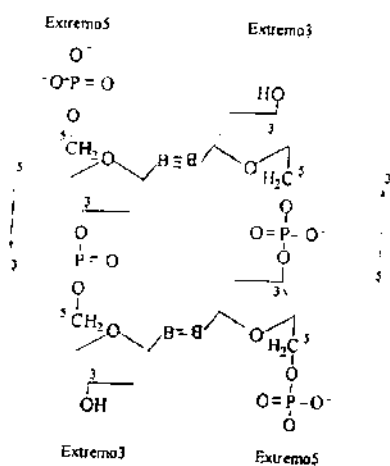


Figura 1.11: Enlace antiparalelo entre cadenas de polinucleótidos.

La característica de cadena doble antiparalela es un requisito necesario para el arreglo estereoquímico (en el espacio) de la estructura helicoidal del ADN. Por conveniencia se hace referencia a la secuencia de bases en la dirección  $5' \rightarrow 3'$  y no es necesario especificar la cadena complementaria.

### 1.3.3 Otro tipo de interacciones

El apareamiento W-C se da principalmente en el ADN que se encuentra como una cadena doble y es la principal interacción que se da de forma horizontal. Esta complementariedad le permite mantenerse inalterado y preservar el código genético. El ARN se encuentra como una sola cadena y al plegarse da origen a los pares W-C en las regiones donde se encuentre esta complementariedad. Esto provoca que el ARN adquiera geometrías diversas de acuerdo a las secuencias de nucleótidos que posee.

Además de las interacciones W-C del ARN, también se encuentran interacciones verticales que le dan la forma específica al ARNt y en especial al ARN catalítico que sobre el cual se enfoca este trabajo [2, 7]. Existen varios tipos de interacciones bien definidas entre las diferentes bases, sin embargo, como solo se dan algunas cuantas en el ARN "Hammerhead", estas serán explicadas cuando se vea la estructura terciaria de esta ribozima en la sección 1.6.2.

Aparte de las interacciones anteriores, existen interacciones entre los átomos debido principalmente a que ciertos grupos se ionizan en un pH fisiológico, permitiendo la formación de zonas cargadas electrostáticamente. Esto provoca atracción entre los diversos grupos que poseen los nucleótidos así como la formación de enlaces muy débiles y de muy corta duración. Los posibles enlaces que tendremos serán entre átomos electronegativos como el oxígeno, carbono, nitrógeno y grupos  $OH^-$ , que atraeran principalmente al  $H^+$ . La interacción de los metales divalentes con el ARN se discute en otra sección.

## 1.4 El ARN catalítico

Las enzimas son catalizadores, es decir, son moléculas capaces de aumentar la tasa de reacción pero no son ni substrato ni producto de esa reacción. Por lo general esta característica enzimática se le asocia a una proteína. Una proteína

es un polímero formado a partir de la combinación de 20 moléculas distintas. El monomero de este polímero se conoce como aminoácido y están unidos a través de enlaces peptídicos [3].

Recientemente se descubrió que no solo las proteínas presentan comportamiento catalítico sino que también la molécula de ARN puede realizar funciones catalíticas. Esto es lo que se conoce como **ribozima** y fue descubierta por Thomas Cech de la Universidad de Colorado y Sidney Altman de la Universidad de Yale en protozoarios ciliados [12, 13]. Hasta este descubrimiento se creía que sólo las proteínas presentaban este comportamiento.

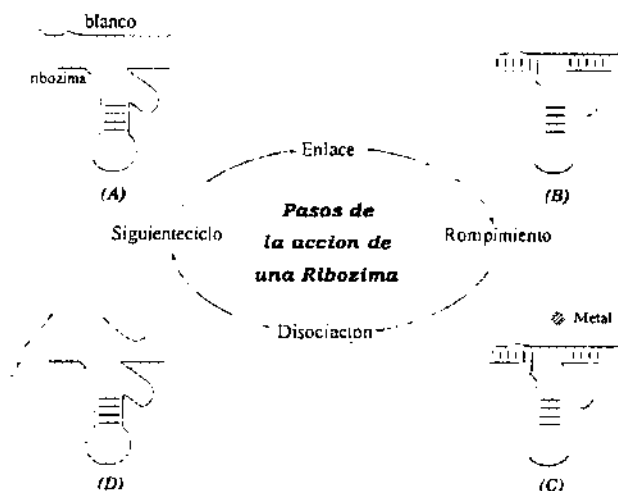


Figura 1.12: Pasos de acción de una ribozima: (A) la complementariedad de la ribozima permite que interaccione con un ARN blanco, (B) se une la ribozima con el blanco formando pares W-C, (C) un ión metálico divalente activa la reacción de catálisis, (D) el blanco se rompe y disocia. La ribozima puede repetir el ciclo.

Las ribozimas son entonces, moléculas de ARN que pueden realizar las mismas funciones que una enzima, es decir:

1. Son específicas de un substrato.
2. Aumentan la tasa de reacción.
3. Emergen de la reacción sin cambio en su estructura.

La función de una ribozima es unirse a una estructura del ARN llamada blanco a través de un apareamiento W-C, figura 1.12. La ribozima rompe al ARN-blanco en un sitio específico de su estructura dejándola inactiva.

La mayoría de los ejemplos conocidos de las ribozimas están asociados con la interacción con ácidos nucleicos cuyas actividades son con el rompimiento y procesamiento de ARNm. Las ribozimas más conocidas se presentan en la Tabla 1.1 en donde *nt* significa nucleótido [8].

Tabla 1.1 Tipos de ribozimas				
Ribozima	Tamaño	Reacción	Nucleófilo	Localización
Grupo I intrón	413 nt	Transesterificación resultando en la unión 3'-5' y un término 3'-OH	3'-OH de una guanosina	Células eucariontes eubacterias y virus
Grupo II intrón	887 nt	Transesterificación resultando en la unión 2'-5' y un término 3'-OH	3'-OH de una guanosina	Organelos de eucariontes y virus
RNasa P	350-410 nt	Rompe el término 5' de un ARNt obteniéndose la unión 2'-5' y un término 3'-OH	agua	Subunidad del ARN en eubacterias
Hammerhead	31-42 nt	Rompimiento del ARN produciendo 2',3' fosfato cíclico y un término 5'-OH	Grupo 2'-OH adyacente	RNA viral de plantas y viroides
Hairpin	50 nt	Rompimiento del ARN produciendo 2',3' fosfato cíclico y un término 5'-OH	Grupo 2'-OH adyacente	ARN satelital del virus del tabaco
Virus hepatitis	84 nt	Rompimiento del ARN produciendo 2',3' fosfato cíclico y un término 5'-OH	Grupo 2'-OH adyacente	HDV
Neurospora	881 nt	Rompimiento del ARN produciendo 2',3' fosfato cíclico y un término 5'-OH	Grupo 2'-OH adyacente	mitocondria Neurospora

La ribozima que se va a estudiar en este trabajo es la ribozima "Hammerhead" que es la más pequeña que se conoce y por lo tanto la que presenta mayor facilidad para ser analizada. La primera estructura crista-

lográfica de una ribozima "Hammerhead" fue obtenida por Dave McKay en Standford [10]. Esta ribozima se presenta en detalle en la sección 1.6.

Los sustratos que se han encontrado en la naturaleza para el ARN catalítico son todos ácidos nucleicos, sin embargo, se han sintetizado sustratos con *ester - amida* y otro tipo de enlaces, de hecho también se ha sintetizado ADN catalítico.

De la estructura del ARN se puede ver fácilmente que ésta tiene potencial para ser catalítico, ya que tiene la flexibilidad para formar una estructura tridimensional en forma globular, la cual puede traer grupos catalíticos en contacto con el sustrato.

### 1.4.1 Aplicaciones

Las ribozimas se pueden utilizar en el estudio de la función genética y en las terapias genéticas contra enfermedades, debido a que un gran número de virus están formados por ARN.

Las ribozimas que están diseñadas para atacar la transcripción de un ARNm se pueden unir a este ARNm y romper la molécula. esto se ha comprobado por experimentos *in vitro*. Estos experimentos muestran la ventaja de utilizar ribozimas en lugar de utilizar terapia genética, la cual disminuye la expresión genética. Con las ribozimas existe la posibilidad de atacar al gen mediante diferentes tipos de ribozimas las cuales tienen diferente eficiencia, permitiendo tener control sobre el número de genes que se desea anular.

Otra aplicación importante de las ribozimas está en el desarrollo de nuevas drogas y protocolos para la terapia genética. Las ribozimas se cree representan una nueva herramienta para atacar agentes infecciosos. Con diferentes enfermedades como el VIH se ha estado investigando en busca de productos genéticos que permitan atacar el problema de una forma más efectiva y controlada. La ventaja de este método es que una ribozima no necesariamente mata a la célula en la que se encuentra, en comparación con muchos fármacos a base de proteínas que si destruyen gran parte del tejido vivo.

## 1.5 Papel de los iones metálicos en la catálisis y plegamiento del ARN

Aproximadamente tres cuartas partes de todos los elementos químicos son metales. Los iones metálicos están asociados con el ARN en solución en un gran número de formas que son fáciles de encontrar en el agua de mar y dentro de las células desempeñando papeles vitales en el plegamiento del ARN. En esta sección examinamos el papel de los iones divalentes en el plegamiento del ARN así como la evidencia experimental de la unión de iones metálicos al ARN en solución y en cristales. Se da un enfoque especial a los iones metálicos que promueven la catálisis en rompimientos que dan lugar a un *fosfato 2, 3-cíclico* y un *5 - hidroxilo* como productos.

### 1.5.1 Propiedades de los iones metálicos

Las concentraciones de iones metálicos en medio intracelular varían en gran medida y se mantienen y ajustan por medio de mecanismos de transporte pasivo y activo<sup>4</sup>. Dentro de las células, las concentraciones de iones metálicos libres son mucho menores que aquellas concentraciones totales de iones metálicos, ya que estos iones metálicos se encuentran unidos a macromoléculas intracelulares.

En soluciones acuosas, los iones metálicos se coordinan para formar complejos con agua o solutos. El número de coordinación (número de enlaces que se forman) y la geometría de coordinación (geometría en la que se acomodan los enlaces) depende de las características tanto el ión metálico como de los enlaces. La mayoría de los iones metálicos forman complejos hexacoordinados con seis moléculas ligantes arreglados en los vértices de un octaedro. En forma menos frecuente, la coordinación de los complejos de cuatro enlaces ocurren en un arreglo tetraédrico [9].

Los iones metálicos muestran preferencia a unirse a diversos complejos. Con base a esta preferencia, tanto los iones metálicos y los enlaces pueden ser clasificados como "fuertes" y "suaves". En soluciones los iones metálicos

---

<sup>4</sup>Transporte pasivo es el que se efectúa por el movimiento aleatorio de las moléculas a través de la membrana, el transporte activo es el transporte en contra de un gradiente de concentración y por lo tanto requiere de energía [4].

fuertes (por ejemplo,  $Li^+$ ,  $Na^+$ ,  $K^+$ ,  $Mg^{++}$  y  $Ca^{++}$ ) se enlazan preferencialmente a ligaduras fuertes (por ejemplo,  $H_2O$ ,  $NH_3$ ,  $OH^-$ , fosfatos, éteres y alcoholes) y los iones metálicos suaves (por ejemplo,  $Pb^{++}$ ,  $Zn^{++}$ ,  $Ni^{++}$ ,  $Co^{++}$  y  $Mn^{++}$ ) se unen preferencialmente a ligaduras suaves (por ejemplo, cianido, fosforothioados, heterocíclicos aromáticos que contienen nitrógenos y tioles). Esta designación fuerte/suave se explica con base a algunas teorías químicas y no es definitiva<sup>5</sup>. Para propósitos de estudios del ARN, la coordinación de los enlaces a través de sulfuros o de átomos de nitrógeno se consideran como suaves y la coordinación de los enlaces a través de oxígeno son considerados como fuertes.

Los iones metálicos intercambian ligaduras de aguas a tasas que varían 18 órdenes de magnitud. Otras ligaduras que no sean agua también tienen afinidad a mostrar grandes diferencias en la tasa de intercambio. Mientras que los iones divalentes más comunes como el  $Mg^{++}$ ,  $Ca^{++}$  y  $Mn^{++}$  tienen un intercambio rápido, en otros iones metálicos como el  $Co^{+++}$  y  $Cr^{+++}$  el intercambio es tan lento que son considerados cinéticamente inertes. La cinética y la termodinámica de estas ligaduras por iones metálicos no se correlacionan de una forma intuitiva; los iones metálicos que tienen la tasa de intercambio más lenta no son necesariamente aquellos que tienen las ligaduras más fuertes [9].

Las moléculas de agua que se unen a iones metálicos suelen ser más ácidas que las moléculas de agua libres. El ión hidroxilo unido a un ion metálico es un potente nucleófilo. De acuerdo con esto, los iones metálicos hidratados pueden participar en la química ya sea de forma general o/y como un ácido o una base [9].

Aunque la oxidación o reducción de los iones metálicos enlazados al ARN bajo condiciones fisiológicas no han sido estudiados, es importante recordar que muchos iones metálicos pueden fácilmente ganar o perder electrones reaccionando con oxígeno disuelto para formar óxidos metálicos insolubles.

---

<sup>5</sup>Huhey, J. E. 1983, *Inorganic Chemistry*.

### 1.5.2 Unión de iones divalentes al ARN

Los iones metálicos divalentes pueden interactuar electrostáticamente con el fosfato de la estructura del ARN. Los experimentos que han medido la unión de iones metálicos divalentes a cadenas (singulares) de ARN alcanzan saturación cuando se tiene una concentración de un ion metálico por cada dos fosfatos. Esto sugiere que cada ión metálico se une a dos (presumiblemente adyacentes) fosfatos de la cadena de ARN [9]. Sin embargo, el puente entre los dos fosfatos a través de un ión metálico, requiere que la conformación de la estructura fosfodiéster esté constreñida con un acompañamiento en la reducción de entropía y en el costo de energía.

Los iones metálicos divalentes son más afines a unirse a hélices homopoliméricas donde los fosfatos están constreñidos estructuralmente y los dos fosfatos en cadenas opuestas tienen la forma de un ARN tipo A para unirse a un solo ión metálico. Ya que los iones divalentes reducen la repulsión electrostática entre las cargas negativas de los fosfatos, esta diferencia en la afinidad entre una sola cadena y una cadena doble de RNA para unirse a un ión divalente, significa que los iones metálicos divalentes son muy efectivos para estabilizar la doble cadena de ARN.

En las estructuras heterogéneas de ARN, los iones metálicos divalentes interactúan electrostáticamente con los fosfatos en las regiones helicoidales y de esta forma estabilizan la estructura secundaria del ARN. Además, existen estructuras donde los fosfatos de una sola cadena son constreñidos de tal forma que se encuentran lo suficientemente cerca en espacio para formar una "bolsa"<sup>6</sup>. Es en estas "bolsas" donde los iones metálicos se pueden unir de forma más afín que a una hélice de ARN. La formación de estas "bolsas" y la unión de un ión metálico se espera sea dependiente del plegamiento completo del ARN.

Es importante remarcar que aunque la discusión anterior se enfoca a las interacciones electrostáticas entre los iones metálicos y el fosfato, los iones metálicos se pueden unir al ARN de otras maneras. Esto es notorio en los metales "suaves" en donde son más comunes las interacciones con las bases. De esta manera, las "bolsas" de iones metálicos pueden formarse tanto en la región de los fosfatos como en las bases de la cadena de ARN.

---

<sup>6</sup>Del inglés pocket.



Se suelen realizar diversos experimentos para determinar la región a la cual se enlazan los iones metálicos. Estos experimentos no miden de forma directa la unión de estos iones metálicos, sino las consecuencias de la unión de este metal. En las técnicas de Rayos X en las cuales no se puede determinar con precisión la localización de estos iones metálicos, se utilizan otros iones divalentes mucho mas grandes y que se espera posean algunas de las propiedades químicas del ión divalente en cuestión.

Las estructuras cristalinas tienen una resolución de 2.5 a 2.7 Å. Esto significa que la baja densidad electrónica de los iones  $Mg^{++}$  hace que sean difíciles de localizar a menos de que se encuentren en sitios con una posición altamente restringida, solamente un número limitado de iones de  $Mg^{++}$  serán localizados dentro de la resolución accesible. Para obtener información acerca de la afinidad de estos iones en el cristal, se requerirá de un experimento que mida la ocupación cristalográfica de cada sitio como función de la concentración de iones de magnesio, los fosfatos se enlazan a otros grupos de tal forma que disminuye el enlace de los iones de  $Mg^{++}$ .

## 1.6 Ribozima "Hammerhead"

### 1.6.1 Estructura primaria y secundaria

El nombre de ribozima "Hammerhead" se le da por la similitud entre su estructura secundaria y la cabeza de martillo, figura 1.13. La ribozima "Hammerhead" es uno de las pocas especies de RNA catalítico que se ha identificado y caracterizado. Al igual que otras ribozimas tiene dirección contraria de ARN aunque difiere en los otros ARN convencionales al ser mas flexible, permitiendo dejar expuestas sólo ciertas regiones activas.

La figura 1.13 muestra la secuencia de nucleótidos indicada por las letras A, C, G y U para la adenina, citosina, guanina y uracilo respectivamente y esta secuencia esta numerada del 1 al 17. En su estructura se puede distinguir que ciertas partes forman enlaces de Watson-Crick (línea punteada en la figura 1.13), para dar lugar a ramas, mientras que otras partes quedan sin aparear para originar bucles (compare figura 1.13 y 1.14). Este apareamiento W-C entre el substrato (S) y la enzima (E) es lo que le da especificidad a la molécula. En el centro de esta misma figura, se distingue un conjunto de bases que no tienen acoplamiento W-C, las bases que comprenden este núcleo no helicoidal

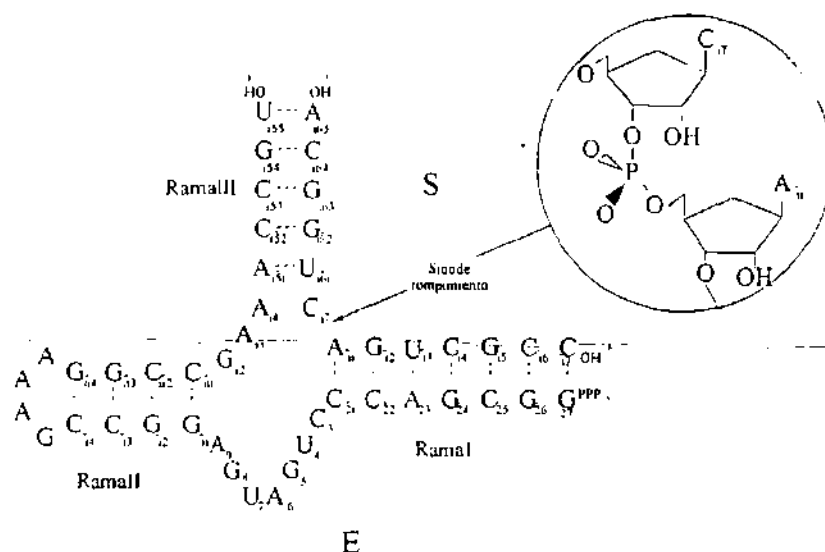


Figura 1.13: Estructura secundaria de la ribozima "Hammerhead".

son  $C_3 - A_9$  y  $G_{12} - A_{14}$ . En esta parte de la molécula también se distingue una región denominada **desacoplamiento GA** que se caracteriza en que las bases que la conforman, guanina (G) y adenina (A), no son complementarias.

La molécula cuenta con tres ramas de bases apareadas que se originan a partir de los nucleótidos  $A_{1.1}$ ,  $C_{2.1}$ ,  $G_{10.1}$ ,  $C_{11.1}$ ,  $A_{15.1}$  y  $U_{16.1}$ . Las bases que conforman estas ramas se clasifican con subíndices con un número entero que indica la base a partir de la cual se originaron. El sitio de rompimiento como se indica en la figura 1.13 se encuentra entre los nucleótidos citosina 17 ( $C_{17}$ ) y adenina 11 ( $A_{1.1}$ ).

## 1.6.2 Estructura terciaria

En la figura 1.14 (izq) se puede observar la estructura terciaria de la molécula "Hammerhead". Esta estructura tiene la forma de una letra "Y", donde las ramas I y II forman los brazos de la "Y" y la rama III forma la base [17]. Las tres ramas forman helices tipo  $A^7$ . Las ramas I y II forman entre sí una vuelta en "U", mientras que las ramas II y III forman casi una línea recta [10].

<sup>7</sup>En el sentido contrario a las manecillas del reloj.

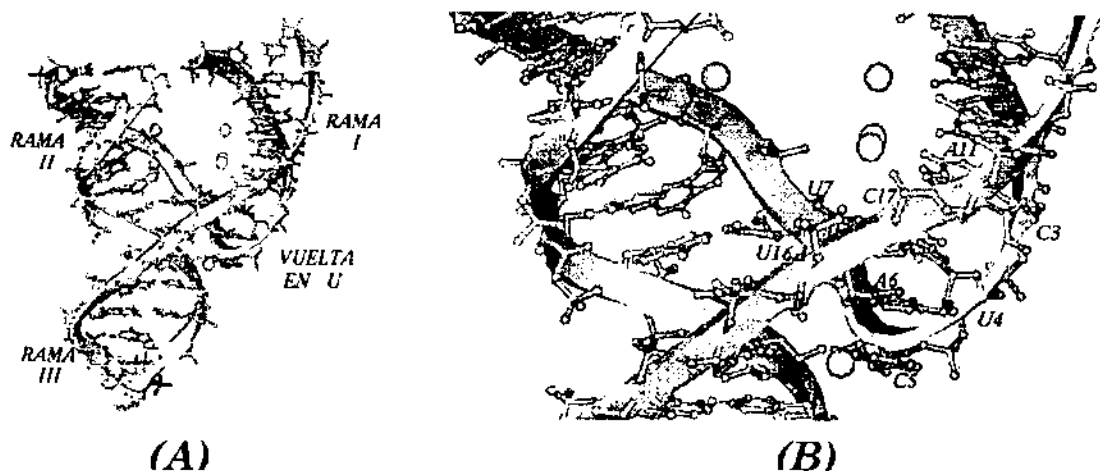


Figura 1.14: (A) Estructura terciaria de la ribozima "Hammerhead". (B) acercamiento a la zona catalítica.

El núcleo central es necesario para la actividad catalítica y esta formado por dos estructuras principales o dominios <sup>8</sup> [10]. Los nucleótidos de la cadena  $C_3-U_4-G_5-A_6$  forman el **Dominio I** y dan un marcado giro en forma de "U" siguiendo la rama I. Los nucleótidos  $U_7-A_9$  junto con  $G_{12}-A_{14}$  forman una cadena doble que no tiene apareamiento Watson-Crick, esta parte es el **Dominio II**. Ver figura 1.14 (der).

En las tres estructuras cristalográficas que se han resultado del "Hammerhead" existen consistencias sobre su estructura terciaria [10, 14, 17], incluso en las bases que conforman a los puentes de hidrógenos en estas estructuras. Sin embargo, se logran observar diferencias como por ejemplo el ángulo entre ramas. De aquí que surjan preguntas sobre si esta flexibilidad relativa observada en el "Hammerhead" es esencial para su actividad o si sólo depende de la forma como se fabricó el cristal en cada estructura.

### Dominio I.

La secuencia de nucleótidos del  $C_3-A_9$  forman una vuelta en "U" muy parecida a un anticodón en el ARN de transferencia. Las interacciones terciarias que le dan estabilidad a esta zona son los enlaces entre hidrógenos del  $N_3$  del  $U_4$  con el O2P del  $U_7$ , así como puentes de hidrógeno entre el hidroxilo del  $U_4$  y el  $N_7$  de la  $A_6$ , figura 1.15 (der). Estas interacciones terciarias han sugerido un

<sup>8</sup>Un dominio es una forma característica que sobresale en la estructura.

concenso sobre cuál podría ser su forma general proponiendo que sea **U-N-R**, donde **N** es cualquier nucleótido y **R** es una base Purina. Sin embargo, a pesar de este concenso, la alteración de cualquiera de estas bases provoca un efecto supresor de la catálisis. Esto sugiere que ese dominio tiene un papel muy importante en la actividad catalítica.

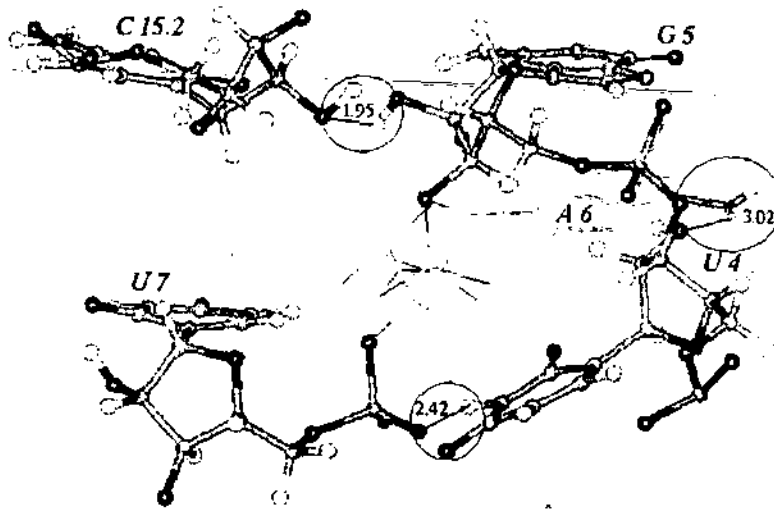


Figura 1.15: Dominio I, puentes de hidrógeno.

El grupo hidroxilo de  $G_5$  está unido por puentes de hidrógeno a la  $C_{15.2}$  en la rama III, la distancia entre oxígenos es de aproximadamente  $2.8 \text{ \AA}$ , esta interacción terciaria da la vuelta en "U" con la rama III y también es muy importante para su conformación, figura 1.15 (der). Si se reemplaza el grupo hidroxilo del azúcar de  $G_5$  por otro grupo funcional se reduce la actividad catalítica. El efecto de alterar el grupo hidroxilo en  $C_{15.2}$  aún no se conoce bien.

## Dominio II

Las tres parejas de nucleótidos  $U_7-A_9$  y  $G_{12}-A_{14}$  forman dos cadenas con bases incompatibles. Estas cadenas son casi paralelas y continuas con las bases de las ramas de tal manera que no se forman más bucles entre las ramas II y III. En la figura 1.16 se observa como tanto la base y el azúcar de  $G_8$  presentan puentes de hidrógeno con los nucleótidos  $A_{13}$  (der., arriba) y  $G_{12}$  (izq., arriba). Esta conformación específica de  $G_8$  y sus interacciones complejas parecen tener un papel crucial en la estructura completa del dominio II.

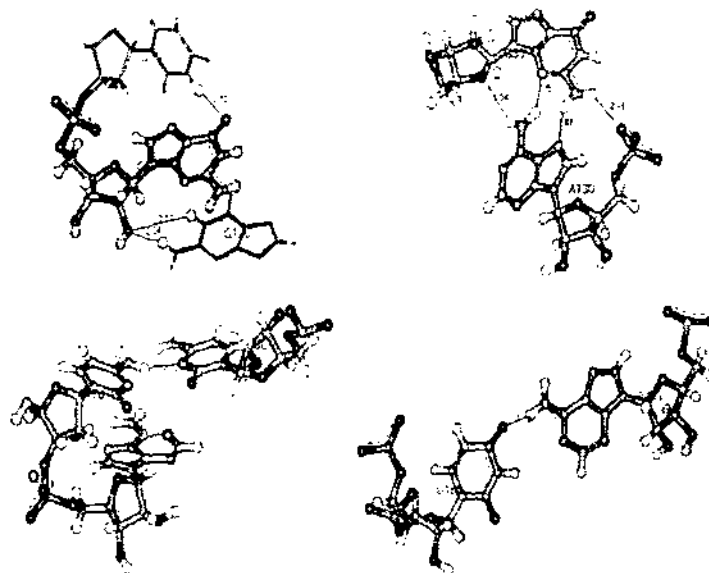


Figura 1.16: Puentes de hidrógeno entre  $U_7 - G_8 - G_{12}$ ,  $G_{12} - A_{13}$  (arriba),  $A_{15.1} - U_{16.1}$  y  $C_3 - C_{17}$  (abajo).

Otras interacciones que no se encuentran relativamente cerca de la vuelta en "U" se observan en la figura 1.16. Vemos por ejemplo la interacción entre la  $A_{15.1}$  y  $U_{16.1}$ , (abajo, der.). A pesar de ser muy similares al apareamiento W-C de A-T, este no se forma ya que solo existe un puente de hidrógeno. También vemos que el nucleótido del sitio de rompimiento, el  $C_{17}$  forma un único puente de hidrógeno con la  $C_{30}$ , figura 1.16 (abajo, izq.)

A pesar de que en las estructuras cristalográficas no se observan los iones metálicos, se sabe que los iones metálicos divalentes (magnesio) intervienen en la catálisis. Algunas de las estructuras resueltas han utilizado iones divalentes como  $Mn^{2+}$  y  $Cd^{2+}$  para identificar los sitios a los cuales se podría adherir el magnesio. Sin embargo, la coordinación de estos metales no es la misma que la del magnesio y estos sitios donde se observan los metales no corresponden necesariamente al sitio real donde interviene el metal en la catálisis. En la siguiente sección se explicarán algunas de las teorías que existen sobre los posibles canales catalíticos del "Hammerhead" y el posible papel del ión metálico divalente para que esta se efectue.

### 1.6.3 El sitio de rompimiento y la catálisis del “Hammerhead”

La ribozima “Hammerhead” es una metaloenzima, es decir, requiere de uno o más metales para realizar su actividad catalítica. Aunque los análisis por rayos X sólo revelan un ión metálico potencial, el número exacto de iones requeridos para la catálisis permanece desconocido [14][10].

La ribozima “Hammerhead” cataliza el sitio específico de un enlace fosfodiéster para generar un grupo libre 5'OH y un fosfodiéster 2', 3' - *cíclico* [9]. La especificidad del sitio de rompimiento está marcado por el apareamiento Watson-Crick en las ramas I y II, como se ve en las figuras 1.14 y 1.15. Cualquier nucleótido excepto Guanosina en la posición 17 permite un rompimiento eficiente.

La ribozima “Hammerhead” requiere uno o más iones catalíticos. Estos iones son divalentes como el  $Mg^{+2}$ ,  $Mn^{+2}$ ,  $Co^{+2}$  o  $Pb^{+2}$ . Uno de estos metales ioniza el 2' - *hidroxil* en el sitio de rompimiento.

Se han propuesto varios modelos para explicar el rompimiento del enlace fosfodiéster del ARN en las ribozimas “Hammerhead”; la mayoría de éstos incluyen un metal solvatado con un ión hidróxido en el centro catalítico, aunque también existe un modelo alternativo con dos metales. A continuación se explican estos dos modelos de acuerdo al diagrama de la figura 1.17.

#### Mecanismo I

En este modelo el ión de *hidróxido metálico* se une al oxígeno *pro - R* en el fosfato en el sitio de rompimiento y activa la ribozima, esto significa que remueve un protón del 2'OH del nucleótido atacado [17]. Este ión así activado, 2'O<sup>-</sup>, ataca el enlace fosfodiéster.

Los pasos generales de este mecanismo son:

- (a) La conformación del estado base, figura 1.15.a.
- (b) El grupo 2' - *hidroxilo* es activado por un enlace hidróxido metálico. Este nuevo nucleófilo ataca al fósforo que es estabilizado por la interacción con el metal (la interacción se da con el oxígeno fosforilado *pro - R<sub>p</sub>*).
- (c) La geometría intermedia es una pirámide doble triangular.
- (d) Los productos son un 2', 3' *fosfodiéster cíclico* y un grupo libre 5' - OH de nucleótido en la posición 1.1.

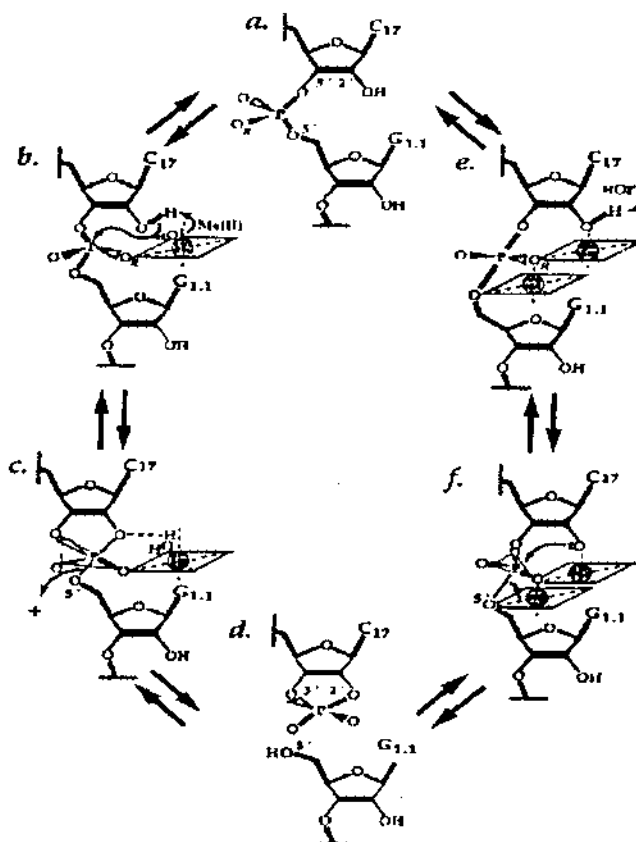


Figura 1.17: Mecanismo de reacción con uno y dos metales

### Mecanismo II

Este mecanismo fue propuesto con el fin de estabilizar la carga que se forma en el 5' oxianión, ya que esto puede entorpecer la efectividad del grupo que se desprende [17]. Los pasos que comprende este mecanismo son:

(e) Se sugiere la intervención de dos metales. En este caso, un grupo hidróxido quita un protón del  $O2' - OH$ . Los metales se adhieren al  $2'OH$ , al oxígeno *pro - R* de la  $A_{1.1}$  y al grupo  $O5'$ .

(f) El  $2'$  oxianión ejecuta un ataque en línea sobre el fosfato, estabilizando los iones divalentes. La geometría intermedia es una pirámide doble triangular.

El grupo que se desprende se estabiliza por la asociación directa con el metal.

(d) Se forma el enlace fosfodiéster y la molécula se separa.

Para la catálisis, en cualquiera de los dos mecanismos, es esencial la ribosa  $2' - OH$  para el rompimiento. Se cree que el ion  $Mg^{2+}$  es el ión catalítico natural para el Hammerhead, pero puede ser substituido por otros incluyendo  $Mn^{2+}$ ,  $Cd^{2+}$  y  $Co^{2+}$  que tienen tasas de rompimiento mayores que el  $Mg^{2+}$  bajo las mismas condiciones, y por el  $Ca^{2+}$  que tiene un orden de magnitud más bajo. El  $Pb^{2+}$  no tiene eficiencia en el rompimiento. La tasa de rompimiento es mayor conforme mayor sea la afinidad del ión metálico solvente por donar protones, una interpretación de esta correlación es que el hidróxido metálico coordinador puede participar en la catálisis.

Se ha demostrado que un ión metálico debe unirse al oxígeno *pro - R* del grupo que se rompe llevando a cabo una reacción. Al hacer esto, la carga negativa del fosfato se neutraliza y facilita el ataque en el átomo de fósforo. Este ión metálico divalente o posiblemente otro adicional pueden llevar a cabo la estabilización requerida de la carga negativa del grupo  $O5'^-$ .

## 1.7 Origen de los datos estructurales de las moléculas

La mayoría de las estructuras cristalográficas pueden obtenerse por tres métodos: cristalografía por Rayos X, Resonancia Magnética Nuclear (NMR) o Modelado Teórico. Los dos primeros métodos son experimentales. Los resultados empíricos de estos métodos pueden describir de forma aproximada estructuras 3D de una molécula en el estado particular en el cual fue resuelta. La cristalización en ocasiones distorsiona las partes de una estructura debido a los contactos con otras moléculas vecinas dentro del cristal. Sin embargo, los cristales utilizados para difracción están altamente hidratados de tal forma que las estructuras determinadas de los cristales no son muy diferentes de las estructuras de una solución acuosa. Las moléculas que han sido estudiadas tanto por cristalografía de rayos X y por NMR muestran gran concordancia.

La difracción del cristal por rayos X no puede resolver las posiciones de los átomos de hidrógeno o distinguir con precisión entre un nitrógeno, un oxígeno o un carbono. Esto significa que la identidad de un átomo es sugerida por



el cristalógrafo con base al ambiente químico. Por lo que en algunas ocasiones existe incertidumbre acerca de si el oxígeno de una molécula de agua o un ión metálico está ligado a la estructura. Algunas nuevas estructuras cristalográficas de rayos X ya incluyen los hidrógenos, sin embargo, estos se colocaron utilizando modelado molecular. Solo para un número muy limitado existe una resolución de 1.2 Å que les permita localizar las posiciones de los hidrógenos.

Antes de que se determine una estructura 3D por difracción de rayos X en el cristal, se deben realizar una gran variedad de experimentos. En primer lugar, la molécula debe ser cristalizada y los cristales deben ser singulares y de calidad casi perfecta. Aquí es donde fallan la mayoría de los experimentos en determinar la estructura 3D. Muchas moléculas importantes no tienen estructura cristalográfica debido a inconvenientes para obtener su forma cristalina.

Una vez que se tiene el cristal, este es irradiado con rayos X para obtener un patrón de difracción. Este patrón consiste de miles de puntos localizados en una placa. La posición y la intensidad de cada punto es relativamente fácil de localizar, pero también se deben determinar las fases de las ondas que dieron lugar a cada punto con el fin de determinar el mapa de densidad electrónica.

Después de resolver el problema de fases, algunos derivados de este mismo cristal se irradian en presencia de iones metálicos pesados. Este método requiere que los iones metálicos sean introducidos al cristal sin cambiar su estructura.

Los Rayos X son difractados por los electrones de las moléculas en el cristal, por lo que se obtienen con esto imágenes planas (mapa de densidad electrónica) de las nubes electrónicas tridimensionales de las moléculas. Uno interpreta esta imagen al construir un modelo de la molécula que se ajuste a este mapa. Una vez que se hace esto se obtienen las coordenadas cartesianas de todos los átomos de la molécula con excepción de los hidrógenos.

Los datos estructurales de la mayoría de las moléculas se pueden encontrar en Bases de Datos de acceso Mundial como son "Protein Data Bank" (PDB) para todas las moléculas de importancia biológica y en "Nucleic Acid Data Bank" (NDB) para ácidos nucleicos. La forma como funciona esta base y los archivos que origina, los archivos PDB, se discuten en la siguiente sección.

### 1.7.1 Bases de datos de estructuras moleculares

Las principales bases de datos son[27, 28]:

<b>PDB</b>	Protein Data Bank,
<b>CSD</b>	Cambridge Structural Database
<b>NDB</b>	Nucleic Acid Database

El formato en que se puede encontrar es PDB, este tipo de formato proporciona información que se agrupa en dos partes: la información *primaria* que incluye coordenadas atómicas, datos cristalográficos, referencias bibliográficas, condiciones de cristalización, método y descripciones estructurales. La siguiente parte es la información *derivada* que se calcula de las coordenadas atómicas e incluye longitud de los enlaces y ángulos, ángulos torsionales, longitudes de enlaces virtuales y ángulos de los átomos de fósforo.

Por ejemplo, un renglón que contiene las coordenadas atómicas para el fósforo de la citosina con clasificación  $C_{170}$  es:

Átomo	No.	Tipo	Mol	X	Y	Z	Datos cristalográficos	
ATOM	751	P	C B	33.402	4.414	29.238	1.00	30.54 P

En donde el tipo de átomo se da de acuerdo al nombre y la posición del átomo en la molécula según la IUPAC.

### 1.7.2 Ribozima “Hammerhead” cristalográfico

La estructura cristalográfica utilizada fue la ribozima “Hammerhead” que contiene iones de magnesio. Se obtuvo del Banco de Datos de Ácidos nucleicos del “Laboratorio Nacional de Brookhaven (Nucleic Acid Data Bank, NDB)” con clave:301D obtenida por W. G. Scott et al [14]. La molécula contiene 873 átomos (no incluye hidrógenos) más 5 átomos de magnesio en los sitios mostrados en la figura 1.18. La resolución que obtuvieron fué de 3.00 Å.



Figura 1.18: Ribozima "Hammerhead" cristalográfico

La secuencia de residuos o de nucleótidos que contiene la molécula ribozima "Hammerhead" está formada por 40 nucleótidos divididos en dos cadenas a y b:

**Cadena a**

Son 16 nucleótidos y es el sustrato formado por la secuencia

G U G G U C U G A U G A G G C C

**Cadena b**

Son 24 nucleótidos y es la enzima formada por la secuencia

G G C C G A A A C U C G U A A G A G U C A C C A C

La región de interés se encuentra en dos nucleótidos la  $C_{17}$  y la  $A_{1,1}$  de acuerdo a la nomenclatura de la figura 1.13. En la figura 1.19 se puede apreciar un acercamiento a esta zona con la nomenclatura asignada de acuerdo al tipo de átomo por CHARMM (programa computacional) y "Protein Data Bank". En esta misma figura se puede ver una posición tentativa del átomo de magnesio.

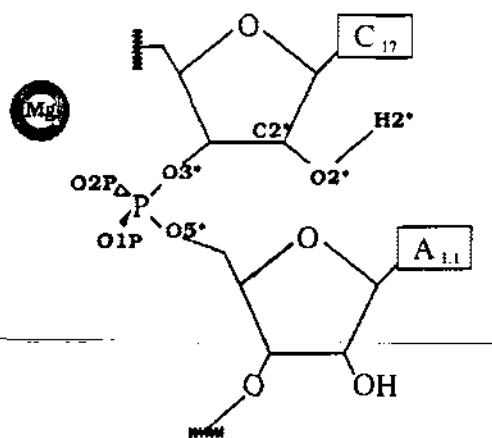


Figura 1.19: Esquema de los nucleótidos involucrados en la catálisis.

## 1.8 Planteamiento del problema

En la figura 1.19 vemos la región de interés catalítico que consta de sólo dos nucleótidos. Sin embargo, aún no se sabe cómo se da el rompimiento de esta molécula. Existen varias dudas con respecto al mecanismo de reacción y cómo se presenta. A continuación se hace un resumen sobre cuáles son las incógnitas que hay sobre esta biomolécula y las posibles hipótesis que se hacen al respecto con base en evidencias experimentales.

1. Recientemente se determinó la estructura cristalográfica del "Hammerhead". Esta resolución es de 3 Å la cual es muy pobre para determinar coordenadas atómicas y formación de puentes de hidrógeno, además de que no se ha observado la transición de los átomos involucrados hacia una conformación previa al rompimiento con las técnicas cristalográficas utilizadas [10, 14, 15].
2. Aún cuando se sabe que sólo influyen dos nucleótidos, no se sabe si es necesario que sea el par guanina-adenina, ni tampoco la influencia de su conformación. Es decir, la importancia de una secuencia determinada para el rompimiento. Se ha observado que al cambiar ciertas bases en el "Hammerhead", se inhibe el rompimiento. Estas bases no necesariamente están cerca del área de rompimiento, sino que se encuentran en la "bolsa" de la estructura terciaria [7].

3. El oxígeno  $O2'$  de la  $C_{17}$ , tiene un papel muy importante, ya que al sustituir la ribosa del ARN por desoxirribosa u otro grupo no existe rompimiento. Así mismo, el oxígeno pro-R del Fosfato despliega una función importante, ya que al ser sustituido por un azufre, no se da tampoco el rompimiento [10, 17].
4. No se sabe el número de magnesios que participan en la catálisis. Existen dos mecanismos propuestos que involucran a uno y dos metales, sin embargo, no se ha determinado con exactitud cuál es el correcto. Se ha propuesto que uno de los magnesios debe estar ligado al oxígeno pro-R del Fosfato  $C_{17}$ , pero esto ha sido determinado por métodos indirectos, ya que el magnesio no se puede observar en la cristalografía por rayos X por ser de baja densidad electrónica [15, 17].

La forma como este problema se va a atacar es a través de Dinámica Molecular, la cual permitirá conocer dentro de un cierto margen de aproximación el mecanismo de reacción que efectúa una ribozima al romper una molécula. Esta técnica es una simulación utilizando equipo computacional y es descrita en el siguiente capítulo.

## Capítulo 2

# Dinámica Molecular

### Introducción

En este trabajo vamos a estudiar a una macromolécula de importancia biológica vista como un sistema de muchas partículas en movimiento que interaccionan entre sí. Como no existe todavía un método para resolver de forma exacta el movimiento de muchos cuerpos utilizaremos un método que resuelve las ecuaciones de la mecánica clásica de forma numérica y estadística. Uno de los métodos que emplea esta técnica es la Dinámica Molecular (DM).

La DM es un método utilizado para modelar los movimientos de un sistema de partículas. Esta basado en la mecánica newtoniana, ideas de dinámica no lineal, teoría cinética, teoría de muestreo y métodos numéricos [29]. Se puede aplicar tanto a sistemas tan pequeños como un átomo o una molécula diatómica en una reacción, como a una galaxia. En todos los casos, los elementos esenciales de una DM son el conocimiento del potencial de interacción de las partículas, del cual se derivan las fuerzas y con esto las ecuaciones de movimiento que gobiernan la dinámica de las partículas [31]. Las ecuaciones así obtenidas con mecánica newtoniana son adecuadas para muchos sistemas, incluyendo las biomoléculas, sin embargo, para algunos problemas como reacciones, las correcciones cuánticas son muy importantes.

Las simulaciones dinámicas tienen dos características que le han permitido tener un papel esencial en el desarrollo de diferentes técnicas y en la aplicación a diferentes áreas de la ciencia. En primer lugar, las simulaciones proporcionan los movimientos de las partículas como función del tiempo de

tal forma que pueden ser probadas en intervalos de tiempo mas largos que los experimentos para contestar a las preguntas detalladas sobre las propiedades del sistema. En segundo lugar, aunque el potencial utilizado es aproximado, este potencial esta bajo el control del usuario, lo que le permite alterar cualquier contribución y así determinar su peso en cierta propiedad que se quiera evaluar del sistema. Otros métodos como la simulación Monte Carlo no dan cuenta de dependencias temporales y no permiten establecer una secuencia entre los estados del sistema.

---

En el método DM la fuerza que se ejerce sobre un átomo dado depende de las posiciones de los átomos de alrededor y estas fuerzas rigen la evolución temporal del sistema. En las simulaciones por computadora de DM se fija la posición inicial y la velocidad de cada átomo y se integran las ecuaciones de movimiento clásicas (newtonianas) para determinar un nuevo conjunto de posiciones y velocidades para todos los átomos del sistema. La integración puede llevarse a cabo manteniendo diferentes variables físicas en forma constante tales como el número de partículas, la energía, temperatura, presión o volumen [37]. Suelen usarse diversos algoritmos para la integración de las ecuaciones de movimiento. Dentro de la DM también se pueden contemplar constricciones (holonómicas)<sup>1</sup> que se introducen mediante el método de Lagrange.

En todos los métodos de mecánica molecular se requiere la introducción de una función de energía potencial  $U$  necesaria para describir las interacciones que se dan entre las partículas que constituyen al sistema. En el caso de moléculas de importancia biológica la función de energía potencial es complicada y por lo mismo, tiene un origen principalmente empírico. La energía así obtenida es generalmente dependiente de las coordenadas atómicas y recibe el nombre de **Campo de Fuerza**. Una vez que se cuenta con la energía potencial  $U$  se puede conocer la función de fuerza  $F$  al obtener su gradiente espacial. Esta fuerza puede utilizarse para obtener el comportamiento dinámico del sistema al resolver las ecuaciones de Newton para el átomo como función del tiempo.

A continuación veremos como se establecen las ecuaciones de Newton que después utilizaremos en la solución de la dinámica del sistema de partículas.

---

<sup>1</sup>Una constricción es algo que limita el movimiento y si estas constricciones se pueden expresar como ecuaciones que conectan a las coordenadas de las partículas (y el tiempo) con la forma  $f(r_1, r_2, \dots, t) = 0$  entonces se dice que las constricciones son holonómicas [32].

## 2.1 Ecuaciones de movimiento y teoremas de conservación

El modelo microscópico más rudimentario de una sustancia se basa en partículas esféricas rígidas para simular átomos. Para el caso de estos sistemas simples se utilizan las ecuaciones de Newton, pero conforme las partículas adquieren estructuras más complejas se requiere una formulación más adecuada para las ecuaciones de movimiento. La formulación Lagrangiana de la mecánica clásica da una base general para resolver problemas complicados, aunque también se puede hacer utilizando la formulación de Hamilton.

En la interpretación newtoniana de la mecánica, el movimiento traslacional de una molécula esférica  $i$  es causada por una fuerza  $F_i$  [36]. Esta fuerza es ejercida por un agente externo. El movimiento y la fuerza aplicada están relacionados explícitamente por la segunda ley de Newton:

$$F_i = m\ddot{r}_i \quad (2.1)$$

Aquí  $m$  es la masa de la molécula. Se asume que la masa es independiente de la posición, velocidad y tiempo (no se contemplan efectos relativistas). La aceleración está dada por:

$$\ddot{r}_i = \frac{d^2 r_i}{dt^2}, \quad (2.2)$$

donde  $r_i$  es un vector que localiza a la molécula con respecto a un sistema de ejes coordenados fijos. Una molécula inicialmente en reposo ó moviéndose con velocidad uniforme continuará en su estado de movimiento hasta que una fuerza externa actúe sobre esta molécula. Esta es la primera ley de Newton.

La segunda ley (2.1) puede utilizarse también para obtener la tercera ley de Newton. Consideremos un sistema aislado que contiene dos moléculas esféricas, **1** y **2**. Por definición, un sistema aislado no tiene fuerzas externas, por lo que la fuerza total es cero

$$F_{total} = 0. \quad (2.3)$$

Por lo tanto cualquier fuerza ejercida por la molécula **1** sobre la molécula **2** debe estar balanceada por una fuerza ejercida por **2** sobre **1**

$$F_{total} = F_{12} + F_{21} = 0.$$



Esto implica

$$\mathbf{F}_{12} = -\mathbf{F}_{21}, \quad (2.4)$$

y esta es la tercera ley de Newton. Con estas variables (posición, velocidad y aceleración) se define la energía cinética como

$$E_k = \frac{1}{2}m\dot{r}^2, \quad (2.5)$$

donde  $\dot{r}$  representa la velocidad de la partícula. Así mismo, el momento  $p_i$  de una molécula  $i$  se define como el producto de la masa por la velocidad:

$$p_i = m\dot{r}_i. \quad (2.6)$$

Utilizando la segunda ley de Newton tenemos que la fuerza está relacionada con la razón de cambio temporal del momentum y se puede expresar como:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(m\mathbf{v}). \quad (2.7)$$

El primer teorema de conservación se refiere al **momento lineal** de una partícula. Si se tiene una partícula libre, es decir, una partícula sobre la que no actúa ninguna fuerza, entonces por la ecuación 2.7 tenemos que  $\dot{\mathbf{p}} = 0$ . Por lo tanto,  $\mathbf{p}$  es un vector constante en el tiempo y el teorema de conservación dice:

- I. *Cuando la fuerza total sobre una partícula es cero entonces el momento lineal  $\mathbf{p}$  total de la partícula se conserva.*

Se define el **momento angular**  $\mathbf{L}$  de una partícula en la posición  $\mathbf{r}$  con respecto a un origen como:

$$\mathbf{L} \equiv \mathbf{r} \times \mathbf{p} \quad (2.8)$$

y la **torca**  $\mathbf{N}$  con respecto al mismo origen como:

$$\mathbf{N} \equiv \mathbf{r} \times \mathbf{F}, \quad (2.9)$$

donde  $\mathbf{r}$  es el vector de posición del origen al punto donde la fuerza  $\mathbf{F}$  es aplicada. Como  $\mathbf{F} = m\dot{\mathbf{v}}$ , la torca se puede expresar alternativamente como

$$\mathbf{N} = \mathbf{r} \times m\dot{\mathbf{v}} = \mathbf{r} \times \dot{\mathbf{p}}. \quad (2.10)$$

Si derivamos el momento angular con respecto al tiempo

$$\dot{\mathbf{L}} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = (\dot{\mathbf{r}} \times \mathbf{p}) + (\mathbf{r} \times \dot{\mathbf{p}}) \quad (2.11)$$

y usamos

$$\dot{\mathbf{r}} \times \mathbf{p} = \dot{\mathbf{r}} \times m\mathbf{v} = m(\dot{\mathbf{r}} \times \dot{\mathbf{r}}) \equiv 0 \quad (2.12)$$

entonces el primer miembro de (2.11) se hace cero y llegamos a que

$$\dot{\mathbf{L}} = \mathbf{r} \times \dot{\mathbf{p}} = \mathbf{N}. \quad (2.13)$$

Si ninguna torca actúa sobre una partícula (es decir, si  $\mathbf{N} = 0$ , entonces  $\dot{\mathbf{L}} = 0$  y  $\mathbf{L}$  es un vector constante en el tiempo. El segundo teorema importante de conservación dice:

II. *Si una partícula no está sujeta a una torca, entonces el momento angular de tal partícula se conserva.*

El trabajo que se aplica a una partícula para pasarlo de la posición **1** a la posición **2** por una fuerza  $\mathbf{F}$  esta definido como:

$$W_{12} \equiv \int_1^2 \mathbf{F} \cdot d\mathbf{r}. \quad (2.14)$$

Si  $\mathbf{F}$  es la fuerza resultante actuando sobre la partícula,

$$\begin{aligned} \mathbf{F} \cdot d\mathbf{r} &= m \frac{d\mathbf{v}}{dt} \cdot \frac{d\mathbf{r}}{dt} dt = m \frac{d\mathbf{v}}{dt} \cdot \mathbf{v} dt = \frac{m}{2} \frac{d}{dt} (\mathbf{v} \cdot \mathbf{v}) dt \\ \mathbf{F} \cdot d\mathbf{r} &= \frac{m}{2} \frac{d}{dt} (v^2) dt = d \left( \frac{1}{2} m v^2 \right). \end{aligned} \quad (2.15)$$

Por lo que la integral (2.14) es una diferencial exacta y el trabajo hecho por la fuerza total  $\mathbf{F}$  es igual a la diferencia de energía cinética:

$$W_{12} = E_{k2} - E_{k1}. \quad (2.16)$$

Si el campo de fuerza es tal que el trabajo hecho a lo largo de una órbita cerrada es cero, es decir,

$$\oint \mathbf{F} \cdot d\mathbf{r} = 0 \quad (2.17)$$

entonces se dice que la fuerza y el sistema son conservativos [36]. Físicamente un sistema no puede ser conservativo si están presentes efectos de fricción u otra fuerza disipativa. Esto es debido a que la integral anterior no se puede anular. Por el teorema de Stokes<sup>2</sup>, la condición de fuerzas conservativas se puede escribir

$$\nabla \times \mathbf{F} = 0 \quad (2.18)$$

y como el rotacional de un gradiente es siempre cero,  $\mathbf{F}$  debe ser el gradiente de algún escalar

$$\mathbf{F} = -\nabla V, \quad (2.19)$$

donde  $V$  es el llamado potencial o energía potencial.

La existencia de  $V$  se puede establecer sin el uso de teoremas de cálculo vectorial. Si la ecuación (2.17) es válida entonces el trabajo debe ser independiente de la trayectoria de integración entre dos puntos 1 y 2. Se sigue entonces que debe ser posible expresar  $W_{12}$  como el cambio en una cantidad que depende sólo de las posiciones de los puntos finales. Esta cantidad puede designarse por  $-V$ , entonces para una diferencia en la trayectoria de integración se tiene la relación

$$\mathbf{F} \cdot d\mathbf{r} \quad (2.20)$$

ó

$$\mathbf{F}_r = -\frac{\partial V}{\partial r} \quad (2.21)$$

que es equivalente a la ecuación (2.19). Nótese que en (2.19) se le puede sumar a  $V$  una cantidad constante en el espacio sin afectar el resultado. Por lo que el nivel cero es arbitrario. Para un sistema conservativo el trabajo ejercido por una fuerza es

$$W_{12} = V_1 - V_2 \quad (2.22)$$

combinando (2.22) con (2.16) se tiene

$$E_{k1} + V_1 = E_{k2} + V_2. \quad (2.23)$$

Con estas condiciones tenemos el tercer teorema de conservación:

III. *La energía total  $E$  de una partícula en un campo de fuerza conservativo se conserva en el tiempo.*

---

<sup>2</sup>El Teorema de Stokes establece la relación entre la integral de línea de una función con la integral de superficie de la derivada de esa función:  $\oint F \cdot d\mathbf{r} = \int_S \nabla \times F$  [50].

## 2.2 Métodos de integración

Resolver la ecuación de movimiento de Newton implica resolver una ecuación diferencial cuya variable es el tiempo. Generalmente esta ecuación no es exacta y uno de los métodos utilizados para resolverla es mediante **integración numérica**. El algoritmo de integración de un programa de DM es una de las bases principales de los cálculos.

Los algoritmos de integración están basados en métodos de diferencias finitas donde se toma un intervalo de tiempo sobre una malla finita. El intervalo de tiempo  $\Delta t$  es la separación entre dos puntos consecutivos en la malla.

A partir de las posiciones y algunas de sus derivadas temporales al tiempo  $t$ , el algoritmo de integración proporciona las cantidades para un tiempo posterior  $t + \Delta t$ . Repitiendo este procedimiento se puede seguir la evolución del sistema para largos periodos de tiempo.

Todos estos métodos son aproximados y existen errores asociados a ellos. Se pueden distinguir dos tipos principales de errores[39]:

1. **Errores de truncación:** los métodos de diferencias finitas están basados principalmente en una truncación de la expansión de Taylor en algún término. Estos errores son intrínsecos al algoritmo usado.
2. **Errores de redondeo:** están relacionados a una implementación particular del algoritmo, por ejemplo, el número de dígitos utilizados en la aritmética computacional.

Estos dos tipos de errores pueden reducirse al tomar valores más pequeños de  $\Delta t$ . Para valores grandes de  $\Delta t$  dominan los errores de truncación y decrecen rápidamente conforme  $\Delta t$  disminuye. Por ejemplo para el algoritmo de Verlet, discutido posteriormente, el error de truncación es proporcional a  $\Delta t^4$  para cada intervalo de integración. En cambio, los errores de redondeo decrecen más despacio conforme  $\Delta t$  disminuye, la precisión en este caso puede aumentar si se utiliza mayor número de dígitos, lo que significa mayor trabajo computacional.

Los esquemas más utilizados son conocidos como el método de **Verlet** y el método **predecir-correr**. Estos dos métodos se presentan a continuación incluyendo el método de "**Leap-Frog**" que es una forma modificada del algoritmo de Verlet.

Estos métodos están basados en la expansión de Taylor de la posición al tiempo  $t + \Delta t$

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{a(t)\Delta t^2}{2} + \frac{b(t)}{6}\Delta t^3 + \dots \quad (2.31)$$

donde  $v(t)$  es la primera derivada de la posición  $r(t)$ ,  $a(t)$  es la segunda derivada de la posición,  $b(t)$  es la tercera derivada, etc. Como la integración se realiza en forma discreta, es conveniente reescribir la expansión anterior en términos de las variables de paso  $n$  y  $n + 1$

$$r_{n+1} = r_n + v_n\Delta t + \frac{a_n\Delta t^2}{2} + \frac{b_n}{6}\Delta t^3 + O(\Delta t^4). \quad (2.32)$$

Los diferentes algoritmos de integración variarán de acuerdo a la forma cómo se implementa ésta expansión básica.

### 2.2.1 Integrador de Verlet

Es el integrador utilizado con más frecuencia, se basa en dos expansiones de Taylor de tercer orden para la posición  $r(t)$ , una hacia adelante y la otra hacia atrás con respecto al tiempo

$$r_{n+1} = r_n + v_n\Delta t + \frac{a_n}{2}\Delta t^2 + \frac{b_n}{6}\Delta t^3 + O(\Delta t^4) \quad (2.33)$$

$$r_{n-1} = r_n - v_n\Delta t + \frac{a_n}{2}\Delta t^2 - \frac{b_n}{6}\Delta t^3 + O(\Delta t^4). \quad (2.34)$$

Las cuales al sumarse dan origen al formalismo básico de la Integración de Verlet [30, 33, 39, 40]:

$$r_{n+1} = 2r_n - r_{n-1} + a_n\Delta t^2 + O(\Delta t^4). \quad (2.35)$$

Como se esta integrando la ecuación de Newton,  $a(t)$  es la fuerza dividida por la masa y la fuerza es a la vez una función de la posición  $\mathbf{r}(t)$ :

$$\mathbf{a}(t) = -\frac{1}{m}\nabla U(\mathbf{r}(t)). \quad (2.36)$$

Como se puede ver, el error de truncación del algoritmo es del orden de  $\Delta t^4$  aunque la derivada de tercer orden no aparezca explícitamente. Este algoritmo es a la vez simple de implementar, da una buena aproximación y es estable en la dinámica.

Surge un problema con esta versión del algoritmo de Verlet y es que las velocidades no se generan de manera directa. Las velocidades son necesarias para obtener la energía cinética  $E_k$  y la evaluación de la energía cinética es necesaria para comprobar la conservación de la energía total  $E = E_k + V$ . Esta prueba es una de las más importantes para verificar que una simulación DM está procediendo correctamente. Las velocidades se pueden computar a partir de las posiciones utilizando

$$v_n = \frac{r_{n+1} - r_{n-1}}{2 dt} . \quad (2.37)$$

Sin embargo, el error asociado a esta expresión es del orden de  $\Delta t^2$  en lugar de  $\Delta t^4$ .

### Ventajas

1. La integración no requiere las velocidades, sólo da la información de la posición.
2. Sólo hay una evaluación de la fuerza por ciclo de integración. (La evaluación de la fuerza es la parte más costosa computacionalmente de la simulación).
3. La formulación está basada en expansiones temporales hacia adelante y hacia atrás, y es naturalmente reversible en el tiempo (una propiedad de la ecuación de movimiento).

### Desventajas

1. Introducción de errores numéricos grandes. Un término par  $O(\Delta t^2)$  es sumado a un término  $O(\Delta t^0)$ .
2. Las velocidades son calculadas en forma aproximada.

### 2.2.2 Integrador de velocidad de Verlet

Este esquema está diseñado para mejorar las evaluaciones de la velocidad. El algoritmo es el siguiente:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + a(t)\frac{\Delta t^2}{2} \quad (2.38)$$

$$v\left(t + \frac{\Delta t}{2}\right) = v(t) + \frac{a(t)}{2}\Delta t \quad (2.39)$$

$$a(t + \Delta t) = -\left(\frac{1}{m}\right)\nabla(U(r(t + \Delta t))) \quad (2.40)$$

$$v(t + \Delta t) = v\left(t + \frac{\Delta t}{2}\right) + \frac{a(t + \Delta t)\Delta t}{2} \quad (2.41)$$

Esto significa que cada ciclo de integración involucra cuatro pasos:

1. Calcular  $r(t + dt)$
2. Calcular la velocidad en el punto medio del intervalo  $dt$ :
 
$$v\left(t + \frac{dt}{2}\right) = v(t) + a(t)\frac{dt}{2}$$
3. Calcular  $a(t + dt)$
4. Calcular  $v(t + dt) = v\left(t + \frac{dt}{2}\right) + a(t + dt)\frac{dt}{2}$

#### Ventajas

Tiene una de las mejores evaluaciones de las velocidades entre los algoritmos de diferencias finitas.

#### Desventajas

Computacionalmente es más caro que el algoritmo de Verlet o "Leap-Frog".

### 2.2.3 Integrador "Leap-Frog"

El integrador "Leap-Frog" es una variación del algoritmo de Verlet diseñado para mejorar la evaluación de la velocidad. Su nombre viene del hecho de que las velocidades son evaluadas en el punto medio de las posiciones y viceversa.

El algoritmo es el siguiente:

$$v\left(t + \frac{\Delta t}{2}\right) = v\left(t - \frac{\Delta t}{2}\right) + a(t)\Delta t \quad (2.42)$$

$$r(t + \Delta t) = r(t) + v(t + \frac{\Delta t}{2})\Delta t . \quad (2.43)$$

Esto significa que cada ciclo de integración involucra tres pasos:

1. Calcular  $a(t)dt$  (basado en  $r(t)$ )
2. Calcular  $v(t + \frac{dt}{2})$
3. Calcular  $r(t + dt)$

La velocidad instantánea en un tiempo  $t$  es entonces calculada como

$$v(t) = v(t + \frac{\Delta t}{2}) + \frac{v(t - \frac{\Delta t}{2}) + a(t)\Delta t}{2} . \quad (2.44)$$

### Ventajas

1. Mejora la evaluación de velocidades.
2. La evaluación directa de las velocidades es útil para controlar la temperatura de la simulación y otras variables dependientes de la velocidad.
3. Reduce el problema de error numérico del algoritmo de Verlet. Aquí los términos  $O(\Delta t^1)$  son sumados a términos  $O(\Delta t^0)$ .

### Desventajas

1. Las velocidades al tiempo  $t$  son todavía aproximaciones.
2. Involucra más pasos computacionales que el algoritmo de Verlet.

## 2.2.4 Métodos de predecir-correr

Los métodos *predecir – correr* están compuestos de tres pasos: predicción, evaluación y corrección. Si por ejemplo se tiene la posición  $r(t)$  y la velocidad  $v(t)$ , los pasos a seguir son los siguientes:

1. Predecir la posición  $r(t + \Delta t)$  y la velocidad  $v(t + \Delta t)$  al final del siguiente paso.
2. Evaluar las fuerzas en  $t + \Delta t$  usando la posición predicha.
3. Corregir las predicciones usando alguna combinación de lo predicho y los valores previos de la posición y la velocidad.



## 2.3 Condiciones de frontera

Se puede simular una gran variedad de condiciones experimentales con la dinámica molecular. Las primeras simulaciones de sistemas macromoleculares se realizaron sin ningún solvente en sus alrededores. Estas simulaciones fueron modificadas después para incluir agua en la vecindad de moléculas de proteína, tal y como se puede encontrar en un cristal. Para sistemas homogéneos simples, tales como una caja con moléculas de agua, se utilizan condiciones de frontera periódicas. Estas condiciones se basan en la representación de un sistema a partir de copias de un modelo pequeño que se repite periódicamente en todas las direcciones. En el caso más común se utiliza una geometría cúbica, pero para sistemas inhomogéneos como proteínas, se requieren otros tipos de geometría como arreglos tetraédricos.

## 2.4 Ensamblés

El método de Dinámica molecular descrito anteriormente es un esquema para estudiar la evolución temporal de un sistema clásico conformado por  $N$  partículas en un volumen  $V$ , donde la energía  $E$  es una constante de movimiento. Esto implica que la trayectoria de la simulación se mueve en una hipersuperficie de energía total constante en el espacio fase. Asumiendo que los promedios temporales son equivalentes a promedios sobre los ensambles, los promedios obtenidos en una simulación DM convencional corresponden a promedios sobre el ensamble microcanónico (NVE), donde se mantienen constantes las variables  $N$ ,  $V$  y  $E$  [29].

Sin embargo, esta formulación de la DM tiene dos pequeños inconvenientes. En primer lugar, las simulaciones de fluidos o soluciones son llevadas a cabo a volumen constante, mientras que las fluctuaciones de volumen son críticas para un gran número de fenómenos. En segundo lugar, la integración numérica de las ecuaciones de movimiento producen una trayectoria en la cual la energía del sistema se conserva, mientras que en los sistemas reales existe intercambio de energía con los alrededores.

Por lo que en ocasiones es más conveniente realizar la simulación DM en otro tipo de ensambles. La lista de ensambles más relevantes en DM se muestra en la Tabla 2.1. En el caso de los sistemas biológicos el ensamble más indicado es el NTP [26]. Para adaptar la DM a otros ensambles existen varios métodos, uno

de ellos se puede realizar al acoplar la celda de simulación a un baño térmico y/o a un pistón para controlar la temperatura y/o la presión, respectivamente.

*Tabla 2.1 DM en varios ensambles*

Condición	Constantes durante la simulación <sup>3</sup>	Ensamble
Dinámica libre	$N, V, E$	<i>Microcanónico</i>
Temperatura Constante	$N, T, V$	<i>Canónico</i>
Presión Constante	$N, P, H$	<i>Isobárico-Isoentálpico</i>
Temperatura y Presión constante	$N, T, P$	<i>Isotérmico-Isobárico</i>

<sup>3</sup>N número de partículas, E energía, T temperatura, P presión, V volumen, H entalpía.

A continuación se describe el tipo de modificaciones que se deben realizar al algoritmo dinámico para permitir simular otro tipo de ensambles[37].

### 2.4.1 Dinámica molecular a temperatura constante

Existen varios métodos DM a temperatura constante. El método más sencillo se basa en un reescalamiento de las velocidades a cada paso de la integración para mantener la energía cinética total en un valor constante.

Tenemos que la temperatura está relacionada con la Energía cinética a través de la ecuación:

$$\langle E_k \rangle = \frac{3}{2} N k_B T. \quad (2.45)$$

En donde  $k_B$  es la constante de Boltzmann. Si la temperatura en el tiempo  $t$  es  $T(t)$ , las velocidades son reescaladas o multiplicadas por un factor  $\lambda$  de tal forma que el cambio de temperatura asociado  $\Delta T = T_{req} - t(t)$  se puede calcular de la siguiente manera[41]:

$$\Delta T = \frac{1}{2} \sum_{i=1}^N \frac{2 m_i (\lambda v_i)^2}{3 N k_B} - \frac{1}{2} \sum_{i=1}^N \frac{2 m_i v_i^2}{3 N k_B}. \quad (2.46)$$

A partir de esta expresión obtenemos:

$$\Delta T = \frac{1}{2} \sum_{i=1}^N \frac{2}{3} \frac{m_i v_i^2}{N k_B} (\lambda^2 - 1) \quad (2.46)$$

pero sabemos que  $T(t) = \frac{1}{2} \sum_{i=1}^N \frac{2}{3} \frac{m_i v_i^2}{N k_B}$ , por lo que

$$\Delta T = (\lambda^2 - 1)T(t) \quad (2.47)$$

despejando finalmente el valor de  $\lambda$  tenemos

$$T_{req} - T(t) = (\lambda^2 - 1)T(t)$$

$$\lambda = \sqrt{\frac{T_{req}}{T(t)}} \quad (2.48)$$

donde  $T_{req}$  es la temperatura requerida o deseada. Por lo que el método más simple para controlar la temperatura es multiplicar las velocidades en cada paso de integración por un factor  $\lambda = \sqrt{(T_{req}/T_{act})}$ , donde  $T_{act}$  es la temperatura actual calculada de la energía cinética.

## 2.4.2 Dinámica molecular a presión constante

En una simulación de DM con condiciones periódicas a la frontera el volumen se encuentra fijo, mientras que a presión constante el volumen del sistema puede fluctuar. Para incluir el volumen como una variable dinámica, las ecuaciones de movimiento se determinan mediante un análisis del sistema en el cual la posición y el momento de todas las partículas son reescaladas por un factor proporcional a la raíz cúbica del volumen del sistema. Si denotamos al volumen como  $V$ , las nuevas ecuaciones del movimiento para cada átomo real  $i$  serán[37]:

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m_i} + \mathcal{V} \frac{\mathbf{r}_i}{3} \quad (2.49)$$

$$\dot{\mathbf{p}}_i = \sum_j \mathbf{F}_{ij} - \frac{\mathcal{V} \mathbf{p}_i}{3} \quad (2.50)$$

$$M\ddot{V} = \alpha + \frac{1}{3V} \left[ \sum_j (\mathbf{p}_i \cdot \mathbf{p}_j) - \sum_{ij} \mathbf{r}_{ij} U'(\mathbf{r}_{ij}) \right] \quad (2.51)$$

donde  $U'$  representa la derivada del potencial con respecto a la posición del átomo  $i$  y

$$\mathcal{V} = \frac{d}{dt}(\ln V). \quad (2.52)$$

Y tenemos que  $\alpha$  y  $M$ , son dos parámetros que deben obtenerse por fuera de la simulación. La primera es la presión a la cual se va a desarrollar la dinámica, mientras que  $M$  puede interpretarse como la masa del pistón que aplica la presión  $\alpha$ .

En la segunda ecuación la suma es sobre todos los valores que  $i \neq j$ , y la convención de imagen mínima es como sigue: para un valor dado de  $j$ , la interacción entre la pareja  $ij$  solo se considera entre el átomo  $i$  y cualquier átomo cercano a este, el átomo real  $j$  o la imagen más cercana de  $j$ . Ambas sumas de la tercera ecuación se toman solo sobre átomos reales, y la segunda suma solo se calcula para valores tales que  $j > i$ . Esto garantiza que las fluctuaciones de volumen reflejen la variación debida a la presión externa  $\alpha$  y la presión que surge de las interacciones atómicas dentro del volumen. En cada paso de la dinámica, las distancias para las celdas repetidas de las condiciones periódicas son reescaladas de tal forma que la presión tiene su valor correcto.

Los promedios de las trayectorias calculadas de la simulación serán independientes de  $M$  y se puede utilizar cualquier valor siempre y cuando  $M$  sea finita y positiva. Sin embargo, si se quiere realizar una aproximación precisa de las fluctuaciones del sistema, es importante escoger un valor con significado físico para  $M$ . Este valor se puede determinar por ensayo y error siguiendo la prescripción de que  $M$  debe ser escogida de tal forma que la escala temporal de las fluctuaciones en el volumen sea burdamente igual a  $V^{2/3}$  dividido por la velocidad del sonido en el fluido [37].

A partir de las coordenadas y velocidades instantáneas se calculan las propiedades como energía potencial, cinética y total, temperatura, presión, etc., que posteriormente se promedian en el tiempo para obtener el valor de expectación correspondiente. En general, estas simulaciones se usan para estudiar sistemas cercanos al equilibrio.

## 2.5 Variables estadísticas

Las cantidades que se pueden medir en Dinámica Molecular son promedios temporales de propiedades físicas sobre la trayectoria del sistema. Estas propiedades son por lo común funciones de las coordenadas y velocidades de la partícula. Se define el valor instantáneo de una propiedad física  $A$  en el tiempo  $t$  como [33]

$$A(t) = A(r_1(t), \dots, r_N(t); v_1(t), \dots, v_N(t)) \quad (2.53)$$

y entonces se obtiene su promedio

$$\langle A \rangle = \frac{1}{N_t} \sum_{t=1}^{N_t} A(t). \quad (2.54)$$

donde  $t$  es un índice que indica el número de pasos temporales desde 1 hasta  $N_t$ . En la práctica existen dos formas equivalentes para realizar este promedio:

1. El programa de DM calcula  $A(t)$  para cada paso de tiempo mientras está corriendo. La suma  $\sum_t A(t)$  es modificada en cada paso y al final de la corrida se obtiene el promedio al dividir por el número total de pasos. Este es el método preferido cuando la cantidad a medir es simple y/o particularmente importante. Por ejemplo, la temperatura del sistema.
2. Mientras está corriendo el programa, las posiciones (y posiblemente las velocidades) son enviadas periódicamente a un archivo, donde un programa por separado procesa la trayectoria y saca las cantidades deseadas para su monitoreo. Este método puede requerir mucho espacio en disco, sin embargo, se utiliza en ocasiones cuando las cantidades a calcular son complejas o cuando son dependientes de otro parámetro adicional que puede ser desconocido cuando el programa de DM está corriendo.

Las cantidades físicas más comúnmente medidas son las siguientes: energía potencial, energía cinética, energía total, temperatura, curva calórica, dispersión cuadrática media de las posiciones, presión, volumen, etc. Algunas de éstas se explican a continuación.

### Energía Potencial

El promedio de la energía potencial  $V$  se obtiene al promediar sus valores instantáneos. Este cálculo se hace usualmente al mismo tiempo que los cálculos de la fuerza. Por ejemplo, para el caso de un potencial de dos cuerpos

$$V(t) = \sum_i \sum_{j>i} \Phi(|r_i(t) - r_j(t)|) . \quad (2.55)$$

Esta energía se requiere para monitorear la conservación de la energía total.

### Energía Cinética

Otra cantidad necesaria para verificar el principio de conservación de energía es la energía cinética que está dada por

$$\langle E_k \rangle = \frac{1}{2} \sum_i m_i [v_i(t)]^2 . \quad (2.56)$$

### Energía Total

En la dinámica newtoniana la energía total se conserva  $E = E_k + V$ . Sin embargo, es común calcularla en cada paso de tiempo para comprobar que en realidad permanece constante en el tiempo.

En la práctica existen pequeñas fluctuaciones en la energía del orden de  $10^{-4}$  o menos. Estas fluctuaciones son debidas a errores en la integración temporal.

También se pueden observar pequeños cambios en la energía total cuando las corridas de DM son muy largas. Los cambios en la energía pueden deberse a un  $\Delta t$  muy grande. Los brinco temporales provocan más disturbios que las fluctuaciones porque el estado termodinámico del sistema también está cambiando junto con la energía. Estos errores pueden prevenirse si la corrida se parte en corridas más cortas y restaurando en cada una la energía a su valor nominal.

### Temperatura

La temperatura está relacionada directamente con la energía cinética por la fórmula de equipartición <sup>4</sup>, asignando un promedio de energía cinética  $\frac{k_B T}{2}$  por

<sup>4</sup>El principio de equipartición dice que la energía de un gas monoatómico esta distribuida de manera equitativa entre los grados de libertad del movimiento traslacional por una cantidad de  $\frac{1}{2}kT$  por molécula [42].

grado de libertad:

$$E_k = \frac{3}{2} N k_B T . \quad (2.57)$$

Por lo que una estimación de la temperatura se puede obtener directamente del promedio de energía cinética  $E_k$ . También se puede definir la temperatura instantánea por una relación igual a la anterior.

### Curva Calórica

Una vez que son medidas la energía total  $E$  y la temperatura  $T$  en diferentes corridas correspondiendo a diferentes estados termodinámicos, se puede construir la *curva calórica*  $E(T)$ . Esta es una herramienta para determinar el orden de las transiciones de fase, ya que de acuerdo con la clasificación de las transiciones de fase se observa un salto en  $E(T)$  si ésta es de primer orden, o en la derivada de  $E(T)$  si es de segundo orden o mayor.

La transición que más comúnmente se observa en una simulación es la fusión, en la que observa un salto en  $E(T)$  cuando un sólido pierde su estructura cristalina y se convierte en un líquido desordenado. Este salto corresponde al calor latente de fusión.

### Dispersión cuadrática media (DCM)

La dispersión cuadrática media en las posiciones puede calcularse de acuerdo a

$$DCM = \left\langle |r(t) - r(0)|^2 \right\rangle , \quad (2.58)$$

donde lo escrito en corchetes indica promedio sobre todos los átomos.

La DCM contiene información de la difusión atómica, si el sistema es un sólido, la DCM se satura a un valor finito mientras que si el sistema es líquido la DCM crece linealmente con el tiempo.

### Presión

La medición de la presión en una simulación de Dinámica Molecular está basada en la función del virial de Clausius:

$$W(r_1, \dots, r_N) = \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i^{TOT} , \quad (2.59)$$

donde  $\mathbf{F}_i^{TOT}$  es la fuerza total actuando sobre el átomo  $i$ . El promedio estadístico  $\langle W \rangle$  se obtiene sobre un promedio de la trayectoria de DM:

$$\langle W \rangle = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dr \sum_{i=1}^N \mathbf{r}_i(r) \cdot m_i \ddot{\mathbf{r}}_i(r), \quad (2.60)$$

donde se ha utilizado la segunda ley de Newton. Al integrar por partes:

$$\langle W \rangle = - \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dr \sum_{i=1}^N m_i |\dot{\mathbf{r}}_i(r)|^2 \quad (2.61)$$

Esto es dos veces la integral del promedio de la energía cinética, por lo que por la ley de equipartición de la mecánica estadística

$$\langle W \rangle = -DNk_B T, \quad (2.62)$$

donde  $D$  es la dimensión del sistema, (2 o 3),  $N$  es el número de partículas y  $k_B$  la constante de Boltzmann.

Ahora, se puede ver a la fuerza que actúa sobre una partícula como una fuerza compuesta de dos contribuciones

$$\mathbf{F}_i^{TOT} = \mathbf{F}_i + \mathbf{F}_i^{EXT} \quad (2.63)$$

donde  $\mathbf{F}_i$  son las fuerzas internas (que surgen de interacciones atómicas) y  $\mathbf{F}_i^{EXT}$  es la fuerza externa ejercida por las paredes del contenedor. Si las partículas están encerradas en un contenedor paralelepípedo de lado  $L_x$ ,  $L_y$  y  $L_z$  y volumen  $V = L_x L_y L_z$ , y el origen de coordenadas esta en una de sus esquinas, la parte del trabajo externo  $\langle W \rangle$  debido al contenedor puede ser evaluado usando

$$\langle W^{EXT} \rangle = L_x (-PL_y L_z) + L_y (-PL_x L_z) + L_z (-PL_x L_y) = -DPV, \quad (2.64)$$

donde  $-PL_y L_z$  es la fuerza externa  $F_x^{EXT}$  ejercida por la pared  $yz$  a lo largo de la dirección  $x$  sobre las partículas localizadas en  $x = L$ , etc. Reescribiendo la ecuación anterior

$$\left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle - DPV = -DNk_B T$$

o también

$$PV = Nk_B T + \frac{1}{D} \left\langle \sum_i \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle. \quad (2.65)$$

Este resultado se conoce como la *ecuación del virial* y es una forma simple de medir la presión a partir de cantidades obtenidas en la simulación.



## 2.6 Métodos de minimización de la energía

Existen varios algoritmos para la minimización de la energía de un sistema. Todos ellos involucran el cálculo de la segunda derivada de la energía potencial y a partir de esta información se ajustan las coordenadas con el fin de encontrar la menor energía. A continuación se presenta una descripción de los algoritmos más utilizados.

### 2.6.1 Método del descenso más profundo

Este es el método más simple. En cada paso de este proceso iterativo se ajustan las coordenadas en la dirección negativa del gradiente. Tiene solo un parámetro ajustable, el tamaño del paso, que determina la frecuencia con la que ajustan las coordenadas a cada paso. El tamaño del paso depende del valor de la energía. Si la energía disminuye, se incrementa el tamaño del paso por un 20 % para acelerar la convergencia. Si la energía aumenta se disminuye el paso a la mitad. En general este método no converge, pero puede mejorar rápidamente una mala conformación.

### 2.6.2 Técnica del gradiente conjugado

Este método tiene una mejor convergencia que el anterior. Es iterativo y hace uso de la historia previa de los pasos de minimización, así como del gradiente actual para determinar el mínimo de energía en  $N$  pasos para una superficie de energía donde  $N$  es el número de grados de libertad del sistema. Debido a que varios términos del potencial son cuadráticos, se requiere de una menor evaluación de la energía comparado con el método del descenso más profundo.

### 2.6.3 Método Newton-Raphson

Este procedimiento requiere la solución iterativa de las ecuaciones de minimización de Newton-Raphson. El procedimiento aquí es encontrar un punto donde el gradiente se haga cero asumiendo que el potencial es cuadrático. Las ecuaciones de Newton-Raphson pueden resolverse de varias maneras, pero el método que se utilizará (y que utiliza CHARMM) involucra la diagonalización de la segunda derivada en forma matricial y encontrar el tamaño de paso óptimo sobre cada eigenvector. Cuando existe uno o más eigenvalores

negativos, se puede encontrar un punto silla en el potencial. Para evitar este problema, se realiza una determinación de la energía y su gradiente para un valor adicional. Con este dato se aproxima la función de energía por un potencial cúbico y se adopta el tamaño del paso que minimiza la función. Las ventajas de este método son que la geometría no permanece en un punto silla, como suele ocurrir con los procedimientos anteriores y converge con mayor rapidez cuando el potencial es cuadrático (o cúbico). La mayor desventaja de este procedimiento es que para moléculas grandes requiere un almacenamiento excesivo de información, así como un tiempo de cómputo muy largo, por lo que está restringido a sistemas con menos de 200 átomos.

#### 2.6.4 Método de base adaptada Newton-Raphson

Esta rutina realiza la minimización de energía utilizando el algoritmo de Newton-Raphson aplicado a un subespacio del vector de coordenadas expresado en términos de las coordenadas de la última posición. La matriz de la segunda derivada se construye numéricamente a partir del cambio en los vectores gradiente y se invierte por un análisis de eigenvectores permitiendo a la rutina que omita puntos silla en la superficie de energía. En cada paso se calcula el vector gradiente residual y se utiliza para agregar un paso de descenso más profundo en un paso Newton-Raphson, incorporando una nueva dirección en el conjunto base.

#### 2.6.5 Dinámica y enfriamiento

En este caso la minimización de energía es equivalente a enfriar el sistema<sup>5</sup>. En este método las ecuaciones clásicas de movimiento se integran para un sistema mientras que la energía del sistema es reescalada periódicamente al reescalar las velocidades mediante un término de fricción. Este método es útil para encontrar la estructura con menor energía cercana a la configuración, sin embargo, suele dar resultados muy lejanos a la configuración inicial, por lo que este método no se utiliza para estudiar mínimos locales.

A continuación se presentará un resumen de la teoría de la dinámica molecular *ab initio*. No se presentarán los detalles de todo el método debido a que este es muy extenso. En cambio se presentan varias referencias de donde se puede extraer un material más completo.

---

<sup>5</sup>Quenched dynamics.

## 2.7 Dinámica molecular cuántica

Para calcular propiedades moleculares la herramienta más adecuada es la Mecánica Cuántica. Los cálculos que no utilizan la mecánica cuántica sólo son aceptables en cuanto a que son capaces de reproducir burdamente los resultados cuánticos.

Una aproximación cuántica de un sistema molecular implica que se debe resolver la ecuación de Schrödinger para un sistema de átomos

$$\mathbf{H}\Psi = E\Psi, \quad (2.66)$$

en donde  $\mathbf{H}$  es el operador hamiltoniano de un sistema de electrones y núcleos descritos por los vectores de posición  $\mathbf{r}$  y  $\mathbf{R}$  respectivamente.

Un hamiltoniano que contiene las interacciones más importantes para una molécula de  $n$ -electrones en ausencia de campos externos es:

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_1 + \mathbf{H}_2 \quad (2.67)$$

con

$$\mathbf{H}_0 = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} \quad (2.68)$$

$$\mathbf{H}_1 = \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>1}^M \frac{Z_A Z_B}{R_{AB}} \quad (2.69)$$

$$\mathbf{H}_2 = \sum_{i=1}^N \xi(r_i) \mathbf{l}_i \cdot \mathbf{s}_i, \quad (2.70)$$

donde  $r_{iA}$  es la distancia del núcleo  $A$  al  $i$ -ésimo electrón,  $r_{ij}$  es la distancia entre los electrones  $i$ -ésimo y  $j$ -ésimo,  $M_A$  es la razón de masa del núcleo  $A$  a la masa de un electrón y  $Z_A$  es el número atómico del núcleo. El término  $\mathbf{H}_0$  contiene la energía cinética y potencial de los núcleos y electrones,  $\mathbf{H}_1$  representa la repulsión coulombiana entre pares de electrones y de núcleos,  $\mathbf{H}_2$  es la interacción espín-órbita, que bajo ciertas consideraciones de simetría se puede incluir en la descripción molecular.

Las contribuciones debidas a  $\mathbf{H}_1$  y  $\mathbf{H}_2$  dependen del número atómico  $Z$ . Para  $Z$  pequeño  $\mathbf{H}_1$  es dominante comparado con  $\mathbf{H}_2$  y contrariamente para  $Z$

mayores  $H_2$  es dominante comparado con  $H_1$ . En los casos límite tendremos los siguientes hamiltonianos

$$H_I = H_0 + H_1$$

$$H_{II} = H_0 + H_2$$

Debido a que en general los sistemas biológicos están constituídos de átomos ligeros,  $Z$  pequeños, es razonable tomar el primer límite  $H_I$  para representar el hamiltoniano total del sistema. Por lo cual el hamiltoniano total que se considerará para  $N$ -electrones y  $M$ -núcleos es en unidades atómicas

$$H = H_I = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>1}^M \frac{Z_A Z_B}{R_{AB}} \quad (2.71)$$

Para resolver la ecuación 2.71 de un sistema molecular se recurre a la aproximación de **Born-Oppenheimer** que se explica en la siguiente sección.

### 2.7.1 Aproximación de Born-Oppenheimer

Una discusión cualitativa de la aproximación de Born-Oppenheimer se basa en el siguiente hecho:

*Debido a que los electrones se mueven mucho más rápido que los núcleos atómicos, por ser estos últimos más pesados, es posible desacoplar en forma paramétrica su movimiento del movimiento nuclear.*

Es decir, se puede separar la ecuación (2.66) en una ecuación para el **problema electrónico** y otra para el **problema nuclear** mediante el siguiente procedimiento.

1. Consideremos que los electrones se mueven bajo el campo producido por los núcleos fijos con coordenadas  $R'$ , por lo que en el hamiltoniano (2.71) se puede despreciar el término de energía cinética del movimiento nuclear y el término de repulsión nuclear será considerado como una constante. Definiéndose el hamiltoniano electrónico como

$$H_e = -\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>1}^M \frac{Z_A Z_B}{R_{AB}} \quad (2.72)$$

2. Reescribiendo la ecuación de Schrödinger bajo la suposición anterior, se obtiene:

$$\mathbf{H}_e \Psi_e(r, R') = E_e(R') \Psi_e(r, R'), \quad (2.73)$$

donde la función  $\Psi_e(r, R')$  es la función de onda electrónica, la cual depende paramétricamente de las posiciones nucleares y  $E_e(R')$  es el eigenvalor de la energía correspondiente a cada configuración nuclear  $R'$ .

3. La función de onda se obtiene como el producto de una función de onda electrónica y una función de onda nuclear, tenemos

$$\Psi(r, R) = \Psi_e(r, R) \Psi_n(R). \quad (2.74)$$

De la ecuación (2.66) y (2.73) se obtiene que

$$\mathbf{H}_n \Psi_n(R) = E \Psi_n(R), \quad (2.75)$$

en donde

$$\mathbf{H}_n = - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + E_e. \quad (2.76)$$

El significado físico de esta aproximación es que la función de onda nuclear de la ecuación (2.74) se calcula por el **potencial efectivo electrónico**  $E_e$ , el cual se obtiene a partir de la ecuación (2.73) para cada configuración nuclear.

Debido a su complejidad, la búsqueda de la solución de la ecuación (2.73) por diferentes métodos ha dado lugar a gran cantidad de códigos de programación. Estos códigos en general se agrupan en dos tipos de cálculos moleculares conocidos con los nombres de *ab initio*<sup>6</sup> y semiempíricos. Los métodos semiempíricos utilizan un hamiltoniano más simple que el hamiltoniano molecular correcto y parámetros cuyos valores son ajustados con datos experimentales o resultados de cálculos *ab initio* para obtener una función potencial. Los cálculos *ab initio* generalmente tratan de resolver el hamiltoniano como se presenta

<sup>6</sup>El nombre *ab initio* proviene de la misma palabra en latín que significa "desde el principio" e indica que los cálculos son basados en principios fundamentales.

en la ecuación (2.72), para lo cual se han desarrollado una gran variedad de métodos. El método que se utilizará para los cálculos realizados en el este trabajo se basa en la **Teoría de funcional de la densidad** que se explica a continuación.

## 2.7.2 Teoría de funcional de la densidad

La mayoría de los métodos *ab initio* utilizan la aproximación de Hartree-Fock (HF) en donde se resuelven las ecuaciones HF para obtener los orbitales que pueden ser utilizados para construir la función de estado[52]. Sin embargo, el uso de estos métodos para obtener la función de estado de un sistema está altamente restringido debido a la dificultad de realizar cálculos exactos con conjuntos grandes de bases y mucho menos aplicarlo a moléculas que tienen un gran número de átomos.

Recientemente ha surgido un método alternativo llamado **Teoría de Funcional de la Densidad (DFT)**. En contraste con el método HF que utiliza una función de estados configuracionales, el método **DFT** inicia con el concepto de densidad de probabilidad electrónica.

La idea básica de DFT es que la energía electrónica del sistema puede escribirse en términos de la densidad de probabilidad electrónica  $\rho$  (Teorema de Hohenberg-Kohn). Para un sistema de  $n$  electrones,  $\rho(\mathbf{r})$  denota la densidad electrónica total en un punto particular del espacio  $\mathbf{r}$ . La energía electrónica  $E$  es un **funcional** de la densidad electrónica denotado por  $E[\rho]$  en el sentido de que para una función  $\rho(\mathbf{r})$  existe una sola energía correspondiente.

El concepto de funcional de la densidad se basa en el método de Thomas-Fermi y en el método de Hartree-Fock-Slater, aunque el método de DFT fue propuesto hasta 1964 cuando se dió una prueba de que la energía del estado base y otras propiedades electrónicas del estado base están determinadas de forma única por la densidad electrónica. Sin embargo, el teorema de **Hohenberg-Kohn** es un teorema que establece la existencia de la dependencia entre la energía y la densidad electrónica, pero no dice cómo se construye esta función.

El siguiente paso en el desarrollo de DFT fué la derivación de un conjunto de ecuaciones para un electrón de las cuales se puede obtener la densidad electrónica  $\rho$ . Como lo demuestran Kohn y Sham [53], la energía electrónica  $E$  exacta del estado base de un sistema de  $n$  – electrones se escribe como

$$E[\rho] = -\frac{\hbar^2}{2m_e} \sum_{i=1}^n \int \psi_i^*(\mathbf{r}_1) \nabla_1^2 \psi_i(\mathbf{r}_1) d\mathbf{r}_1 - \sum_{I=1}^N \int \frac{Z_I e^2}{4\pi\epsilon_0 r_{I1}} \rho(\mathbf{r}_1) d\mathbf{r}_1 + \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{4\pi\epsilon_0 r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 + E_{XC}[\rho], \quad (2.77)$$

donde los orbitales de un electrón  $\psi_i (i = 1, 2, \dots, n)$  son las soluciones de la ecuación anterior y se conocen como los **orbitales de Kohn-Sham**. La densidad de carga exacta para el estado base en un punto  $\mathbf{r}$  está dada por

$$\rho(\mathbf{r}) = \sum_{i=1}^n |\psi_i(\mathbf{r})|^2, \quad (2.78)$$

donde la suma se efectúa sobre todos los orbitales ocupados de Kohn-Sham y se conoce una vez que los orbitales han sido calculados. El primer término en la ecuación (2.77) es la energía cinética de los electrones; el segundo término representa la atracción nuclear entre todos los núcleos  $N$  con índice  $I$  y número atómico  $Z_I$ ; el tercer término representa la interacción coulombiana entre la distribución de carga total (sumada sobre todos los orbitales) en  $\mathbf{r}_1$  y  $\mathbf{r}_2$ ; el último término es la energía de correlación del sistema, la cual es un funcional de la densidad y tiene lugar en todas las interacciones no-clásicas entre electrones. De los cuatro términos presentados en (2.77),  $E_{XC}$  es el único término que no se conoce de forma exacta. Aunque el teorema de Hohenberg-Kohn dice que  $E$  y por lo tanto  $E_{XC}$  son funcionales de la densidad electrónica, no conocemos la forma analítica exacta y estamos forzados a utilizar formas aproximadas de ella.

Los orbitales de Kohn-Sham (KS) pueden encontrarse al resolver la ecuación de Kohn-Sham. Esta ecuación se deriva al aplicar el principio variacional [50] a la energía electrónica  $E[\rho]$  con la densidad de carga dada por la ecuación (2.78).

Las ecuaciones KS para los orbitales de un electrón  $\Psi(\mathbf{r}_1)$  tienen la forma

$$\left\{ -\frac{\hbar^2}{2m_e} \nabla_1^2 - \sum_{I=1}^N \frac{Z_I e^2}{2\pi\epsilon_0 r_{I1}} + \int \frac{\rho(\mathbf{r}_2)}{r\pi\epsilon_0 r_{12}} d\mathbf{r}_2 + V_{XC}(\mathbf{r}_1) \right\} \Psi_i(\mathbf{r}_1) = \epsilon_i \Psi_i(\mathbf{r}_1) \quad (2.79)$$

donde  $\epsilon_i$  son las energías de los orbitales KS y el potencial de correlación  $V_{XC}$  es la derivada funcional de la energía de correlación:

$$V_{XC}[\rho] = \frac{\delta E_{XC}[\rho]}{\delta \rho}. \quad (2.80)$$

Si  $E_{XC}$  es conocida, entonces  $V_{XC}$  puede obtenerse. El significado de los orbitales KS es que estos permiten que la densidad  $\rho$  sea calculada de la ecuación (2.78).

Las ecuaciones KS se resuelven por el método autoconsistente. Inicialmente se propone la densidad de carga  $\rho$  para lo cual se utilizan las densidades atómicas de todos los sistemas moleculares. Utilizando una forma aproximada de la dependencia funcional de  $E_{XC}$  en la densidad, se calcula  $V_{XC}$  como función de  $\mathbf{r}$ . El conjunto de ecuaciones KS es resuelto para obtener una densidad inicial de la ecuación (2.78). A continuación el conjunto de orbitales se utiliza para calcular una nueva densidad mejorada de la ecuación (2.78) y el proceso se repite hasta que la densidad y la energía de correlación-intercambio converjan a un límite de tolerancia. La energía electrónica es calculada de (2.77).

En cada iteración se calculan los orbitales KS y esto puede hacerse de forma numérica o pueden ser expresados en términos de un conjunto de funciones base. En este último caso, encontrar las ecuaciones KS significa encontrar los coeficientes de la expansión en el conjunto base.

### 2.7.3 Orbitales moleculares y atómicos

La función de onda por lo general es difícil de resolver, por lo que existen métodos aproximados que tratan de predecir la forma que esta función debe tener. Uno de estos métodos consiste en expandir la función de onda molecular como una combinación de orbitales atómicos (LCA). Cada orbital se supone que es una función gaussiana, es decir, si escribimos la función de onda como

$$\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n) = |\psi_1(\vec{x}_1)\psi_2(\vec{x}_2)\dots\psi_n(\vec{x}_n)\rangle, \quad (2.81)$$



entonces podemos escribir cada orbital atómico  $\psi$  como

$$\psi_k(\vec{x}) = \sum_{i=1}^N \sum_{j=1}^{N_i} a_{ij}^k \chi_{ij}(\vec{x}) \quad (2.82)$$

las sumas son tomadas sobre el número total de átomos  $N$ , a continuación las capas atómicas se expanden en términos de Gaussianas para facilitar la solución de las ecuaciones integro-diferenciales que aparecen en el método de campo-autoconsistente-(SCF<sup>7</sup>) [51, 52]:

$$\chi(\vec{x}) = \sum_{k=1}^L \beta^k g_k(\alpha^k, \vec{x} - \vec{R}^k) \quad (2.83)$$

Las bases más simples son aquellas que utilizan un conjunto mínimo de Gaussianas ( $L=3$ ) y en la literatura se reconoce como **STO-3G** en donde **STO** significa **Orbitales tipo Slater**<sup>8</sup>. Estas funciones son capaces de reproducir las propiedades básicas de las capas atómicas y están centradas en el núcleo atómico [51]. Las bases que incluyen a tres funciones gaussianas para orbitales de cada capa cerrada y que además incluye un término de polarización se conoce como **6-31G\***. Debido a que los parámetros (los coeficientes  $\alpha$  y  $\beta$  en (2.83)) que se encuentran en las Gaussianas son fijos, la forma de la función de onda está completamente definida por los coeficientes de los orbitales moleculares  $a_{ij}^k$ . Esta es la razón por la que todos los esfuerzos están concentrados en encontrar los mejores coeficientes  $a_{ij}^k$  utilizando por lo general un método variacional que minimiza la energía **DFT**.

En este capítulo hemos visto las bases para resolver las ecuaciones de Newton por diferentes métodos de integración así como los principios fundamentales de una Dinámica Molecular cuántica. También hemos visto cómo se calculan ciertas variables termodinámicas necesarias para la estabilidad de la integración. Pero la simulación de Dinámica Molecular empieza con el conocimiento de la energía del sistema, la cual se puede calcular utilizando mecánica cuántica, sin embargo, para sistemas grandes, como polímeros, se convierte en un problema muy complicado, de ahí que para determinar la energía del sistema se recurra a cálculos empíricos (potenciales modelo). Esta forma de obtener la energía del sistema se conoce como Campo de Fuerza (Force Field) y su descripción en detalle se dará en el siguiente capítulo.

<sup>7</sup>Self Consistent Field.

<sup>8</sup>Slater type orbitals.

# Capítulo 3

## Campos de Fuerza

### Introducción

La simulación de DM empieza por suponer una forma específica de interacción entre las partículas que componen a nuestro sistema. Aunque es posible realizar cálculos con la mecánica cuántica para determinar las propiedades moleculares de un sistema pequeño, el uso de métodos de primeros principios involucra la solución de ecuaciones muy complicadas tanto matemáticamente como computacionalmente. Por otro lado, los polímeros involucrados en sistemas biológicos son muy grandes, por lo que se recurre al uso de la mecánica molecular utilizando funciones de potencial empíricos. Los potenciales están en términos de coordenadas internas y de constantes apropiadas para poder reproducir los datos experimentales. A este tipo de potenciales de interacción se les conoce genéricamente como **Campos de Fuerza**. Existen varios tipos de Campos de Fuerza según el sistema que se desea estudiar y la aplicación que se le quiera dar.

En este capítulo se hace la descripción general de las componentes de un Campo de Fuerza y después la descripción de la función de potencial que se usará. El potencial a describir corresponde al programa **CHARMM**<sup>1</sup> [43] creado por Martin Karplus y colaboradores de la Universidad de Harvard. Este programa es adecuado para modelar sistemas macromoleculares.

---

<sup>1</sup>Chemistry at Harvard Macromolecular Mechanics.

### 3.1 Factores para la formulación de un campo de fuerza

Para tratar un sistema de átomos en dinámica molecular se utilizan varios principios que difieren de la mecánica cuántica. En dinámica molecular el electrón y el núcleo están unidos para formar una sola partícula esférica que simula un átomo. El radio de la partícula se obtiene de manera experimental o teórica y tiene una carga neta obtenida de la teoría. Las interacciones entre los átomos están basadas en osciladores armónicos y potenciales clásicos y se preasignan a un conjunto específico de átomos.

La siguiente aproximación que se hace consiste en expresar al potencial de interacción como una función de las coordenadas atómicas y con esto obtener las fuerzas interatómicas. Luego estas fuerzas se utilizan con la ecuación de Newton para determinar la evolución (trayectoria) del sistema de partículas. De esta manera, un problema que resulta demasiado complejo a nivel de mecánica cuántica, puede resolverse utilizando sólo ecuaciones clásicas. Para distinguir los varios tipos de función de energía se deben tomar en cuenta los siguientes atributos:

1. **Coordenadas:** si  $N$  es el número de átomos en la molécula, el conjunto de coordenadas más apropiado podría ser el conjunto de  $3N - 6$  coordenadas normales que se derivan a partir del análisis vibracional de la molécula. Este conjunto es completo en el sentido que es el número exacto para definir la geometría interna de la molécula y no son redundantes. Algunas expresiones de energía se dan en términos de coordenadas internas como por ejemplo los ángulos de enlace, de torsión, etc. El problema de estos conjuntos de coordenadas está en que se tienen aproximadamente  $5N$  coordenadas o más representando la estructura lo cual provoca redundancia en la descripción del sistema.
2. **Orden de la expansión:** la expresión de la energía se forma de una expansión en serie alrededor del punto de energía mínima. Esta serie puede tener diversos grados de aproximación como la armónica, aunque también se podrían incluir términos cúbicos. En algunos casos se incluyen términos anarmónicos como la torsión o los términos de interacción de larga distancia (de no enlace). La elección de estos diferentes grados de exactitud depende del sistema.

3. **Formas funcionales:** es la expresión analítica que describe los tipos de fuerzas entre cada par de átomos. Por ejemplo:

### TÉRMINOS DE ENERGÍA INTERNA

#### Potencial de enlace

$K_b(r - r_0)^2$	armónico
$K_b(r - r_0)^2[1 + k_1(r - r_0)]$	término cúbico anarmónico
$D[1 - \exp(-A(r - r_0))]^2$	función de Morse

#### Potencial de ángulo de enlace

$K_\theta(\theta - \theta_0)^2$	armónico en $\theta$
$K_\theta(\cos\theta - \cos\theta_0)^2$	armónico en $\cos\theta$
$K_\theta(\theta - \theta_0)^2 + F(q - q_0)^2 + F'(q - q_0)$	forma de Urey-Bradley

#### Potencial de torsión

$K_\phi[1 + \cos(n\phi - \delta)]$	forma trigonométrica
$\sum V_n \cos(n\phi)$	series trigonométricas

### INTERACCIONES DE NO ENLACE

#### Interacción de van der Waals

$\frac{A}{r^{12}} - \frac{B}{r^6}$	Lennard-Jones
------------------------------------	---------------

#### Potencial electrostático

$\frac{q_i q_j}{\epsilon r_{ij}}$	interacción Coulombiana
$E_{el} = \sum_{excl(ij)=1} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} SW(r_{ij}^2 r_{on}^2 r_{off}^2)$	dieléctrico dependiente de la distancia
$E_{el} = \sum_{excl(ij)=1} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \left(1 - \frac{2r_{ij}^2}{r_{cut}^2}\right)$	dieléctrico ajustado
$D[\exp(-2A(r - r_0)) - 2\exp(-A(r - r_0))]$	forma de Morse

4. **Tipos de átomos:** cada término de la energía se refiere a la deformación de alguna coordenada interna definida por sus átomos vecinos. Cada término contiene parámetros empíricos: los coeficientes  $K_b, K_\theta, K_\phi, R_0, \theta_0$ , etc. Uno debe escoger y asignar un parámetro único a cada término o tratar cada grupo químico en un marco consistente.

La elección de un tipo de átomo determina la sensibilidad del Campo de Fuerza para distinguir entre los diferentes medios químicos. Por ejemplo, pueden existir 20 tipos de átomos de carbono si tomamos en cuenta el tamaño del anillo que forma, posibilidades de resonancia y sustituyentes. Para moléculas grandes se pueden suponer átomos unidos vistos como una sola partícula, como por ejemplo un carbono con todos los hidrógenos unidos.

5. **Datos de parametrización:** son los datos usados para calibrar la expresión de energía, se obtienen de manera experimental o tratando de ajustar resultados obtentidos por medio de la mecánica cuántica. Con estos parámetros se absorben y reproducen los valores de propiedades determinadas por la mecánica cuántica pero sin efectuar los cálculos cuánticos.

## 3.2 Formas funcionales

Para desarrollar una expresión general empírica de la energía, se toma la conformación molecular que representa un mínimo local o global de la energía y se escribe la expresión de la energía como una expansión en serie de Taylor sobre ese punto. Para tomar el caso más general, tomaremos un sistema de coordenadas generalizadas  $\{S_i\}$  y se expresa la energía en función de estas coordenadas  $E = E(S_i)$ . La expansión de la energía en serie de Taylor para una deformación pequeña de  $S_i$  es:

$$E = E_0 + \sum_i G_i \Delta S_i + \sum_i K_i \Delta S_i^2 + \sum_{i \neq j} F_{ij} \Delta S_i \Delta S_j + \dots \quad (3.1)$$

*términos de tercer orden + términos de mayor orden*

En la suma del segundo término de (3.1), las  $G_i$  son las primeras derivadas con respecto a las coordenadas o los gradientes de la energía, es decir,

$$G_i = \frac{\partial E}{\partial S_i}.$$

Pero en el mínimo todos estos gradientes es cero y la energía toma el valor constante  $E_0$ . Este valor puede tomarse como la energía de enlace residual inherente a la molécula. Conforme uno se aleja del mínimo estos gradientes dejan de ser cero y la suma del segundo término en (3.1) es una aproximación lineal al incremento de energía.

La segunda y tercera sumas son de segundo orden en  $\Delta S_i$  pero están escritas en forma separada para enfatizar la distinción entre los términos diagonales (aquellos que dependen de una sola coordenada  $S_i$ ) y los términos cruzados que contienen explícitamente deformaciones en dos coordenadas,  $S_i$  y  $S_j$  y que representan el acoplamiento de dos coordenadas. Los coeficientes  $K_i$  y  $F_{ij}$  son las segundas derivadas de la energía

$$K_i = F_{ii} = \frac{\partial^2 E}{\partial S_i^2}$$

$$F_{ij} = \frac{\partial^2 E}{\partial S_i \partial S_j}.$$

Los términos de mayor orden se construyen de manera similar. Para deformaciones suficientemente pequeñas, se tiene una buena aproximación al truncar la serie después del término de segundo orden. Esta es una aproximación **armónica**, se nombra de esta manera debido a que sería exacta si las fuerzas entre cada par de átomos se comportaran como fuerzas de resortes ideales u osciladores armónicos. Los términos de mayor orden contribuyen con términos **anarmónicos** cuyos efectos pueden ser evaluados en el experimento pero que generalmente contribuyen poco y pueden ser despreciados.

Los términos armónicos estiman la energía asociada con la vibración en la longitud de equilibrio y es la ecuación de una parábola. En esta aproximación el modelo tiende a fallar cuando el enlace se rompe.

Las formas funcionales de tipo armónico, que se utilizarán en la función de energía son:

$$\text{Potencial de enlace:} \quad E_b = K_b(r - r_0)^2$$

$$\text{Potencial de ángulo de enlace:} \quad E_\theta = K_\theta(\theta - \theta_0)^2$$

$$\text{Torsión impropia:} \quad E_w = k_w(w - w_0)^2$$

Esta aproximación es suficiente para temperaturas ordinarias y no permite el rompimiento de enlaces dada la naturaleza armónica de los términos del potencial, donde a mayor separación se tiene más energía. Nótese que con esta expresión sólo se tendría un modelo basado en esferas y resortes, por eso se han implementado otros términos que han permitido que el modelo se acerque más a la realidad, tales términos son: ángulos de torsión, interacciones de partículas no enlazadas, ángulos de valencia, etc.

A continuación se hace la descripción de los términos del Campo de Fuerza de CHARMM. Esta función de energía está basada en coordenadas internas separables y términos de interacciones de no enlace. La energía total está expresada en la forma:

$$E = E_b + E_\theta + E_\phi + E_w + E_{vdw} + E_{el} + E_{hb} + E_{cr} + E_{c\phi} \quad (3.2)$$

y a continuación se presenta la fórmula para cada uno de estos términos junto con su descripción.

### 3.2.1 Energía de enlace

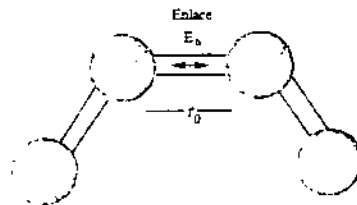


Figura 3.1: Potencial de enlace.

La energía de enlace representa el desplazamiento instantáneo de la longitud de enlace ideal  $r_0$ , figura 3.2.

Su expresión analítica es la siguiente [34, 40, 44]:

$$E_b = \sum k_b (r - r_0)^2 \quad (3.3)$$

donde la suma se hace sobre todos los enlaces.

Está basada en la ley de Hooke, es decir es una aproximación armónica. El parámetro  $k_b$ , definido como la constante de fuerza, determina la flexibilidad del enlace y puede evaluarse a partir de frecuencias de alargamiento en infrarrojo o por cálculos de mecánica cuántica. El valor  $r_0$  define la longitud de equilibrio y se determina por estructuras cristalinas a baja temperatura o por espectroscopía de microondas.

### 3.2.2 Energía de ángulo de enlace

Esta expresión también se basa en la ley de Hooke y la figura 3.2 muestra como se mide este ángulo.

$$E_\theta = \sum_{\text{angulos}} k_\theta (\theta - \theta_0)^2 \quad (3.4)$$

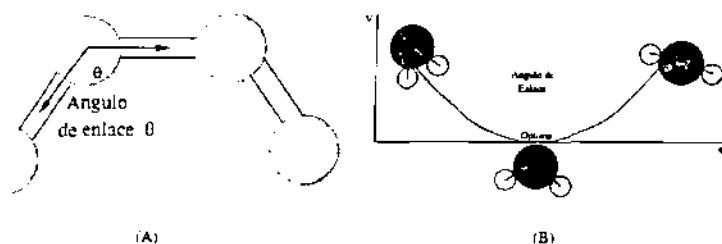


Figura 3.2: (A) Potencial de ángulo de enlace, (B) diferentes valores del ángulo de enlace.

El parámetro  $k_\theta$  controla la rigidez del ángulo, mientras que  $\theta_0$  define el ángulo de equilibrio. Esta ecuación estima la energía asociada con la vibración alrededor del ángulo de equilibrio, figura 3.2.

Los parámetros  $k_b$  y  $k_\theta$  tienen la función de abrir o cerrar la parábola que representa el potencial armónico.



### 3.2.3 Energía de ángulo de torsión

El ángulo de torsión es un término que surge de considerar un arreglo de cuatro átomos basado en el ángulo diedro formado por el plano de los 3 primeros átomos con el de los 3 últimos, figura 3.3 [34, 40, 44].

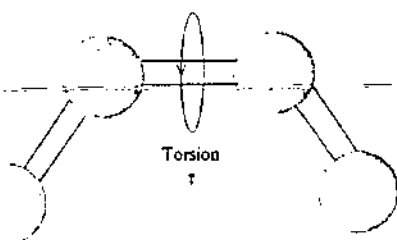


Figura 3.3: Potencial de torsión.

Para escribir la expresión de energía del ángulo de torsión se podría hacer de manera similar que con el ángulo de enlace, con expansión armónica, sin embargo, la experiencia demuestra que se puede realizar una mejor aproximación utilizando una expansión trigonométrica ya que se sabe que la rotación alrededor de enlaces simples puede encontrar pequeñas barreras y se tiene la idea de que el potencial torsional tiene forma periódica. Es, por lo tanto, costumbre escribir la expresión para los grados de libertad torsionales en forma trigonométrica, y la expresión más simple es

$$E_{\phi} = K_{\phi}[1 + \cos(n\phi - \delta)] \quad (3.5)$$

Aquí la altura de la barrera va de 0 a  $2K_{\phi}$  conforme el término del coseno varía de  $-1$  a  $1$ . El parámetro  $n$  controla la periodicidad de la función y el ángulo de fase  $\delta$  es una herramienta matemática para forzar a obtener el mínimo en el lugar correcto.

La energía de torsión se utiliza en dinámica molecular para corregir los términos residuales de energía más que representar un proceso físico.

### 3.2.4 Energía de ángulo de torsión impropia

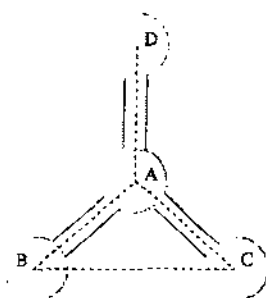


Figura 3.4: Definición de ángulo de torsión impropia.

Otro término que permite tener una mejor aproximación cerca del mínimo de energía es la **torsión impropia**. La torsión impropia se define a partir de la figura 3.4, es el ángulo  $\omega$  definido por el ángulo  $\phi_{ABCD}$  o de manera equivalente, el ángulo entre los planos  $ABC$  y  $BCD$ . La torsión impropia se escribe:

$$E_{\omega} = k_{\omega}(\omega - \omega_0)^2 \quad (3.6)$$

El término de torsión impropia se ha diseñado para mantener los planos de cierto arreglo de átomos con una distorsión de un término cuadrático, sin este término, los potenciales que están fuera del plano tienden a ser cuárticos. Además, este término provee una mejor expresión de energía cerca del mínimo, que es importante para la dinámica y el análisis vibracional.

### 3.2.5 Potencial de Lennard-Jones

Con el potencial tipo Lennard-Jones se supone que los átomos de un sistema interactúan por pares. Este potencial comprende dos tipos principales de fuerzas. La primera modela la resistencia a la compresión y tiene carácter repulsivo a distancias pequeñas. La segunda modela la atracción entre átomos para formar el estado líquido o sólido y tiene carácter de fuerza de larga distancia. En general las funciones de potencial que exhiben estas características pueden adoptar una gran variedad de formas. Cuando estas formas se escogen cuidadosamente se vuelven modelos capaces de describir al sistema con un buen grado de exactitud.

El potencial más conocido propuesto primeramente para el Argón líquido fué el de Lennard-Jones. Para un par de átomos  $i$  y  $j$  localizados en  $r_i$  y  $r_j$ , la energía potencial es:

$$u(r_{ij}) = 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right], \quad \text{si} \quad r_{ij} \leq r_c. \quad (3.7)$$

donde  $r_{ij} = |r_i - r_j|$ , figura 3.5.

El parámetro  $\epsilon$  rige la fuerza de interacción y el parámetro  $\sigma$  define la longitud de escala. La interacción es repulsiva a distancias cortas y luego atractiva, para eventualmente anularse en un límite de separación  $r_c$ . La repulsión en el potencial proviene de nubes electrónicas que se traslapan y su parte atractiva proviene de fuerzas tipo Van der Waals que se deben a la dispersión entre electrones. Las interacciones involucran sólo pares de átomos y cada par es tratado independientemente de los otros.

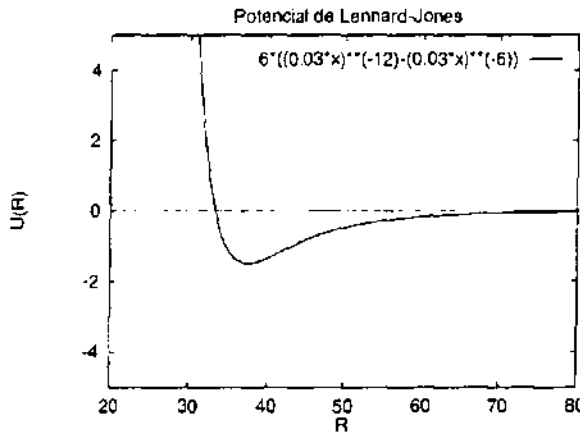


Figura 3.5:  $u(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right]$ ,  $\sigma = 1.5$ ,  $\epsilon = 0.03$

La fuerza correspondiente al potencial de Lennard-Jones es  $f = -\nabla u(r)$ . La fuerza que ejerce el átomo  $j$  sobre el  $i$  es

$$f_{ij} = \left( \frac{48\epsilon^2}{\sigma} \right) \left[ \left( \frac{\sigma}{r_{ij}} \right)^{14} - \frac{1}{2} \left( \frac{\sigma}{r_{ij}} \right)^8 \right] r_{ij}. \quad (3.8)$$

con la restricción de que  $r_{ij}$  sea mucho menor que  $r_c$  y cero en cualquier otra parte. Conforme  $r$  tiende a  $r_c$  la fuerza se anula de tal manera que no hay discontinuidad en  $r_c$ , pero  $\nabla f$  y derivadas de orden más alto son discontinuas, aunque esto no tiene un impacto práctico en la solución numérica.

### 3.2.6 Energía de interacciones de no enlace

Las interacciones que no son de enlace ó de largo alcance son producidas por pares de átomos que no están unidos unos con otros. Estas funciones dependen de las distancias entre los átomos. Estas interacciones consisten en lo general de dos términos uno estereoquímico y otro electrostático. El componente estereoquímico es la interacción de Van der Waals entre dos pares de átomos  $ij$  y se puede representar con el potencial  $n - m$  de Lennard-Jones,

$$E_{ij}(vdW) = \frac{A_{ij}}{R_{ij}^n} - \frac{B_{ij}}{R_{ij}^m} \quad (3.9)$$

donde  $n$  y  $m$  son enteros,  $R_{ij}$  es la separación entre los dos núcleos y los parámetros  $A_{ij}$  y  $B_{ij}$  son específicos para cada par de átomos  $i$  y  $j$ . La forma más común es el potencial 12-6 donde  $n$  es 12 y  $m$  es 6. Las componentes electrostáticas se escriben en términos de las cargas atómicas parciales  $q_i$  y  $q_j$  en los dos átomos,

$$E_{ij}(el) = \frac{Cq_iq_j}{\epsilon R_{ij}} \quad (3.10)$$

donde  $C$  es el factor de conversión que pone el resultado en las unidades correctas, y  $\epsilon$  es la constante dieléctrica. De manera estricta no está bien formulado el concepto de constante dieléctrica a nivel microscópico como se considera aquí, pero se incluye el factor con el fin de reducir o apantallar las interacciones entre cargas a largas distancias. Una aproximación común es considerar  $\epsilon$  como una función de las distancias interatómicas y poner  $\epsilon(R) = KR_{ij}$  donde  $K$  es una constante, de esta forma el denominador de la expresión electrostática contiene una potencia impar de  $R_{ij}$ .

### 3.2.7 Puentes de hidrógeno

Involucra a los átomos *acceptor antecedente* ( $AA$ ), *acceptor* ( $A$ ), *hidrógeno* ( $H$ ) y *átomo pesado donador* ( $D$ ). Su expresión analítica es [43]:

$$E_{hb} = \sum_{i,j} \left( \frac{A'}{r^{i,AD}} - \frac{B'}{r^{j,AD}} \right) \cos^m(\theta_{A-H-D}) \cos^n(\theta_{AA-A-H}) \times \\ \times sw(r^2_{AD}, r^2_{h_{on}}, r^2_{h_{off}}) sw[\cos^2(\theta_{A-H-D}), \cos^2(\theta_{h_{on}}), \cos^2(\theta_{h_{off}})]$$

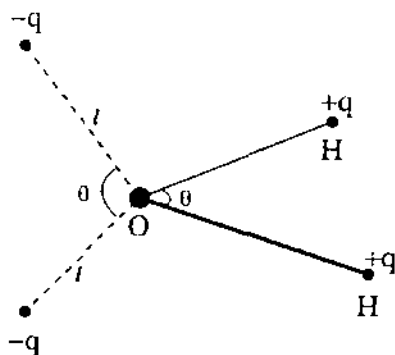
donde  $i$  y  $j$  son enteros positivos,  $m = (0, 2, 4)$  y  $n = (0, 2)$ . La función  $sw$  (switching function) es una función de empalme definida por

$$sw(r, r_{on}, r_{off}) = \left\{ \begin{array}{ll} 1 & \text{cuando } r \leq r_{on} \\ \frac{(r_{off}-r)^2(r_{off}+2r-3r_{on})}{(r_{off}-r_{on})^3} & \text{cuando } r_{on} < r \leq r_{off} \\ 0 & \text{cuando } r > r_{off} \end{array} \right\}$$

La elección del enlace de hidrógeno está basado en la distancia  $A-D$  y en el ángulo  $A-H-D$ . El término completo es cero si el ángulo  $\theta_{A-H-D}$  es menor de  $90^\circ$  o si  $n$  es un entero mayor que 0 y  $\theta_{AA-A-H}$  es menor de  $90^\circ$ . Los exponentes  $i$ ,  $j$  se determinan cuando se lee el archivo de parámetros y los exponentes  $n$  y  $m$  se obtienen a partir del archivo de topología cuando se enlista el donador y el acceptor. El exponente  $n$  depende del tipo de acceptor y el exponente  $m$  del tipo de donador. Cuando los donadores son tratados como átomos extendidos sólo se utilizan los términos radiales. Además de este potencial de puente de hidrógeno, se incluyen las interacciones electrostáticas.

### 3.2.8 Interacciones entre las moléculas de agua

Para la descripción de estas interacciones, CHARMM utiliza el potencial  $ST^2$  [45] o una combinación de los términos anteriores (electrostático, Van der Waals y puentes de hidrógeno), aunque es preferible utilizar el potencial  $ST^2$  en ambos términos para tener mayor aproximación y eficiencia computacional para describir este tipo de interacciones.

Figura 3.6: Modelo de una molécula de agua para el potencial  $ST^2$ .

La función de potencial  $ST^2$  está basada en un modelo de cuatro cargas puntuales rígidas para cada molécula de agua. Esta geometría se muestra en la figura 3.6.

Las cargas positivas  $+q$  se identifican como los protones y están localizados a  $1 \text{ \AA}$  del núcleo de oxígeno,  $O$ . La distancia  $l$  de  $O$  a cada carga negativa  $-q$  es también de  $1 \text{ \AA}$ . El potencial  $ST^2$  consiste en un potencial de Lennard-Jones actuando entre los oxígenos, más un término de Coulomb para los 16 pares de interacciones entre las cargas puntuales, denotando como  $r_{12}$  a la distancia entre oxígenos de dos moléculas de agua, tenemos como expresión general

$$V(1, 2) = V_{LJ}(r_{12}) + S(r_{12})V_{el}(1, 2), \quad (3.11)$$

en donde

$$V_{LJ}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right]$$

$$V_{el}(1, 2) = q^2 \sum_{\alpha, \beta=1}^4 \frac{(-1)^{\alpha+\beta}}{d_{\alpha, \beta}(1, 2)}$$

con  $d_{\alpha, \beta}(1, 2)$  como la distancia entre la carga  $\alpha$  de la molécula 1 y la carga  $\beta$  en la molécula 2 ( $\alpha$  y  $\beta$  son pares para cargas positivas e impares para cargas negativas).

La función moduladora varía suavemente entre el 0 (distancias pequeñas) y el 1 (distancias grandes),

$$S(r_{12}) = \left\{ \begin{array}{ll} 0 & \text{cuando } 0 \leq r_{12} \leq R_L \\ \frac{(r_{12} - R_L)^2(3R_U - R_L - 2r_{12})}{(R_U - R_L)^3} & \text{cuando } R_L \leq r_{12} \leq R_U \\ 1 & \text{cuando } R_U \leq r_{12} \end{array} \right\}$$

Para el potencial  $ST^2$  se tienen los siguientes parámetros [45]:  $\sigma = 3.10 \text{ \AA}$ ,  $\epsilon = 5.2605 \times 10^{-15} \text{ erg}$ ,  $q = 0.2357 e$ ,  $R_L = 2.0160 \text{ \AA}$  y  $R_U = 3.1287 \text{ \AA}$ .

### 3.2.9 Constricciones

En algunas ocasiones no es deseable que los átomos presenten movimientos grandes dentro de la molécula. Con este fin existe forma de restringir los movimientos de los átomos al desarrollar una minimización de energía o una dinámica. Las constricciones que se pueden incluir en la energía son:

	Variable que se modifica
1. Constricción en Átomos armónicos	$E_{cr} = \sum K_i(r_i - r_{i0})^2$
2. Rígido	$\frac{\partial E}{\partial \tau} = 0$
3. Constricciones en los ángulos diedros	$E_{c\phi} = \sum K_i(\phi_i - \phi_{i0})$
4. Constricción en la distancia	$\partial \tau_{ij} = 0$

donde  $K_i$  puede hacerse muy grande y  $\partial \tau = 0$  para todas las operaciones. La constricción de átomo armónico es utilizada principalmente para no permitir que los átomos tengan desplazamientos grandes cuando se está minimizando (manteniendo un contacto fijo) mientras que se permite que otra parte de la estructura se relaje. Las constricciones diédricas se utilizan para mantener cierta conformación local o cuando se necesita examinar una serie de diferentes conformaciones para hacer un mapeo de energía.

### 3.3 Programa computacional CHARMM

En esta sección se resume como funciona el programa computacional utilizado CHARMMv24. Cada una de las Dinámicas se realizó bajo condiciones ambientales diferentes con el fin de determinar la **Conformación Cercana a línea de Ataque (CCA)**<sup>2</sup>.

Las partes que conforman una Dinámica son:

1. Transcribir los datos experimentales de los átomos que conforman a la molécula como son las posiciones atómicas, en términos de parámetros computacionales determinados por CHARMM.
2. Formación del ambiente, en donde se define si la molécula estará en presencia de moléculas de agua o iones metálicos.
3. Periodo de relajación molecular a través de una minimización de la energía de la molécula con la introducción de los parámetros del Campo de Fuerza.
4. Condiciones de frontera.
5. Inicio de la Dinámica Molecular, dadas las posiciones de los átomos, se asignan velocidades aleatorias pero siguiendo una distribución de Maxwell-Boltzmann y se integran las ecuaciones de movimiento de Newton para obtener la nueva posición.
6. Análisis de resultados que arroja el archivo de salida como son datos de temperatura, energía y variables estadísticas, así como las posiciones moleculares de cada paso de la corrida. Con estas coordenadas se pueden a su vez obtener datos de distancias relativas, ángulos y la visualización de la secuencia de imágenes (película) de la dinámica.

En las siguientes secciones se dará la descripción detallada de cada uno de estos pasos y las conclusiones obtenidas a través de los datos que arrojaron las diferentes dinámicas.

---

<sup>2</sup>Near Attack Conformation (NAC) [6].



### 3.3.1 Estructurar la molécula

Una vez que se tienen los datos de las coordenadas de los átomos que formarán nuestra molécula (archivo PDB<sup>3</sup>), estos datos los debe leer el programa con el fin de asignarle los parámetros correspondientes. Cabe señalar que la estructura cristalográfica no contiene hidrógenos y estos deben ser adicionados de forma adecuada por CHARMM. La forma como lo realiza es mediante los archivos de topología y parámetros en los cuales CHARMM reconoce el tipo de molécula con la que se trabaja (en este caso nucleótidos) y de acuerdo a los datos teóricos que se encuentran en estos dos archivos, puede tanto colocar los hidrógenos como distinguir cuáles átomos no forman enlaces, entre otras muchas cosas que se describen a continuación.

#### Archivo de estructura

El archivo de datos de la estructura más fundamental es el PSF (Protein Structure File) que contiene toda la información necesaria para generar la lista de enlaces de hidrógeno y los que no forman enlaces y de esta forma evaluar la energía. La información que presenta este archivo es la siguiente:

1. Residuos y segmentos: se indica cómo están agrupados los átomos para formar residuos y segmentos así como sus nombres de identificación.
2. Información de átomos: nombres de átomos, tipos, cargas, masas y constricciones rígidas.
3. Lista de enlaces, ángulos, diedros y diedros impropios.
4. Lista de donadores y aceptores, la cual se utiliza para construir los enlaces de hidrógenos.
5. Lista de exclusión de interacción electrostática o de van der Waals, debido a que la interacción de algunos átomos ya es está descrita en el hamiltoniano.

CHARMM construye este archivo PSF durante la dinámica cada vez que se quiera mediante un comando del Programa.

---

<sup>3</sup>Protein Data Bank, ver sección 1.7.1

### Archivo de topología

El archivo de topología RTF (Residue Topology File) contiene la información local acerca de los átomos, enlaces, ángulos, etc, para cada tipo de unidad monomérica (residuo) que puede ser utilizada para construir un tipo particular de macromolécula.

El concepto de residuo en CHARMM es fundamental desde un punto de vista organizacional. En el archivo de ácidos nucleicos, un residuo corresponde a un nucleótido a lo largo de una cadena de ARN o ADN. Este archivo contiene también la información de las cuatro distintas bases, grupos terminales y agua. En un típico archivo de topología se puede encontrar la siguiente información de la molécula:

- a) Tipo de átomo: CHARMM asigna *tipo de átomo*, es decir, un átomo puede tener varias presentaciones que corresponden a su ambiente químico, tipo de molécula en que se encuentra, posición y enlaces. El archivo de topología es un intérprete entre las posiciones de átomos como partículas (archivo PDB) y una estructura que distingue la función química de cada átomo para conformar una molécula de una gran importancia biológica. La Dinámica Molecular aproxima a los átomos como partículas con interacciones semiclásicas, por lo que es necesario incluir valores experimentales que permitan tener un mejor acercamiento al comportamiento real.
- b) Carga, de acuerdo al tipo de átomo, en el archivo de topología se le asigna un valor de carga por grupos o individuales y ésta debe corresponder a la de una molécula real.
- c) Enlaces donde se indica que átomos están unidos.
- d) Ángulos, los ángulos que se forman entre dos átomos, o entre tres átomos que pueden ser diedros o impropios.
- e) Átomos donadores de cargas.

Después de que se tiene la estructura de la molécula en coordenadas, se debe tener la energía de la molécula como función de estas coordenadas y para esto se crea un archivo de parámetros.

### Archivo de Parámetros

En este archivo de parámetros se encuentra la siguiente información:

- a) Parámetro de la energía de enlace,  $k_b$ , ecuación 3.3.
- b) Parámetro de la energía de ángulo,  $k_\theta$ , ecuación 3.4.
- c) Parámetro de la energía de ángulo de torsión,  $k_\phi$ , ecuación 3.5.
- d) Parámetro de la energía de ángulo impropio,  $k_\omega$ , ecuación 3.6.
- e) Parámetros de la energía de interacciones de no enlace, como son la electrostática y la de van der Waals.
- f) Átomos que contienen enlaces con hidrógenos. El archivo PDB no contiene esta información y en la primera parte que consiste en la lectura de los valores de topología y del campo de fuerza, el programa CHARMM puede generar estos hidrógenos.

Estos valores de parámetros son muy importantes porque determinan las estabildades relativas de las diferentes estructuras.

#### **Generación de la lista de no enlaces**

Esta lista contiene todas las parejas de interacciones que serán consideradas. Estas parejas de interacciones no se incluyen si los átomos se encuentran muy lejos o si se encuentran muy cerca, en este último caso son excluidos.

Durante la minimización o la dinámica, las posiciones relativas de los átomos pueden cambiar, pero la lista puede ser actualizada de forma periódica, lo cual permite tener una mayor eficiencia computacional.

#### **Generación de la lista de puentes de hidrógeno**

La lista de puentes de hidrógenos se genera a partir de una distancia y un ángulo preespecificados en el archivo de topología. Si existe una distancia menor o igual a una distancia de corte entre un hidrógeno y cualquier otro átomo (carbono, oxígeno, nitrógeno) se genera la lista de enlaces. Durante la minimización y la dinámica, la lista puede mantenerse fija o ser actualizada de forma periódica.

### 3.3.2 Ambiente

Para introducir a la molécula en un ambiente de agua, se construye primero un cubo formado por moléculas de agua cuyas dimensiones deben ser un poco más grandes que las de la molécula a solvatar. La presión a la que se encontrará la molécula depende de la presión de este cubo de agua, es decir, el número de moléculas por volumen que se coloquen en ella. Mediante un programa se traslapan los dos archivos que contienen las coordenadas de la molécula y las coordenadas de las moléculas de agua. Este programa borra las moléculas de agua que se enciman hasta una distancia de corte deseada (en nuestro caso 2.8 Å).

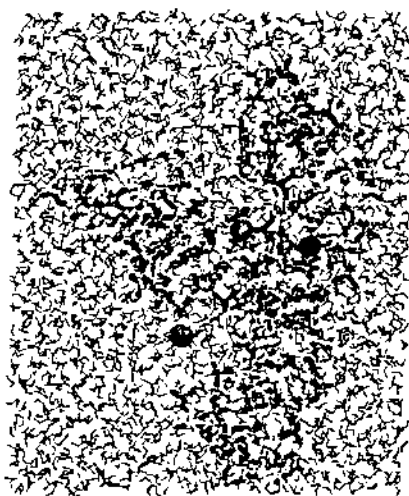


Figura 3.7: Ribozima "Hammerhead" solvatada

### 3.3.3 Minimización

Dada la función de energía potencial, en ocasiones, se requiere de la configuración de energía mínima del sistema. La minimización permite relajar la estructura obtenida de forma experimental, promediar varias estructuras, así como encontrar un mínimo de energía local o global. Para sistemas macromoleculares, el número de mínimos locales es muy grande y es frecuentemente imposible determinar el mínimo de energía global.

Generalmente sólo es posible examinar un mínimo local que se encuentra cercano a la estructura determinada por rayos X.

CHARMM incluye 5 diferentes algoritmos de minimización con el fin de otorgarle mayor flexibilidad a la energía del sistema para una gran variedad de propósitos. Estos métodos se clasifican en no convergentes (descenso más profundo), dos métodos convergentes con aproximaciones armónicas (gradiente conjugado y el método de base adaptada Newton-Raphson o ABNR), el método de segunda derivada explícita (Newton-Raphson) y dinámica de enfriamiento. Los métodos convergentes generalmente terminan cuando han alcanzado un mínimo de energía, los métodos no convergentes nunca terminan, sino que se encuentran navegando sobre toda la superficie de energía. Estos diferentes métodos para minimizar el sistema ya se explicaron en el Capítulo 2. En este caso, el método que más se adecúa al tipo de dinámica y ensamble CPT que se va a utilizar es el de Newton-Raphson.

### 3.3.4 Cristalizar

Cuando se requiere que la presión permanezca constante durante la dinámica, es necesario poner condiciones periódicas a la frontera, esto significa que se colocan cristales cúbicos (imágenes) en cada cara del cristal original (6 en este caso). Esto permite tener la misma cantidad de moléculas en el mismo volumen ya que por cada molécula que se sale de la región, entra otra por el lado opuesto, permitiendo tener entonces presión constante.

### 3.3.5 Dinámica

Las simulaciones dinámicas del sistema se realizan con base a la mecánica clásica e involucra la integración de las ecuaciones de Newton como se describió en el Capítulo anterior. CHARMM resuelve estas ecuaciones de forma numérica para lo cual dispone de dos métodos; el de Verlet y el predicción-corrección.

Debido a que la solución numérica se obtiene de forma secuencial, existe un potencial muy grande de acumular errores de redondeo. Para controlar esto, los algoritmos utilizan variables de doble-precisión para las varias derivadas de las posiciones y para la acumulación de sumas de varios términos. Para incrementar la eficiencia computacional se utiliza aritmética básica de precisión,

es decir, se utilizan sumas en lugar de restas, multiplicaciones en lugar de divisiones, etc.

La precisión de la solución numérica puede ser revisada por diferentes técnicas. En primer lugar, ya que no existen fuerzas externas, la energía del sistema debe conservarse. Los errores numéricos manifestarían en todo caso fluctuaciones en la energía total. Como el calor específico es proporcional a las fluctuaciones de la energía cinética del sistema, las fluctuaciones en la energía total son comparadas con las fluctuaciones de la energía cinética para juzgar la calidad de la solución numérica. En segundo lugar, se puede repetir la dinámica en orden inverso el mismo número de pasos que se dio y la dinámica debería regresar al mismo punto del cual se partió. Este método es más riguroso que el anterior.

Debido a que no hay forma de preparar el sistema en un conjunto de posiciones y velocidades que lo mantengan en equilibrio, los primeros pasos de la dinámica son utilizados para equilibrar el sistema en la temperatura deseada.

Para iniciar la equilibración, a partir del conjunto de coordenadas obtenidas de una estructura por rayos X, el sistema con estas coordenadas se minimiza para eliminar grandes fuerzas debidas a repulsiones por contactos de van der Waals o por geometrías pobremente resueltas. Las minimizaciones se corren por un número pequeño de pasos y en ocasiones se piden constricciones en las posiciones para no permitir que los átomos tengan grandes saltos de sus posiciones originales.

Dadas estas posiciones de equilibración, se asignan entonces las velocidades iniciales para empezar la dinámica. CHARMM permite varias formas de asignar las velocidades; una es asignar velocidad cero a todos los átomos dejando que estos se muevan a partir de las fuerzas restantes, otra es asignar energías cinéticas de forma aleatoria a partir de distribución Gaussiana a una temperatura dada.

Una vez que se tienen las posiciones y velocidades iniciales, el sistema debe colocarse a la temperatura deseada. No existe una forma única de hacer esto. Una forma es asignar las velocidades correspondientes a la temperatura deseada de la dinámica. Esto no es muy conveniente porque introduce una perturbación muy grande en el sistema, que en ocasiones requerirá de un ajuste grande en las velocidades del sistema. La otra forma sería calentar al sistema muy lentamente al asignar primero a las velocidades un valor pequeño que implica un

valor pequeño de temperatura y permitir que la dinámica corra un número de pasos, de nuevo asignar velocidades y así sucesivamente hasta llevar el sistema a la temperatura deseada. Las velocidades en cada caso deben adecuarse a una distribución Gaussiana.

Después de que el sistema es llevado a la temperatura deseada, es necesario en ocasiones dejar que el sistema alcance equilibrio, es decir, hasta que las propiedades estadísticas del sistema (temperatura, fluctuaciones y rms) sean independientes de los periodos temporales para que puedan analizarse.

Cuando el sistema está equilibrado inicia entonces la simulación. El programa CHARMM escribirá las coordenadas y las velocidades en archivos separados a una frecuencia controlada. La trayectoria resultante puede ser analizada mediante algunos programas o paqueterías.

En el curso de la dinámica, los átomos se mueven, así que la lista de interacciones de no enlace debe ser regenerada. El usuario puede especificar a que frecuencia debe ser actualizada cada lista. La elección de las frecuencias se determina por cuánto se mueven los átomos y las funciones de ajuste.

Para un sistema biológico a temperatura de 300K se recomiendan los siguientes valores.

Frecuencia de actualización de no-enlace	8.0 Å
Frecuencia de actualización de enlaces de H	10 pasos (0.010 ps)
Distancia elegida de no-enlace	8.0 Å
Punto final para el conector de no-enlace	7.5 Å
Punto inicial para el conector de no-enlace	6.5 Å
Distancia elegida para el enlace de hidrógeno	5.0 Å
Distancia final para el enlace de hidrógeno	4.5 Å
Distancia inicial para el enlace de hidrógeno	3.5 Å
Ángulo de enlace de hidrógeno elegido	90°
Punto final del ángulo de enlace del hidrógeno	70°
Punto inicial del ángulo de enlace del hidrógeno	50°

Dependiendo de la naturaleza del sistema, estos valores pueden ser modificados.

La siguiente parte consiste en el análisis de los resultados obtenidos. Esto se verá en el capítulo siguiente con los detalles de las Dinámicas Moleculares realizadas.

## Capítulo 4

# Resultados de la Dinámica Molecular clásica

### Introducción

En este capítulo se analizan los resultados de los cálculos de Dinámica Molecular (DM) utilizando uno de los cristales recientemente obtenidos de una ribozima "Hammerhead" [14]. Se colocaron uno o más iones de magnesio en diferentes partes de la molécula y se realizaron simulaciones DM sin constricciones en la estructura inicial con periodos de 200 a 500 ps.

Los resultados obtenidos concuerdan con algunas observaciones experimentales que caracterizan a ciertos sitios donde los átomos de magnesio permanecen asociados. También se logra observar una **Conformación Cercana a la línea de Ataque (CCA)**<sup>1</sup> en algunas partes de la simulación. Estas conformaciones CCA están caracterizadas por el acercamiento de los dos átomos que forman un nuevo enlace,  $PA_{1.1}$  y el  $O2'C_{17}$  hasta una distancia de 3.5 Å y la formación de un ángulo entre los átomos ( $O2' C_{17} - P A_{1.1} - O2' O5' A_{1.1}$ ) mayor a 150° [6]. Durante la DM un magnesio en la posición *VI* ( $Mg_{VI}$ ) se movió inmediatamente a asociarse con el oxígeno pro - R del fosfato ( $A_{1.1}$ ) y el magnesio en la posición *III* ( $Mg_{III}$ ) permaneció en la región donde se supone que se efectúa el rompimiento,  $C_{17}$  y  $A_{1.1}$ . Los resultados fueron la base para un estudio posterior con Mecánica Cuántica Molecular para observar el rompimiento.

---

<sup>1</sup>Near Attack Conformation.



El programa computacional utilizado fue CHARMMv24 [43]. Cada una de las dinámicas se realizó bajo condiciones ambientales diferentes con el fin de explorar la conformación CCA.

Los estudios realizados comprenden el uso de alguna o de todas las siguientes herramientas:

1. Coordenadas Iniciales.
2. Construcción de los enlaces de hidrógenos.
3. Minimización de energía.
4. Solvatación de la molécula y construcción del cristal.
5. Dinámica Molecular.

Estos métodos fueron descritos en los capítulos anteriores y aquí se expondrán los resultados de las dinámicas efectuadas con el fin de estudiar la catálisis de la ribozima "Hammerhead" en presencia de iones de magnesio.

La molécula de ARN cuenta con 1321 átomos (contando los hidrógenos) y en la estructura cristalográfica se encuentran 5 átomos de magnesio. Los átomos de magnesio tienen designados los números I, II, III, IV y VI, según Scott *et al* [14] que se muestran en la figura 4.1.



Figura 4.1: Imagen de las posiciones de los cinco iones de magnesio.

## 4.1 Dinámica de la molécula "Hammerhead" completa sin átomos de magnesio

Para entender los cambios conformacionales que presenta un sistema de muchos cuerpos sujeto a muchas variables como presión, temperatura, pH, etc., en principio partimos del sistema más simple que podemos tener y conforme entendemos el problema le vamos adicionando un mayor número de variables. El sistema que queremos estudiar es una molécula de aproximadamente 1321 átomos y el fenómeno que queremos observar son los cambios conformacionales de esta molécula en presencia de iones de magnesio. Por lo que para poder discernir si un cambio es debido o no a la presencia de un ion metálico, necesitamos un punto de comparación. Este punto de referencia será una dinámica de la molécula completa del "Hammerhead" en ausencia de iones de magnesio pero bajo las mismas condiciones a las que se someterán las otras dinámicas con metales. Es decir, la misma temperatura, presión, número de moléculas de agua y las otras variables de la misma dinámica.

No hay que olvidar que según los datos experimentales, la catálisis de la ribozima existe aún sin la presencia de iones metálicos divalentes, pero esta se desarrolla en tiempos muy largos comparados con aquellos tiempos en presencia de iones divalentes. Los resultados de esta dinámica se mencionan a continuación.

### Datos de la dinámica

La estructura capturada por cristalografía de Rayos X no puede detectar los hidrógenos, así que el primer paso que se da para modelar la molécula es añadir estos enlaces de hidrógeno faltantes. Esto se hace a partir del archivo de topología que tiene integrado CHARMM y que es capaz de reconocer los átomos como nucleótidos y con datos teóricos asignar la mejor posición para estos hidrógenos. A continuación se le agregan las moléculas de agua como se describió en el capítulo anterior bajo condiciones CPT<sup>2</sup> en donde se tuvieron los valores de 1 atm y 300 K para la presión y la temperatura respectivamente.<sup>3</sup> La siguiente parte consiste en formar un cristal para darle libertad de movimiento y conservar el número de moléculas. En esta parte del programa contamos con aproximadamente 9000 átomos.

---

<sup>2</sup>Presión y temperatura constantes.

<sup>3</sup>Es decir, a 27° Celsius, que es la aproximadamente la temperatura de un sistema biológico.

Antes de que la molécula entre en la dinámica esta debe encontrarse relajada, para lo cual se realiza una minimización de energía de 30 pasos del método ABNR<sup>1</sup> y a continuación inicia propiamente la dinámica. Los datos de la dinámica se pueden ver en el apéndice I y la gráfica de la energía se muestra en la figura (A) del apéndice II.

## Resultados

En primer lugar se analiza la zona catalítica para ver si existe algún cambio en la conformación que no sea una reacción que involucrara iones de magnesio.

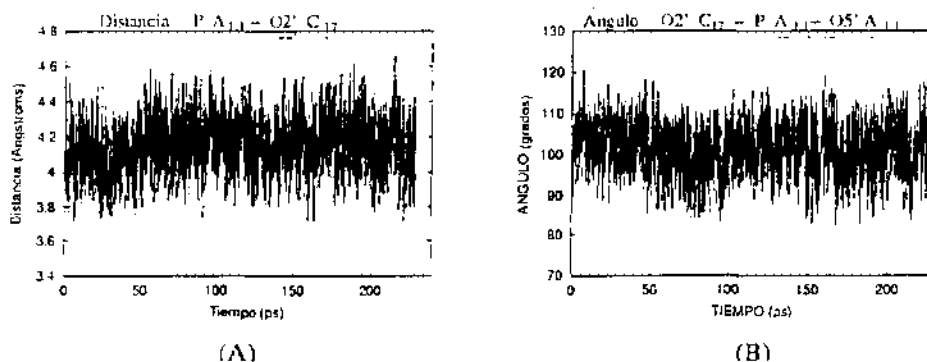


Figura 4.2: Dinámica sin metales: (A) distancia relativa entre  $P-O_2'C_{17}$ , (B) ángulo de la conformación CCA  $O_5'A_{1.1} - PA_{1.1} - O_2'C_{17}$ .

## Distancias y ángulos

Las distancias más importantes de la zona catalítica se presentan en la figura 4.2, estas son la distancia entre el fósforo de la citosina ( $PC_{17}$ ) y el oxígeno de la ribosa de este mismo nucleótido ( $O_2'C_{17}$ ) ya que es el enlace que se forma después de la catálisis. La distancia promedio observada está en el intervalo de 3.8 – 4.4 Å. Es decir los átomos se encuentran muy lejos todavía de tener interacciones para formar un nuevo enlace. La otra gráfica presenta el ángulo CCA formado por los átomos  $O_2'C_{17} - PA_{1.1} - O_5'A_{1.1}$ . Lo que observamos es que este ángulo está muy lejos todavía de formar por lo menos  $150^\circ$ . En promedio permanece en el intervalo  $90-120^\circ$ . Lo cual es un valor muy pequeño para permitir un cambio conformacional que lleve al fósforo cerca del nuevo enlace.

<sup>1</sup>Base adaptada de Newton-Raphson, ver sección 2.6.4.

Los pasos que ejecuta la dinámica puede consultarse en el Apéndice I, ahí están los parámetros y el algoritmo de integración que en este caso es el algoritmo "Leap-Frog" descrito en el capítulo 2. El paso de integración fue de 0.002 ns y en total se realizaron 100,000 pasos, con lo que el tiempo total de la corrida fue de 200 ps. Las velocidades iniciales se asignaron de forma aleatoria de acuerdo a la temperatura. Cada 10 pasos escribe las coordenadas de la trayectoria.

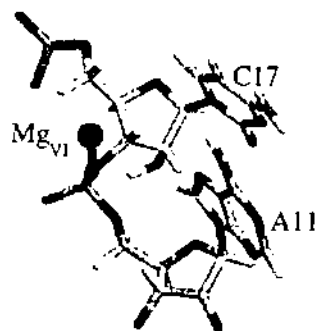


Figura 4.4: Zona catalítica después de 200 ps con  $Mg_{VI}$

## Resultados

En la figura 4.4 se observa la imagen final de la posición de los átomos de la región catalítica. Esta región está formada por el nucleótido citosina 17 ( $C_{17}$ ) y la adenina 1.1 ( $A_{1.1}$ ). La imagen es representativa de la conformación consistente durante toda la dinámica. El magnesio se acerca al oxígeno  $O2PA_{1.1}$  y se mantiene en esa posición durante toda la dinámica. Los cambios en las distancias y ángulos de otros átomos se enuncian en los siguientes párrafos.

Para analizar los resultados de la dinámica, lo primero que se debe observar es la estabilidad del sistema, para lo cual se estudia la gráfica de energía contra tiempo, figura (B) del Apéndice II. En esta gráfica lo que podemos ver es que la energía se conserva, no se han graficado los 10 ps del inicio, debido a que en esta etapa el sistema se está estabilizando y la energía oscila mucho. Los resultados de la dinámica son tomados en cuenta solo donde la energía aparece de manera constante.

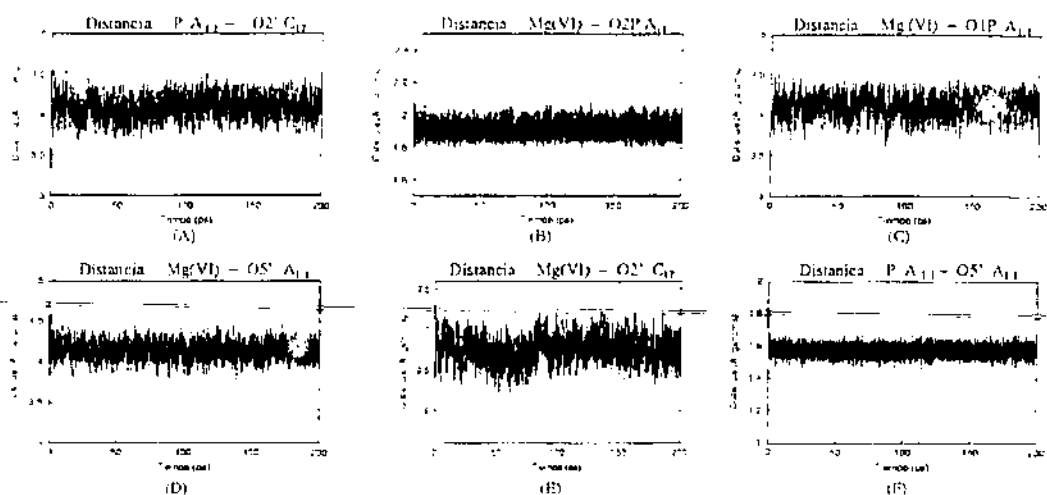


Figura 4.5: Dinámica con un metal,  $Mg_{VI}$ ; distancias relativas

## Distancias y ángulos

La distancia que nos interesa principalmente es la del enlace que se debe formar de acuerdo a los mecanismos de catálisis y el desarrollo de una conformación previa a la catálisis. La primera gráfica que observamos, figura 4.5(A), muestra que esta distancia se mantiene en promedio en 4 Å de separación. Es decir, hay un acercamiento debido a la presencia del magnesio, (en la dinámica sin  $Mg(VI)$  la distancia es de 4.2 Å.) En la gráfica 4.5(B) y 4.5(C) observamos las distancias entre el magnesio y los oxígenos  $O1P$  y  $O2P$  del fósforo de  $A_{1.1}$ . Aquí vemos que el magnesio se une al  $O2P$  y se mantiene a una distancia de entre 1.8 y 2 Å. Esta unión se conserva durante toda la dinámica, lo cual nos permite decir a que parte del ARN se une el magnesio en esta posición. La gráfica 4.5(D) muestra la distancia entre el magnesio y oxígeno perteneciente al enlace que se quiere romper, la distancia oscila entre 3.8 – 4.4 Å, por lo que no se observa un acercamiento notorio. La gráfica 4.5(E) muestra que tan separado se encuentra el magnesio del  $O2'C_{17}$  perteneciente al nuevo enlace. Este oxígeno debe perder un hidrógeno antes de unirse al fósforo según los mecanismos de reacción. La última gráfica muestra el enlace que se debe romper, en la cual no se observa indicio que muestre la intención de la catálisis.

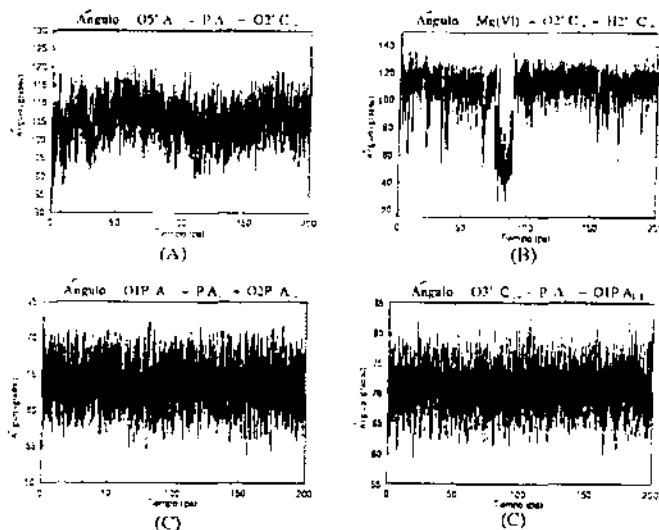


Figura 4.6: Dinámica con un metal.  $Mg_{VI}$ : ángulos relativos

El ángulo de la conformación cercana a la línea de ataque (CCA) se muestra en la figura 4.6(A), el cual está formado por  $O2'C_{17} - P A_{1,1} - O5' A_{1,1}$ . En la figura 4.6, se muestran los valores de algunos ángulos durante la dinámica. Vemos que no existe cambio significativo en estos valores, a excepción del hidrógeno cuya posición oscila debido a que es un elemento ligero atraído por los átomos cercanos.

## Conclusiones

Los resultados de DM con un átomo de magnesio en la posición VI no presentan un cambio significativo en la conformación hacia una línea de ataque previa a la catálisis. Sin embargo, todavía no podemos descartar que sólo sea necesario un átomo de magnesio pues necesitamos explorar otros sitios cercanos a la región de rompimiento en los cuales pueda tener un papel importante la presencia del metal. Debido al gran costo computacional de las dinámicas (cada una se tarda 2 semanas en el equipo de Supercómputo de DGSCA, "bernice8"), se decidió simplificar el sistema y considerar sólo los átomos de la región catalítica y comparar los resultados con una dinámica completa. Estos resultados se explican en las dos siguientes secciones.

### 4.2.2 Posición VI, dinámica corta sin cristal

Esta dinámica se realizó con sólo 8 nucleótidos cercanos a la región catalítica y fueron:

Cadena a:  $G_{2.2}$   $U_{2.1}$   $C_3$   $U_4$

Cadena b:  $U_{16.1}$   $C_{17}$   $A_{1.1}$   $C_{1.2}$

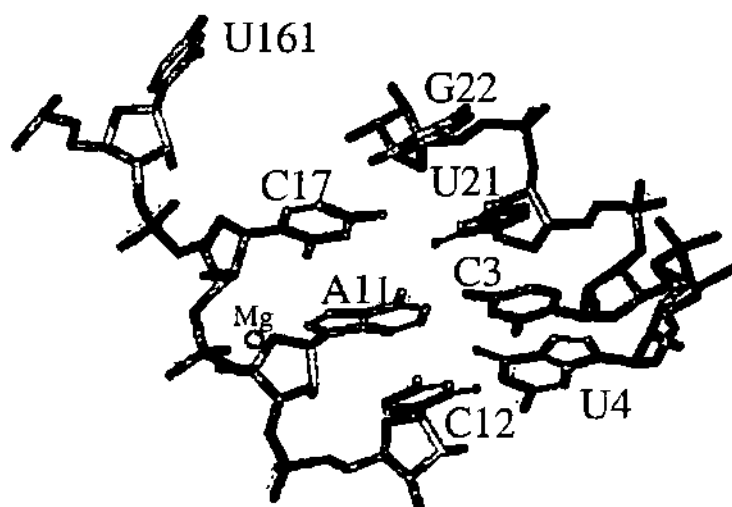


Figura 4.7: Imagen de las posiciones iniciales del modelo corto del "Hammerhead".

La dinámica se realizó con el fin de conocer qué factores pueden marcar diferencia entre una dinámica con toda la molécula y una dinámica con sólo la región de interés, por lo que esta dinámica se realizó con sólo 8 nucleótidos y sin el cristal, lo que significa que no tuvo condiciones de presión constante.

La justificación de esta aproximación se basa en el hecho de que en la molécula completa de ARN vemos que hay partes que tienen muy poca influencia en la región catalítica. El átomo de magnesio sólo puede interactuar con algunos átomos que están a su alrededor. Este tipo de aproximación también nos permite realizar varias dinámicas en muy poco tiempo, lo cual nos facilitará alterar a la molécula con el fin de obtener más información acerca de su rigidez y actividad.

### Datos de la dinámica

Los datos del programa de entrada se presentan en el Apéndice I. A diferencia de las otras dinámicas, en este programa no se pide la formación de un cristal (es decir, se omite la instrucción 50 del código de entrada de CHARMM, apéndice I), por lo que la molécula no está restringida a permanecer en un cubo y por lo tanto no hay presión constante. Las moléculas de agua de los alrededores se separan considerablemente de la molécula.

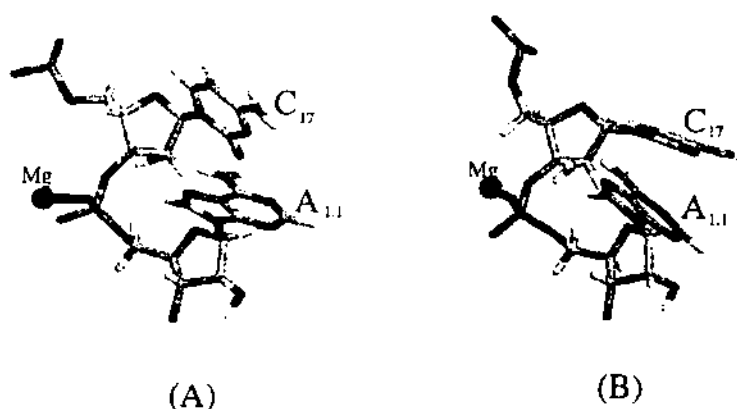


Figura 4.8: Imagen de la zona catalítica en el modelo corto sin cristal: (A) después de una minimización de 30 pasos, (B) después de la dinámica de 200 ps.

### Resultados

La figura 4.8 muestra la zona catalítica minimizada y al final de la dinámica. Se aprecia un gran cambio en la configuración, vemos que las bases están rotadas respecto a su configuración inicial, además de que los átomos de los extremos tendieron a rotar y separarse de la molécula. Las condiciones que no restringen el volumen constante, afectan de manera considerable la estabilidad de la molécula y por lo tanto su actividad biológica puede cambiar notablemente.

En la figura (F) del Apéndice II tenemos la gráfica de la energía durante la dinámica. La energía no es la misma que en la dinámica completa. Además se ve como la energía no permanece estable alrededor de un punto sino que presenta bastantes fluctuaciones. Esta gráfica nos permite concluir que este sistema mucho más pequeño y con menos constricciones es inestable.



### Distancias y ángulos

En las gráficas de las distancias, figura 4.9, vemos muy pocos cambios con respecto a la dinámica con la molécula de completa. Lo que se observa si seguimos la trayectoria de la dinámica es que las moléculas de agua tienden a evaporarse, es decir, tienden a alejarse por fuerzas repulsivas y por la energía cinética que tienen. A pesar de que las distancias de los átomos internos son muy parecidas, esto es solamente durante una parte de la dinámica, ya que si le diéramos más tiempo, las ramas tenderían a separarse.

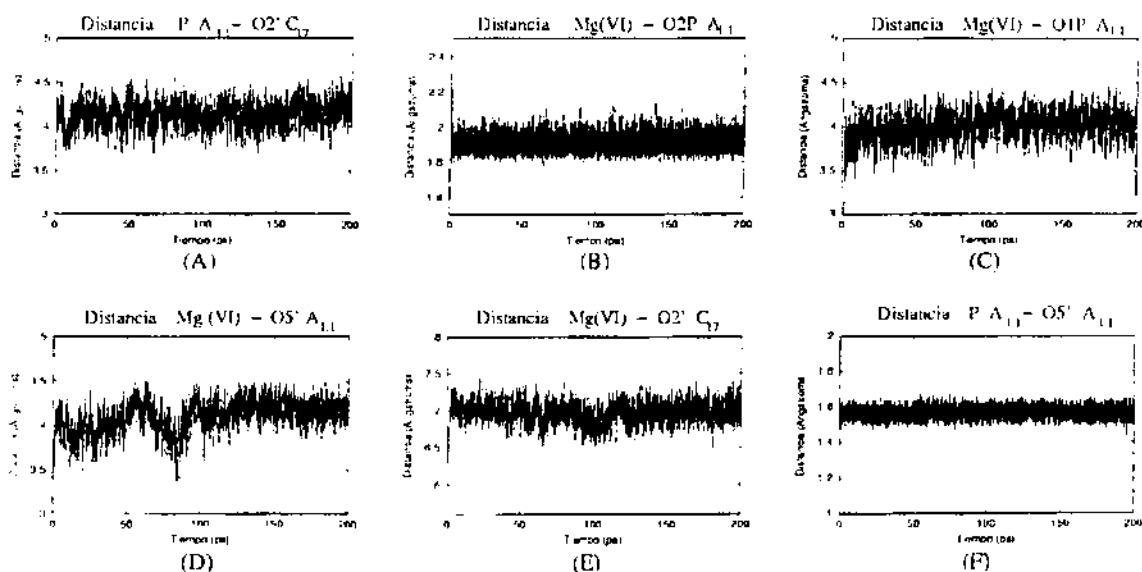


Figura 4.9: Dinámica del modelo corto sin cristal,  $Mg_{VI}$ : distancias relativas.

Igualmente, los ángulos presentados en la figura 4.10, tendieron a mantenerse constantes con excepción del ángulo formado por  $O5' A_{1.1} - P A_{1.1} - O2' C_{17}$  de la primera gráfica, en donde se ve que hacia el final de la dinámica el ángulo entre ellos tiende a ser mayor porque los átomos se van separando, esto es contrario a lo que se quiere, que es una unión entre estos dos átomos. Vemos que el comportamiento de los Hidrógenos sigue siendo casi igual, todos oscilando por las interacciones con átomos vecinos.

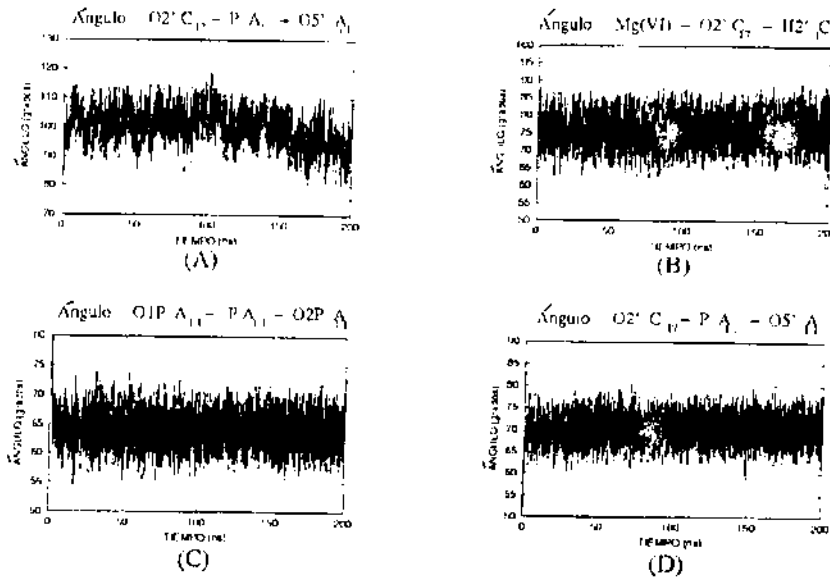


Figura 4.10: Dinámica del modelo corto sin cristal,  $Mg_{VI}$ : ángulos relativos.

## Conclusiones

Con esta dinámica podemos darnos cuenta de varios detalles que son importantes cuando se modela un sistema biológico. El ambiente normal en el que se encuentra la molécula que estamos estudiando es dentro de la célula, es decir, se encuentra rodeada de agua y otros elementos, además de encontrarse a una cierta presión y temperatura. Este tipo de elementos deben ser tomados en cuenta si se quiere simplificar el problema para cuantificar el grado en que nuestro modelo deja de ser adecuado. Por lo que este modelo podría considerarse malo por no preservar estos lineamientos, sin embargo, existen varias justificaciones para haberlo llevado a cabo. En primer lugar, nos da una idea de la importancia de los nucleótidos que se encuentran alejados. A primera vista se pensaría que no influyen, sin embargo, vemos que esto no es obvio hasta después de haber realizado los cálculos. En segundo lugar, al sistema se le brindó una mayor libertad, los átomos tenían menos restricciones de espacio y podían colocarse más rápido en los sitios a los cuales eran más afines. Además el sitio entre nucleótidos que debía presentar la catálisis tuvo una mayor abertura, permitiendo que los átomos que debían tener un arreglo conformacional lo hicieran de forma más rápida al no encontrar obstáculos o repulsiones de

otros átomos. Por último, el sistema es mucho más pequeño, lo que permite repetir las dinámicas cortas de forma fácil y con la oportunidad de estudiar un mayor número de variables. Todo esto nos mostró que valía la pena hacer un segundo estudio con la molécula pequeña, pero cambiando las condiciones que lo hacen ver irreal, es decir, la condición de presión constante. Con esto se evitaría que los nucleótidos finales tendieran a alejarse y que la molécula perdiera su conformación. Se realizó esta dinámica con presión constante y los resultados-obtenidos se-presentan a continuación.

### 4.2.3 Posición VI, dinámica corta con cristal

#### Datos de la dinámica

Esta dinámica se realizó igualmente con los 8 nucleótidos de la figura 4.7, que incluyen a los más cercanos a la zona catalítica. La diferencia con la dinámica anterior fué que se construyó un cristal para mantener las condiciones periódicas, es decir, presión constante. El cristal fué mucho más pequeño que en la dinámica completa por lo que se tuvo un considerable ahorro de tiempo computacional.

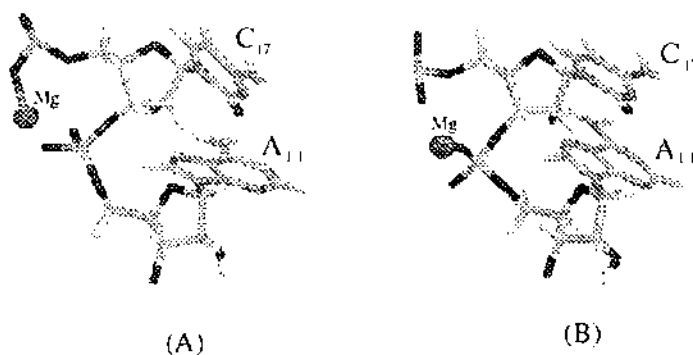


Figura 4.11: Imagen de la zona catalítica en el modelo corto con cristal: (A) después de una minimización de 30 pasos. (B) después de la dinámica de 200 ps.

#### Resultados

En la figura 4.11, se tienen las imágenes de la dinámica, esta se considera mucho más estable que la anterior, aunque también se ve como los nucleótidos de los extremos tienden a rotar separándose de su configuración como molécula

completa. Sin embargo, la zona catalítica presentó la misma conformación por lo que se consideraron como válidos estos cálculos debido a la similitud de las gráficas que se presentan a continuación.

En la figura (G) del Apéndice II, tenemos la gráfica de la energía contra tiempo. Si le quitamos el periodo inicial correspondiente a la estabilización del sistema, vemos que esta permanece constante en la dinámica. Esta condición es fundamental para la validez de los resultados. De nuevo se observa que el valor promedio no corresponde al valor de la dinámica del modelo con la molécula completa debido a que se tiene un número mucho menor de átomos.

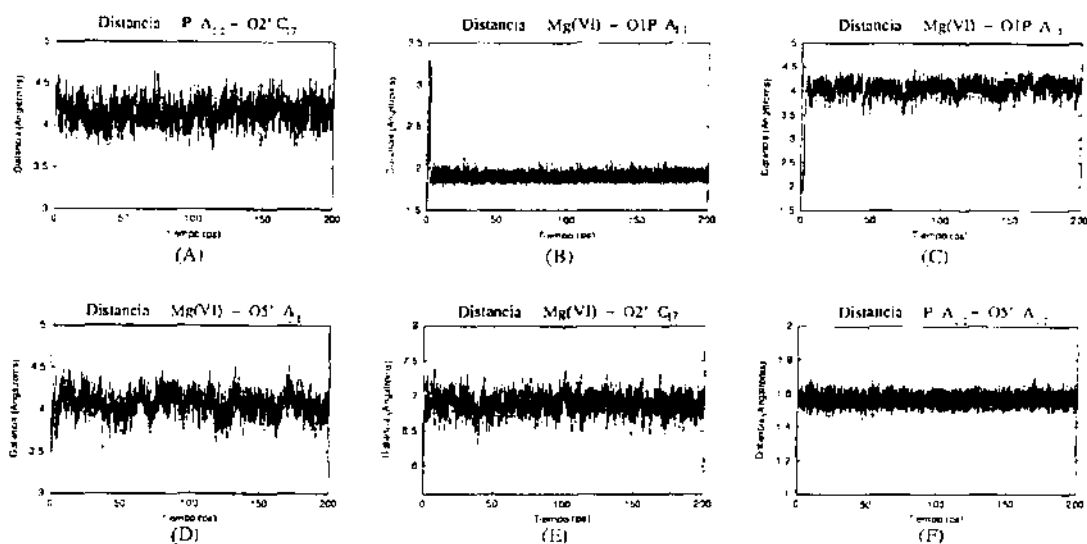


Figura 4.12: Dinámica del modelo corto con cristal,  $Mg_{VI}$ : distancias relativas.

### Distancias y ángulos

En la figura 4.12 se tienen las gráficas de distancias, se tiene aproximadamente el mismo valor promedio que en las gráficas de la figura 4.5. Como se tiene un número menor de nucleótidos es de esperarse que la molécula sea más flexible ya que hay menos interacción o fuerzas que la mantengan rígida.

En la figura 4.13 se muestran los mismos ángulos que en las dos dinámicas anteriores. Como se puede apreciar, no hay variación en estos que nos indiquen una transición hacia una conformación previa a la catálisis de la misma manera como pasa en la dinámica completa con un metal en la misma posición.

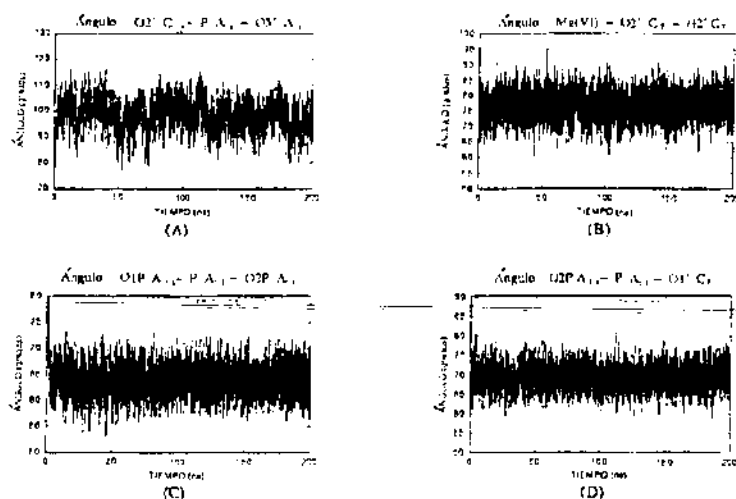


Figura 4.13: Dinámica del modelo corto con cristal,  $\text{MgVI}$ : ángulos relativos.

## Conclusiones

Con base en los datos mostrados en esta dinámica, podemos decir que la estructura corta se aproxima a la dinámica de la molécula completa siempre y cuando se respeten las condiciones de cristal, esto no permite que las moléculas de agua o las de los nucleótidos tiendan a escaparse y de esa forma desnaturar la molécula. Con base en éstas dos muestras se decidió continuar con las dinámicas con sólo 8 nucleótidos con condiciones de cristal. Sin embargo, a pesar de que se realizó un número considerable de corridas con diferentes variables, como posición del magnesio y temperatura, no se obtuvo una conformación cercana a la catálisis. Por lo que se decidió continuar con dinámicas de la molécula del "Hammerhead" completa y explorar la zona catalítica con otras posiciones del magnesio. Ahora lo que se intenta observar es la importancia de que los otros nucleótidos estén presentes para que se realice la catálisis. Esto de acuerdo a lo explicado en el capítulo 1, sobre la necesidad de que ciertos nucleótidos estén presentes para que se lleve a cabo el rompimiento. La siguiente dinámica modela todo el ARN pero colocando al magnesio en una posición diferente.

## 4.2.4 Posición III

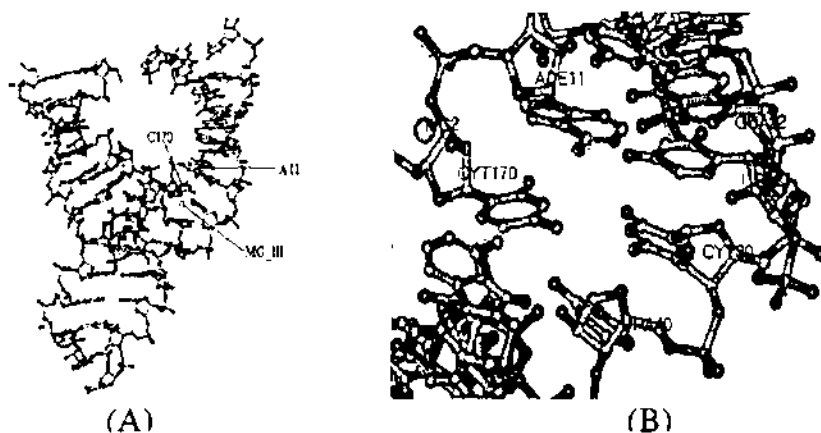


Figura 4.14: (A) Imagen de la posición inicial del  $Mg_{111}$ , (B) acercamiento a la zona catalítica con  $Mg_{111}$ .

Esta dinámica se realizó con toda la molécula de ARN "Hammerhead" con un átomo de magnesio en la posición III, figura 4.14(A) y un acercamiento a esta posición se observa en la figura 4.14(B). Este átomo no es el más cercano al sitio de rompimiento, sin embargo, su posición es muy cercana a las bases de los nucleótidos que presentan rompimiento. Con base en esto se supuso que en esta posición el átomo de magnesio podría influir en el cambio conformacional para que se presente la catálisis. Las distancias del  $Mg_{111}$  a algunos átomos fueron:

$O2'C_{17}$	2.8 Å
$O3'C_{17}$	4.7 Å
$O5'A_{1.1}$	5.6 Å

#### Datos de la dinámica

Al igual que en las otras corridas, a esta estructura se le aplica primero una minimización de energía consistente en 30 pasos del método ABNR. La molécula se colocó también en un cubo de agua con condiciones de presión constante y temperatura de 300 K.

Los pasos que ejecuta el programa se pueden consultar en el Apéndice I, que son similares a los ejecutados por la dinámica sin metales pero con un metal en la posición VI (ver instrucción 26 en el Apéndice I). El número de pasos

ejecutado por la dinámica fué mayor, hasta llegar a los 260 ps.

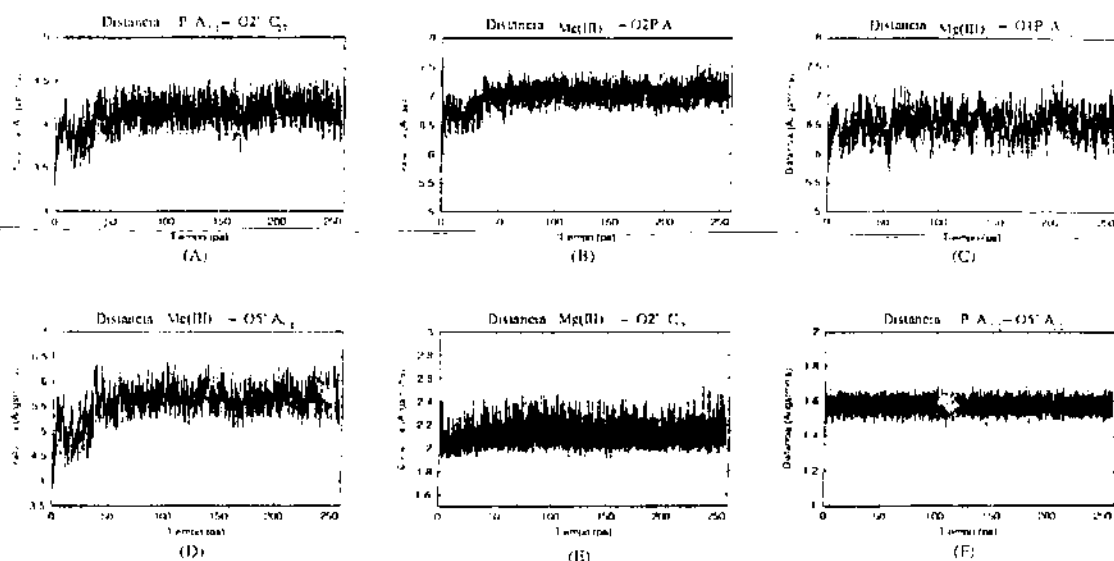


Figura 4.15: Dinámica con un metal,  $Mg_{II}$ : distancias relativas

### Distancias y ángulos

La primera gráfica que observamos en la figura 4.15 es la que corresponde a los átomos que se deben acercar para formar un nuevo enlace:  $O2'C_{17} - P A_{1,1}$ . En esta gráfica vemos que la tendencia inicial fué de mantener una distancia promedio mayor a los 3.5 Å aunque consistentemente en la dinámica muestra una distancia de 4 a 4.5 Å. Esta distancia es muy grande para que los dos átomos interactúen. En los oxígenos  $O1P$  y  $O2P$  del  $A_{1,1}$ , existe carga negativa que podría influir en el comportamiento del magnesio, sin embargo, en las gráficas 4.15(B) y 4.15(C), vemos que la distancia es mayor a 6 Å, lo cual indica que la fuerza de atracción entre estos átomos y el magnesio es muy pequeña, aunque sí se logra observar que el átomo  $O1P A_{1,1}$  presenta el mismo tipo de variaciones que en la dinámica con el  $Mg_{VI}$ , figura 4.5.

También observamos que el magnesio se aleja del  $O5' A_{1,1}$ , indicando que el metal es atraído por otros átomos cercanos. Vemos que uno de estos aparece en la gráfica 4.15(E), donde se observa que el magnesio se enlaza al  $O2' C_{17}$  manteniendo una distancia promedio de 2.2 Å consistentemente durante toda

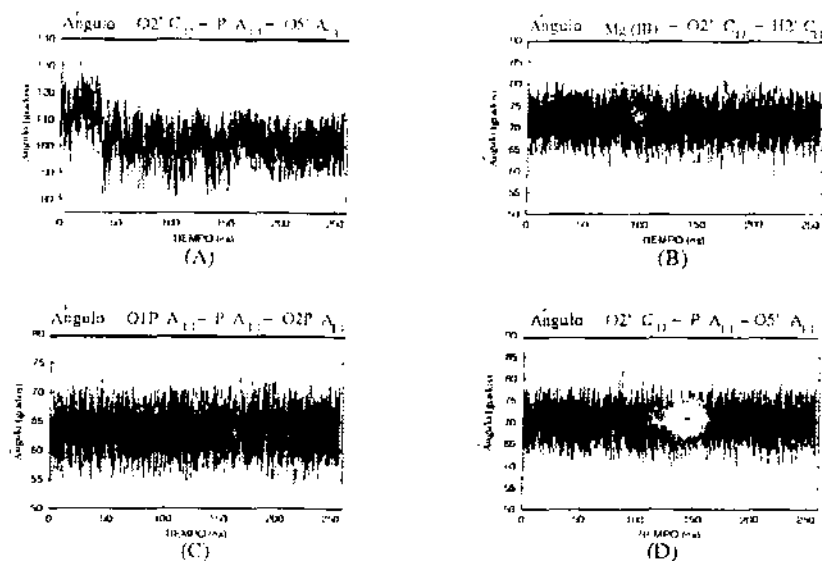


Figura 4.16: Dinámica con un metal,  $Mg(II)$ : ángulos relativos

la dinámica. La formación de este enlace está también caracterizado porque el hidrógeno unido a este oxígeno, el  $H2'C_{17}$ , presenta una menor atracción formando un puente de hidrógeno con el  $O3'C_{17}$  donde se observa una distancia promedio de 2.2 Å durante toda la dinámica. Al inicio existe la formación de un enlace con el  $O5'A_{1,1}$ , sin embargo, el enlace se rompe casi inmediatamente. Cabe señalar que este hidrógeno podría ser el que se une al  $O5'A_{1,1}$  para saturarlo cuando se da la catálisis, aunque también podría realizarlo cualquier otro hidrógeno suelto ya sea de la molécula de ARN o de una molécula de agua.

El enlace que se quiere romper, permanece igual que en todas las dinámicas a una distancia promedio de 1.6 Å.

Con respecto a los ángulos, observamos en la figura 4.16 (A), que debido al alejamiento entre el fósforo  $P A_{1,1}$  y el oxígeno  $O2' C_{17}$ , el ángulo de la línea de ataque mostró una tendencia a disminuir; bajó de  $120^\circ$  a un promedio de  $100^\circ$ , esto es contrario a lo que se requiere para observar un rompimiento (por lo menos con umbral de  $150^\circ$ ). Los otros ángulos también indican que no hubo ningún cambio conformacional hacia la estructura CCA.



## Conclusiones

Con base en lo anterior podemos concluir de forma general que el átomo de magnesio en la posición III por si solo no lleva a la formación de una estructura CCA. De hecho se adhiere a los oxígenos  $O2'C_{17}$  y  $O2C_{17}$ . Este primer átomo participa en la catálisis y al estar unido a este magnesio pierde totalmente su flexibilidad y no puede interaccionar con ningún otro átomo. Por lo que seguramente un átomo de magnesio en esta posición no favorece la catálisis.

Las dinámicas anteriores exploran la zona catalítica con el átomo de magnesio solamente en la posición VI, sin embargo, falta explorar otras posiciones cercanas a esta zona, por lo que se decidió realizar otra dinámica utilizando una posición distinta para el átomo de magnesio [14]. Esta posición es la III en la figura 4.1 y la dinámica que se presenta a continuación.

## 4.3 Dinámicas de la molécula completa con dos átomos de magnesio

En la figura 1.17 se mostraron las dos posibles rutas para el mecanismo de reacción de la ribozima "Hammerhead". En ella tenemos la ruta con un ion metálico y con dos iones metálicos. En esta sección se estudiará si en el cambio conformacional del "Hammerhead" en dirección a un rompimiento se requieren dos iones metálicos. Como se ve en la figura, un ion metálico se adhiere al oxígeno  $O2PA_{1.1}$  y el otro al oxígeno  $O2'C_{17}$ . Así que los iones que debemos escoger para la simulación deberán ser los más cercanos al sitio de rompimiento y a estos átomos. Viendo la figura 4.1 de este capítulo donde se presentan todos los átomos de magnesio que nos proporciona el análisis experimental, tenemos que los posibles iones a elegir se encuentran en las posiciones VI y III, los resultados los se presentan a continuación.

Hay que mencionar que al colocar dos iones metálicos  $Mg^{2+}$  aumentará la complejidad del sistema, debido a que el magnesio es un átomo grande y con afinidad a formar complejos tetraédricos. Es decir, los lugares en donde el magnesio se colocaría en una posición estable son aquellos en los que pueda asociarse con 6 átomos de oxígeno. Estos átomos de oxígeno pueden provenir tanto de la cadena de ARN como de moléculas de agua. Al colocarse en esta conformación creará cambios en la distribución de los átomos de dos maneras: una es adhiriéndose el átomo de oxígeno o adhiriéndose a dos o más átomos de

oxígeno compartiendo el enlace entre estos. El primer cambio será fijo, pero la segunda forma le propiciará una mayor flexibilidad a la molécula que le permitirá tener otro tipo de interacciones. Al colocar dos átomos de magnesio también se provocará mayor repulsión entre ellos y también entre los átomos con los que forme enlaces. Todo esto nos lleva a estudiar con más detalle el papel de dos átomos de magnesio en el rompimiento. Es posible que para alcanzar cambios conformacionales más relevantes influyan directamente dos átomos de magnesio, sin embargo, estos no pueden encontrarse tan cerca uno de otro, por lo que tal vez en la catálisis sólo intervenga un átomo de forma realmente directa. Esto lo veremos en las simulaciones con Mecánica Cuántica en las cuales si se puede observar rompimiento de enlaces y así determinar si la posición del magnesio influye o no en el rompimiento. Estas dinámicas se explican al final del capítulo.

### 4.3.1 Posición III y VI

En esta dinámica intervienen dos átomos de magnesio los cuales se encuentran en la posición *VI*, que es el átomo cercano al fósforo, en la posición *III*, que es el átomo cercano a las bases. En la figura 4.17(A), se observa la posición inicial de estos átomos y un acercamiento a la zona se muestra en la figura 4.17(B). Aquí vemos también las distancias iniciales del magnesio a los oxígenos más cercanos con los que podría interactuar.

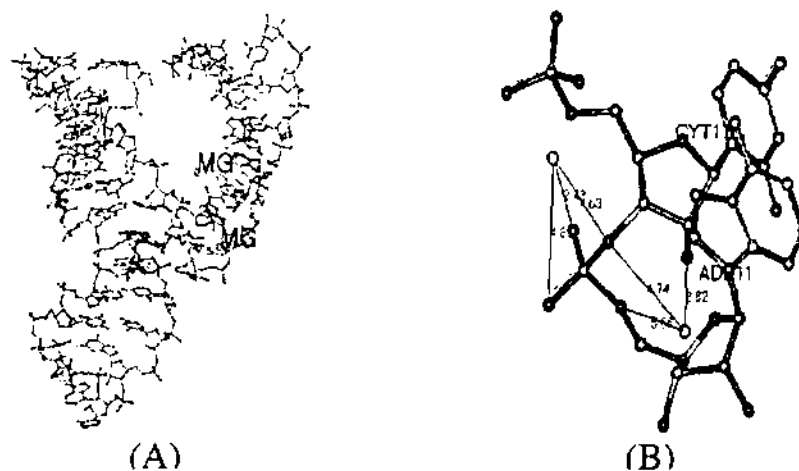


Figura 4.17: (A) Imagen de las posiciones iniciales con dos magnesios,  $Mg_{VI}$  y  $Mg_{III}$ , (B) Acercamiento a la zona catalítica con  $Mg_{VI}$  y  $Mg_{III}$ .

Las distancias entre los átomos de oxígeno más cercanos a los magnesios y los magnesios en la estructura cristalográfica son:

$Mg(III) - O3'C_{17}$	4.63 Å
$Mg(III) - O1P_{A_{1,1}}$	2.43 Å
$Mg(III) - O2P_{A_{1,1}}$	4.83 Å
$Mg(VI) - O3'C_{17}$	4.74 Å
$Mg(VI) - O2'C_{17}$	2.82 Å
$Mg(VI) - O5'A_{1,1}$	5.66 Å

### Datos de la dinámica

En el Apéndice I, tenemos los datos de la dinámica con dos metales. Es totalmente semejante a las anteriores salvo el número de átomos en el cual el archivo "metals.pdb" fue modificado para incluir dos iones de magnesio. La temperatura fue de 300 K y la molécula estuvo sumergida en un cubo de agua a presión constante. En este caso se decidió darle un número de pasos mayor a la dinámica, el cual fue de 300 ps para explorar la posibilidad de que los cambios conformacionales que se investigan requieran de un tiempo mayor.

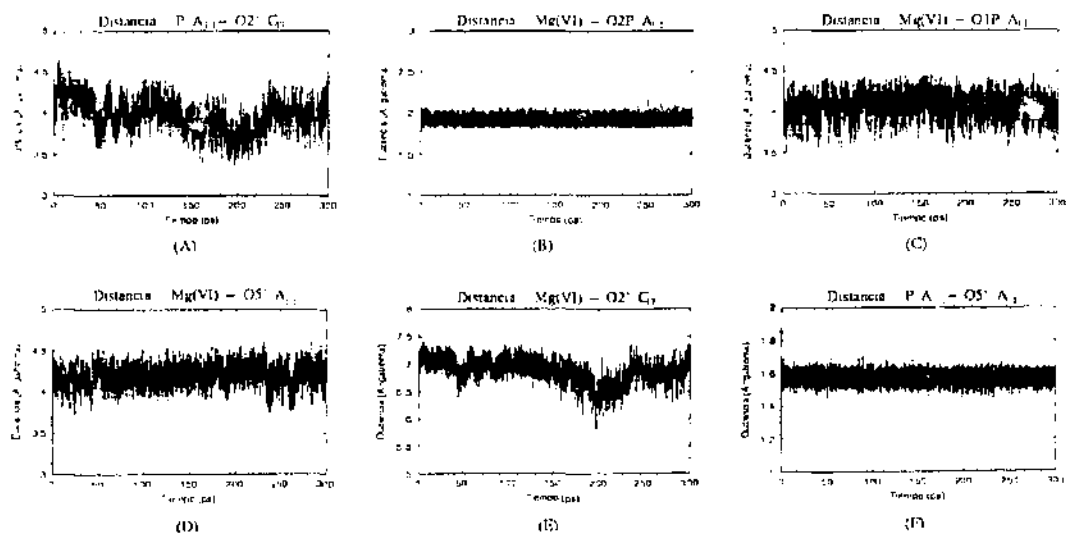


Figura 4.18: Dinámica con dos metales,  $Mg_{VI}$  y  $Mg_{III}$ : distancias relativas.

## Resultados

A diferencia de las dinámicas anteriores, esta dinámica comprendió hasta 300 *ps* en lugar de 200 *ps*. Realizando una rápida observación en las gráficas de las distancias, se logra observar como a tiempos mayores a los 200 *ps*, la estructura todavía presenta cambios significativos en su conformación que de haber utilizado un tiempo menor no se observarían. En este caso se optó por un tiempo mayor debido a que el sistema presenta más variables por lo que sería más difícil que alcanzara un equilibrio estable, ya que el sistema tiene un comportamiento más desordenado. La gráfica de la energía se presenta en la figura (D) del Apéndice II. Veamos los resultados obtenidos con estas dos posiciones.

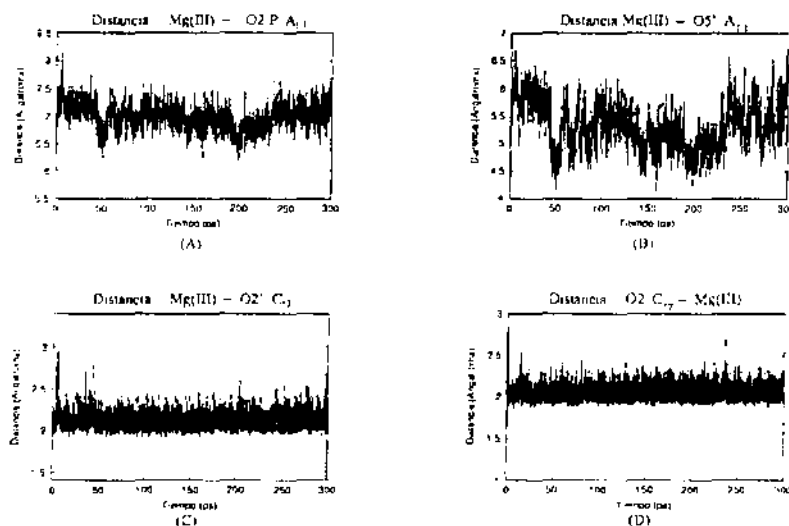


Figura 4.19: Dinámica con dos metales,  $Mg_{VI}$  y  $Mg_{III}$ : distancias relativas.

## Distancias y ángulos

En la figura 4.18 y 4.19 se encuentran las gráficas de distancias. Tenemos primero, la distancia entre el  $O_2C_{17}$  y el  $PA_{1,1}$ , aquí observamos, a diferencia de las dinámicas anteriores, una tendencia a que esta distancia alcance los 3.5 Å en el periodo de tiempo comprendido entre 160 y 210 *ps*. Esto es un cambio notorio con respecto a la dinámica con un metal, debido a que se observa mayor movimiento de estos átomos lo que en algún momento podría

determinar un acercamiento entre ellos, adecuado para la formación de un nuevo enlace. Vemos que el  $Mg_{VI}$  sigue unido al  $O2PA_{1,1}$  consistentemente en toda la dinámica, aunque también observamos que un pequeño intervalo de tiempo, este se acerca al  $O5'A_{1,1}$  hasta encontrarse a 3.5 Å. En la misma figura, 4.18(E), observamos que este mismo magnesio,  $Mg_{VI}$ , se mantiene alejado del  $O2'C_{17}$ . La razón de este alejamiento tal vez sea que este oxígeno se une al magnesio  $Mg_{III}$ , figura 4.18(C), y debido a la repulsión entre magnesios, el oxígeno se mantiene alejado del  $Mg_{VI}$ . El  $Mg_{III}$  también permanece unido al  $O2'C_{17}$ , que es un oxígeno de la citosina. Mientras que su interacción con otros oxígenos es casi nula como el  $O2PA_{1,1}$  o el  $O5'A_{1,1}$ , que están coordinados al otro magnesio. No se observa cambio en el enlace a romper, el  $PA_{1,1}-O5'A_{1,1}$ .

De acuerdo con la descripción de distancias, vemos que cuando hay un mayor acercamiento entre el  $O2'C_{17}$  y el  $PA_{1,1}$ , figura 4.18(A), es cuando mayor aparece el ángulo CCA entre el  $O2'C_{17}-PA_{1,1}-O5'A_{1,1}$ , figura 4.19(A). Los otros ángulos que podrían indicar un cambio conformacional permanecen constantes.

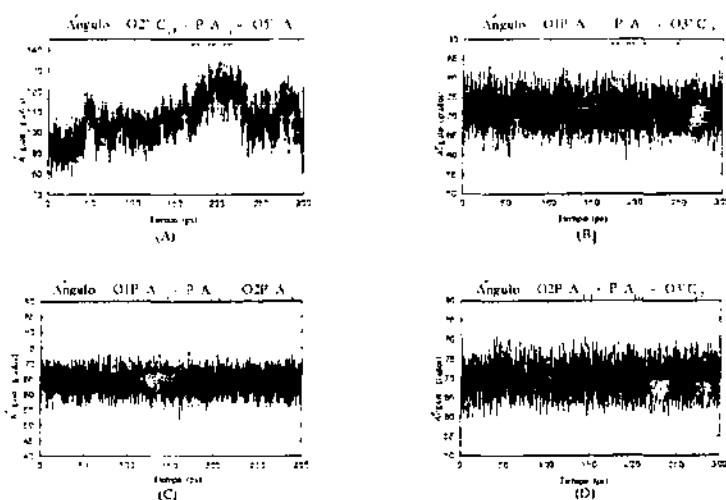


Figura 4.20: Dinámica con dos metales,  $Mg_{VI}$  y  $Mg_{III}$ : ángulos relativos.

## Conclusiones

Con base en la discusión anterior, podemos concluir que esta dinámica arroja más resultados positivos a favor del mecanismo propuesto para la catálisis con dos metales. El problema es saber si en esta misma dinámica se podría observar una mejor conformación para la catálisis si esta se realiza en un periodo temporal más largo. Cabe mencionar que esta dinámica se realizó durante 300 ps, es decir 50000 pasos de integración más que las dinámicas anteriores. Esto se hizo con el fin de observar durante más tiempo el comportamiento del sistema pero no se cuenta en estos momentos con la posibilidad de realizar dinámicas por periodos más largos, debido principalmente a que ocupan grandes cantidades de tiempo computacional y espacio en disco. Además de que un sistema con este número de variables tiende a presentar un comportamiento que no converge, es decir, tiende a seguir cambiando de acuerdo a como se aumente el tiempo sin apreciarse una conformación fija, debido a que el sistema real debería romperse y no oscilar entre diferentes posiciones.

La siguiente dinámica explora otras posibles conformaciones del "Hammerhead" con los mismas posiciones de los magnesios pero con mayor energía. Para esto se le colocó a una temperatura de 600 K. Los resultados de esta dinámica se presentan a continuación.

## 4.4 Dinámica de la molécula completa con cinco átomos de magnesio

### 4.4.1 Posición I, II, III, IV y VI

En esta dinámica se introdujeron los 5 átomos de magnesio que aparecen en estructura cristalográfica [14]. La dinámica se realizó con el propósito de acelerar el proceso de catálisis. Se sabe de los experimentos que a mayor concentración, más rápido ocurre el rompimiento. Los resultados obtenidos nos servirán para observar qué tanto es capaz la Dinámica Molecular Clásica de simular este proceso biológico. Debido a que el número de iones metálicos es mayor que antes, existen más posibilidades de cambios conformacionales, lo cual podría darnos una orientación de cómo se está llevando a cabo este proceso y cuáles podrían ser las siguientes variables a añadir.

El esquema de los iones metálicos está dado por la figura 4.1, en la cual apare-

cen las posiciones de los cinco átomos de magnesio, estos corresponden a los números *I*, *II*, *III*, *IV*, *VI*. Como observamos de las dinámicas anteriores, el  $Mg_{VI}$  corresponde al ión que se une al  $O2PA_{1.1}$  y permanece firme en esta posición durante toda la dinámica. Sin embargo, el otro átomo de magnesio, el  $Mg_{III}$  que se había observado unido al  $O2'C_{17}$  ya no se encuentra tan afín a esta posición, la distancia de máximo acercamiento es de 5 Å. El  $Mg_{III}$  es el magnesio más cercano al oxígeno  $O2'C_{17}$ . Asimismo, la formación de un nuevo enlace o la de una conformación CCA, se encuentra mucho más distante bajo estas condiciones que en aquellas donde había uno o dos átomos magnesio, contrario a la evidencia experimental, donde este proceso debería estar acelerado por la mayor concentración de iones divalentes.

### Datos de la Dinámica

Los datos de esta dinámica aparecen en el Apéndice I. Se mantienen las mismas condiciones que en las dinámicas anteriores, presión y temperatura constante, la presión fué de 1 atm y la temperatura de 600K. Antes de iniciar la dinámica se realizó una minimización de energía, como la estructura presenta un número mayor de variables que podrían desestabilizarla, el número de pasos fue de 200 con el método ABNR descrito en el capítulo 2. El número de pasos de integración de la dinámica fue de 250000 que corresponde a 500 ps.

### Resultados

En la figura (E) del Apéndice II, tenemos la gráfica de la energía, en ella podemos observar que esta se mantiene constante después de un periodo de equilibración de aproximadamente 10 ps. No se observa ningún salto, por lo que no hubo ningún cambio conformacional importante.

### Distancias y ángulos

La primera gráfica que estudiaremos será la de la distancia entre los átomos que formarían el nuevo enlace:  $PA_{1.1} - O2'C_{17}$ , figura 4.21(A). En esta gráfica observamos que la distancia promedio se encuentra entre 4.2–4.6 Å y presenta un ligero brinco a los 360 ps. Este brinco tiende a acercar a los dos átomos a una distancia de 3.25 Å. Sin embargo, el acercamiento no es algo consistente, sino que se da durante un corto periodo de tiempo.

Las siguientes gráficas, muestran qué átomos de magnesio se acercaron al oxígeno  $O2'C_{17}$ . Vemos que el  $Mg_{III}$  es el que logra acercarse más a este oxígeno, pero no forma un enlace consistente como fue el caso de dos iones

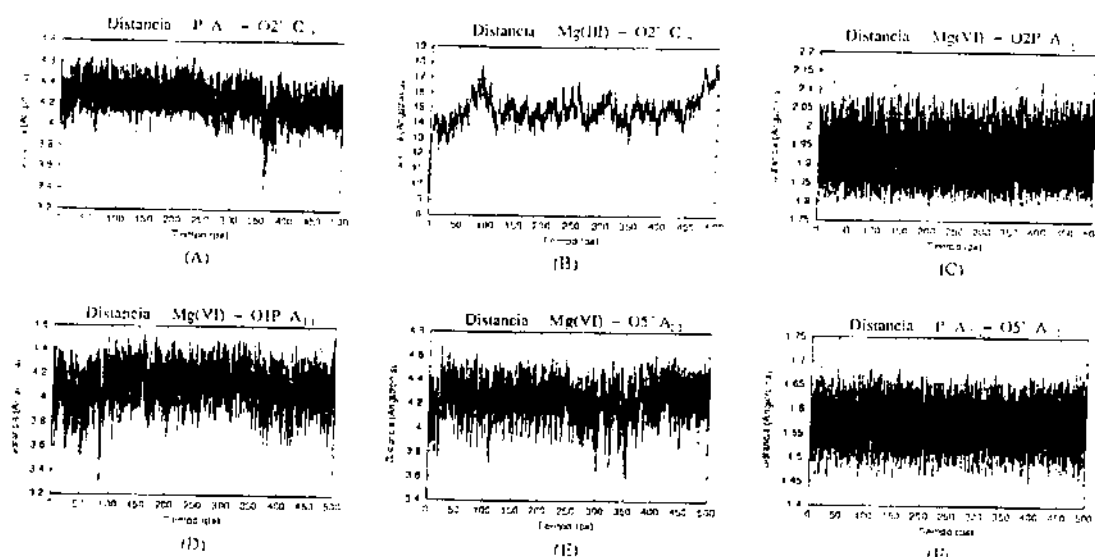


Figura 4.21: Dinámica con cinco magnesios: distancias relativas.

metálicos donde observábamos como este magnesio permanecía firme en esta posición, figura 4.19(C) y que ningún otro magnesio se logra acercar a esta zona. Observamos en la siguiente gráfica, figura 4.21(C), que el  $Mg_{VI}$  se mantiene unido al  $O1PA_{1.1}$  durante toda la dinámica, en la misma forma consistente que en las otras dinámicas.

Al mismo tiempo que el  $Mg_{VI}$  se acerca al  $O2'C_{17}$  vemos que se acerca también a los oxígenos  $O5'A_{1.1}$  y  $O2PA_{1.1}$ , figura 4.21(D) y (E). El enlace que se debe romper:  $PA_{1.1}$  y  $O5'A_{1.1}$ , figura 4.21(F), no muestra ningún cambio importante que indique una repulsión o catálisis.

Respecto a los ángulos, figura 4.22, vemos que el ángulo de formación previa a la catálisis es menor a los  $90^\circ$  durante toda la dinámica, lo cual nuevamente no indica ningún cambio conformacional hacia una catálisis.

## Conclusiones

A partir de los resultados anteriores vemos que en el estudio de la conformación previa a la catálisis de la ribozima "Hammerhead" no es necesario incluir varios iones de magnesio. En un principio se creía que era necesario un cambio conformacional en todo el "Hammerhead" y este cambio implicaría un cambio



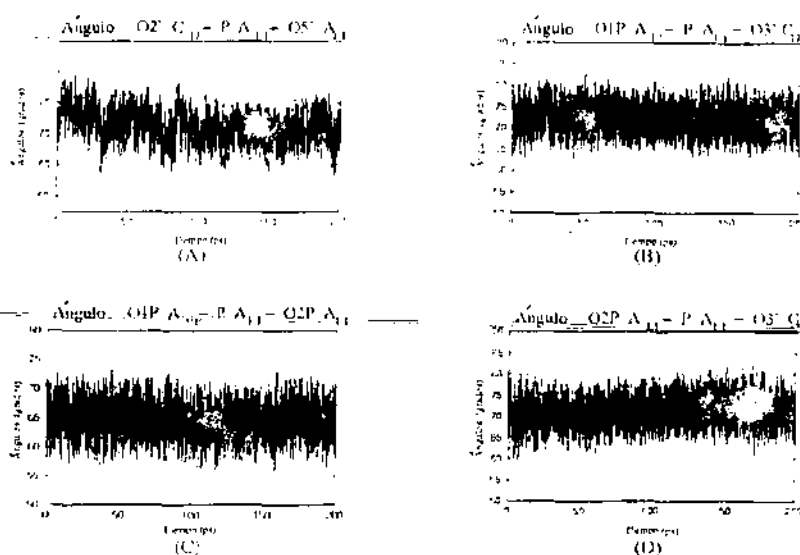


Figura 4.22: Dinámica con cinco magnesios: ángulos relativos.

en la zona catalítica pero los datos muestran que esto no es así. Aunque el papel de los iones de magnesio debe ser importante para la catálisis, es probable que su papel tenga que ver con una estabilidad estructural más que con su participación explícita para la catálisis. En esta dinámica se dió un tiempo mucho mayor que en las dinámicas anteriores con el fin de tener un intervalo mayor para dar lugar a cambios importantes. Vemos que no ocurre así, es probable que se requiera de un tiempo computacional aún mayor.

Antes de presentar los resultados obtenidos con los cálculos *ab initio* realizaremos un estudio sobre los cambios conformacionales presentados en la estructura a fin de que podamos confirmar algún otro dato experimental. Este estudio se basa en el análisis de los puentes de hidrógeno supuestos por los autores de la estructura cristalográfica [14] obtenida por rayos X, además se estudia si el rompimiento de alguno de estos enlaces explica los cambios conformacionales presentados en las diferentes dinámicas y también se esclarecerá si el cambio en la zona catalítica esta relacionado con cambios generales en la estructura del "Hammerhead".

## 4.5 Análisis de puentes de hidrógeno en las dinámicas

Se realizó un estudio sobre los cambios en la estructura debido a los puentes de Hidrógeno. De acuerdo a lo mencionado en el capítulo 1, tenemos básicamente 15 enlaces de este tipo para la estructura propuesta por Scott *et al* [14].

La tabla 4.1 muestra la distancia promedio de los puentes de hidrógeno. También se presenta la desviación estándar que indica el margen de desviación de la distancia promedio relativa. Se presentan únicamente las dinámicas realizadas con la molécula "Hammerhead" completa.

Puentes de hidrógeno	Sin Mg	Un Mg (VI)	Un Mg (III)	Dos Mg (III, VI)	Cinco Mg (Todos)
(N3 - H3)U <sub>7</sub> ... O6G <sub>8</sub>	3.40±0.54	3.21±0.65	4.44±1.26	3.01±0.29	3.16±0.44
(N2 - H21)G <sub>8</sub> ... O2PA <sub>1.3</sub>	3.47±0.53	2.92±0.56	2.98±0.50	3.17±0.73	3.05±0.67
(N2 - H22)G <sub>8</sub> ... N7A <sub>1.3</sub>	2.01±0.14	2.03±0.17	2.03±0.15	1.99±0.14	2.02±0.16
(N6 - H61)A <sub>1.3</sub> ... N3G <sub>8</sub>	2.19±0.21	2.27±0.26	2.30±0.24	2.26±0.26	2.27±0.25
(N6 - H62)A <sub>1.3</sub> ... O4'G <sub>8</sub>	2.56±0.33	2.92±0.43	2.89±0.40	2.84±0.39	2.71±0.36
(N1 - H1)G <sub>1.2</sub> ... O2'G <sub>8</sub>	4.09±0.48	3.03±0.56	2.80±0.52	3.67±0.79	3.76±0.53
(N2 - H21)G <sub>1.2</sub> ... O2'G <sub>8</sub>	3.03±0.55	2.15±0.33	2.09±0.26	2.76±0.60	2.61±0.49
(H2 - H22)G <sub>1.2</sub> ... N7A <sub>9</sub>	2.08±0.18	2.09±0.19	2.05±0.26	2.06±0.17	2.06±0.16
(N6 - H61)A <sub>9</sub> ... N3G <sub>1.2</sub>	2.19±0.19	2.23±0.23	2.25±0.22	2.24±0.26	2.22±0.20
(N6 - H62)A <sub>9</sub> ... O2'G <sub>1.2</sub>	2.16±0.26	2.21±0.31	2.16±0.24	2.24±0.34	2.15±0.26
(N6 - H62)A <sub>9</sub> ... O4U <sub>18.1</sub>	2.37±0.44	4.96±0.92	6.77±1.11	5.72±0.75	4.95±0.57
(N4 - H42)C <sub>17</sub> ... N3C <sub>3</sub>	3.07±0.30	3.13±0.26	3.29±0.30	2.56±0.48	4.89±2.28
(N3 - H3)U <sub>4</sub> ... O1PU <sub>7</sub>	2.73±0.16	2.76±0.13	2.75±0.19	2.79±0.13	2.73±0.12
(O2' - HO2')U <sub>4</sub> ... N7A <sub>8</sub>	2.89±0.13	2.86±0.11	2.86±0.12	2.85±0.13	2.88±0.13
(O2' - HO2')G <sub>5</sub> ... O2'C <sub>15.2</sub>	3.34±0.56	3.27±0.39	3.87±0.58	2.86±0.17	3.89±0.57

Tenemos en primer lugar los puentes de hidrógeno en los que consistentemente permaneció el enlace durante todas las dinámicas, estos fueron:

- i)  $N6 A_{1,3} \quad H61 A_{1,3} \cdots N3 G_8$
- ii)  $N2 G_{1,2} \quad H22 G_{1,2} \cdots N7 A_9$
- iii)  $N6 A_9 \quad H61 A_9 \cdots N3 G_{1,2}$
- iv)  $N6 A_9 \quad H62 A_9 \cdots O2' G_{1,2}$
- v)  $N3 U_4 \quad H3 U \cdots O1P U_7$
- vi)  $O2' U_4 \quad HO2' U_4 \cdots N7 A_6$

La mayoría de estos nucleótidos se encuentran en la "vuelta en U",  $N_3, U_4, N_7, G_8, A_9$ , por lo que se puede decir que estas bases son necesarias para la preservación de la forma del "Hammerhead". El enlace que definitivamente nunca se dió en todas las dinámicas realizadas es:

- i)  $N_6 A_{15,1} - H62 A_9 \cdots O4 U_{16,1}$

De los otros enlaces que permanecen más flexibles en las dinámicas con pocos metales, notamos que en la dinámica con cinco metales son mucho más fuertes, siendo las distancias promedio y la desviaciones estándar más pequeñas, lo que va de acuerdo con la teoría de que los iones divalentes estabilizan la estructura del ARN, con excepción de un enlace, el  $N3C_3O - H42C_{17}$ , que notoriamente aparece a distancias cercanas a los 3 Å en las dinámicas con baja densidad de iones metálicos, mientras que con cinco iones metálicos casi llega a los 5 Å. Tal vez el rompimiento de este enlace se encuentre relacionado con la aumento de la tasa de catálisis de la ribozima al aumentar la concentración de iones divalentes.

En vista de que ni con mayor temperatura ni con mayor concentración de magnesio se observa un cambio conformacional importante, es probable que este cambio no pueda ser observado utilizando dinámica molecular a partir de Campos de Fuerza. Para poder comprobar si efectivamente se puede presentar un rompimiento en el "Hammerhead" se recurrió a dinámicas *ab initio*. No se presentará la base teórica de estas dinámicas, sólo se presentarán los resultados obtenidos y en el siguiente capítulo se compararán los resultados obtenidos clásicamente con los resultados obtenidos utilizando dinámica molecular cuántica.

## 4.6 Dinámicas *ab initio* de la zona catalítica

### 4.6.1 Dinámica con un ión de magnesio

La primera dinámica realizada involucra solo un átomo de magnesio. Debido a que los cálculos *ab initio* son muy costosos computacionalmente, solamente se tomaron en cuenta los átomos cercanos a la catálisis de los nucleótidos  $C_{17}$  y  $A_{1.1}$ .

#### Datos de la Dinámica

Los átomos que intervinieron en estos cálculos consistieron en las dos ribosas de los nucleótidos  $C_{17}$  y  $A_{1.1}$ . Los enlaces correspondientes a sus respectivas bases fueron saturados con hidrógenos. Además se incluyó el átomo de magnesio en la posición VI, que es la posición más consistente en las dinámicas desarrolladas con un metal, dos metales y 5 metales. También se incluyeron las moléculas de agua más cercanas a la región catalítica.

La dinámica se realizó con un total de 78 átomos entre los cuales teníamos hidrógeno, oxígeno, carbono, fósforo y magnesio. Los átomos correspondientes al nucleótido tuvieron bases 6 – 31G\* y los átomos que forman las moléculas de agua tuvieron bases *mínimas*. El programa utilizado fue NWChem<sup>5</sup>. La dinámica se desarrolló durante 160 *fs*.

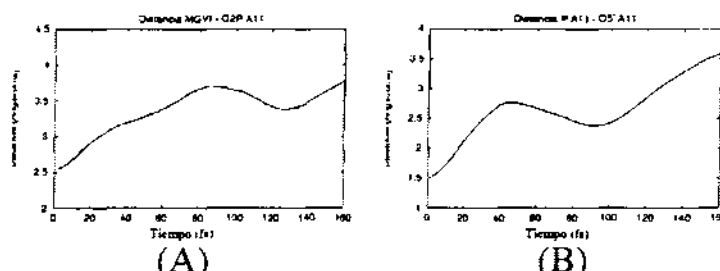


Figura 4.23: Gráficas de distancias interatómicas en la dinámica *ab initio* con un átomo de magnesio

### Resultados

En la figura 4.23 tenemos las gráficas de distancias, en la figura 4.24 se muestran los cuadros de algunos eventos importantes.

<sup>5</sup>High-performance Computational Chemistry group of the Environmental Molecular Sciences Laboratory (EMSL) at the Pacific Northwest National Laboratory (PNNL). (Point of Contact: Rubén Santamaría)

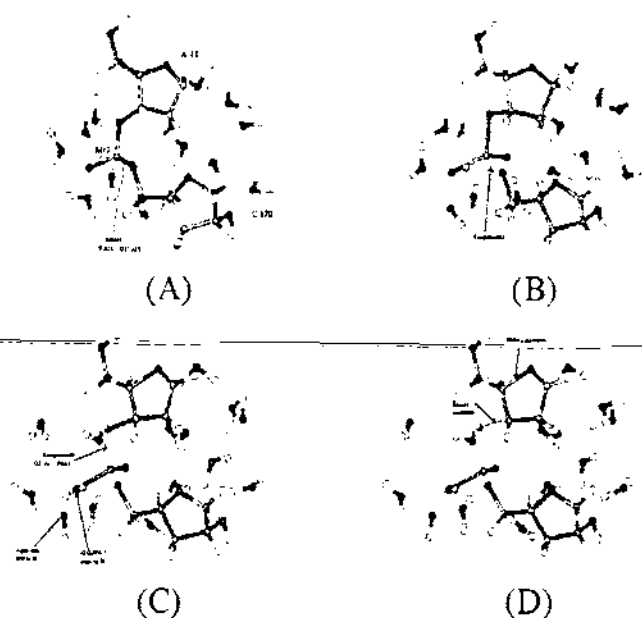


Figura 4.24: Dinámica cuántica con un ion de magnesio

La primera figura 4.23(A) muestra la conformación al inicio de la dinámica, la distancia entre el átomo de magnesio y el  $O2PA_{1.1}$  es de 2.56 Å. Lo primero que observamos en la figura 4.23(B), después de 20 fs de dinámica es el rompimiento del enlace  $PA_{1.1} - O5'A_{1.1}$ , este rompimiento ocurre cuando el ángulo CCA tiene un valor de  $89^\circ$ . A continuación, en la figura 4.24(C), el  $O1PA_{1.1}$  se une a un átomo de hidrógeno a una molécula de agua que se encontraba en su vecindad. El enlace es inestable al inicio, pero después permanece durante toda la dinámica. Por otro lado, vemos que el átomo de magnesio en esta posición tiende a alejarse de la región catalítica al mismo tiempo que los oxígenos del fosfato como se ve en la gráfica de distancias, figura 4.23. En la misma figura 4.24(C) se rompe el enlace  $O3'A_{1.1} - PA_{1.1}$ . Otros enlaces, que se suponen más fuertes, son inestables, como por ejemplo los enlaces que se rompen y vuelven a formar, el  $C3'C_{17} - O3'C_{17}$  y del azúcar  $C4'C_{17} - O4'C_{17}$  en la figura 4.24(D). Esta última figura también muestra el final de la dinámica. Se decidió terminarla aquí debido a que tanto el átomo de magnesio como el fósforo se alejan sin formar nuevos enlaces con los átomos de las pentosas.

### 4.6.2 Dinámica con dos iones de magnesio

La segunda dinámica realizada involucra dos átomos de magnesio y consiste igualmente en las dos ribosas de los nucleótidos  $A_{1.1}$  y  $C_{17}$ , además de las moléculas de agua que se encuentran unidas tanto a los magnesios como las posibles moléculas de agua que podrían ser donadores de hidrógenos. Las distancias iniciales fueron:

- i)  $PA_{1.1} - O5'A_{1.1}$  1.53 Å
- ii)  $O2'C_{17} - PA_{1.1}$  3.19 Å
- iii)  $Mg_{III} - O2'C_{17}$  2.18 Å
- iv)  $Mg_{VI} - O2PA_{1.1}$  2.56 Å

#### Resultados

La figura 4.25 muestra la distancia relativa de los átomos participantes.

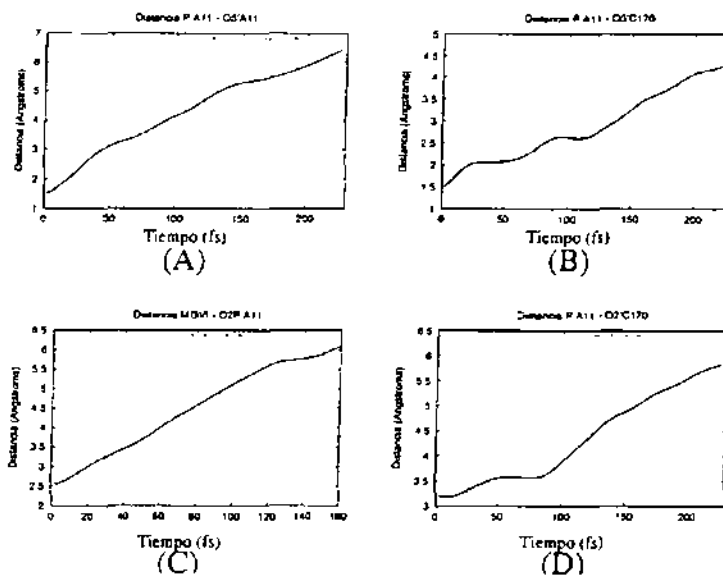


Figura 4.25: Gráficas de las distancias interatómicas en la Dinámica *ab initio* con dos metales

En la primera figura, fig 4.26(A), tenemos la molécula en el estado inicial de la dinámica cuántica. Aquí se observan los dos magnesios correspondientes a las posiciones III y VI y los dos anillos de la pentosa del  $A_{1.1}$  y  $C_{17}$ . Los enlaces correspondientes a las bases de cada pentosa se saturaron con hidrógenos, estos hidrógenos son inestables y rápidamente entran en resonancia separándose de

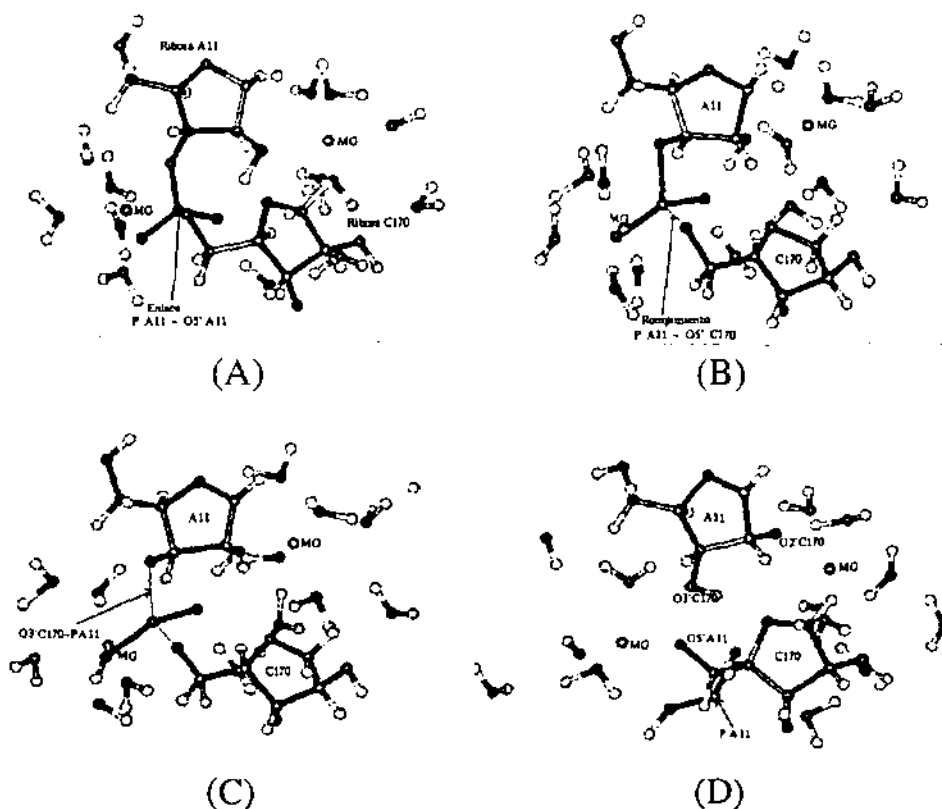


Figura 4.26: Dinámica cuántica con dos iones de magnesio

la molécula principal.

En la figura 4.26(B) vemos el primer rompimiento importante de la molécula que corresponde al enlace  $PA_{1.1} - O5'A_{1.1}$ , de acuerdo a los mecanismos de reacción que involucran a dos metales, figura 4.26(A). Sin embargo, siguiendo la evolución de la dinámica vemos en la figura 4.26(C) que a continuación se rompe el enlace entre el  $PA_{1.1} - O3'C_{17}$ , figura 4.26(B), dejando libre al átomo de fósforo con sus dos átomos de oxígeno  $O1PA_{1.1}$  y  $O2PA_{1.1}$ . En la figura 4.26(C) vemos que existe un rompimiento momentáneo de la molécula de azúcar de la  $C_{17}$ , después de unos pasos la molécula recupera su estructura estable. El átomo de fósforo se encuentra libre de la molécula de azúcar y se aleja del átomo con el cual debería formar un nuevo enlace el  $O2'C_{17}$ , gráfica 4.25(D). Se supone que el  $O2'C_{17}$  se desprende de su hidrógeno para reaccionar con el fósforo, pero esto no ocurre y el oxígeno no se encuentra en libertad de

interaccionar con otros átomos.

El átomo de fósforo suelto, empieza a reaccionar con los átomos de alrededor y el oxígeno  $O2PA_{1,1}$  se satura con una molécula de hidrógeno formando un grupo hidroxilo, figura 4.26(C). Este átomo al igual que el  $MgVI$  continúa alejándose, como se ve en la figura 4.25(D) y en la gráfica 4.25(C). En este momento se decide terminar la dinámica.

A lo largo del capítulo se han presentado los resultados obtenidos hasta ahora. En cada dinámica hemos presentado las conclusiones particulares. Sin embargo, el conjunto completo de dinámicas clásicas y cuánticas nos lleva a conclusiones generales que se presentan en el siguiente capítulo, así como cuáles serían los pasos siguientes para continuar con el estudio de la catálisis del "Hammerhead".



## Capítulo 5

### Conclusiones

En este capítulo se exponen las conclusiones generales del estudio realizado sobre el mecanismo de reacción de la ribozima “Hammerhead” y el papel de los iones metálicos divalentes en la catálisis. Los estudios comprenden simulaciones tanto de Dinámica Molecular clásica de la molécula completa como cálculos *ab initio* del sitio de rompimiento.

Los cambios conformacionales de la ribozima “Hammerhead” se efectuaron con base en la estructura resuelta por W. Scott [14]. Aunque la estructura es la que mejor describe a la ribozima “Hammerhead”, presenta las limitantes inherentes a cualquier macromolécula resuelta por cristalografía de rayos X. Entre estas tenemos que la estructura presentada no se obtuvo en un ambiente biológico, además de las técnicas utilizadas con el fin de obtener una estructura cristalina. Finalmente, el archivo PDB utilizado contiene posiciones de 5 átomos de magnesio que fueron determinadas por métodos indirectos.

Al investigar la catálisis del “Hammerhead” nos hemos basado en las dos hipótesis que existen hasta el momento sobre el mecanismo de rompimiento de la molécula [17]. Mientras un mecanismo se inclina hacia la intervención de un sólo metal divalente [6], el segundo mecanismo favorece la participación de dos iones metálicos divalentes. Sin embargo, ambos mecanismos presuponen que antes de la catálisis el “Hammerhead” debió haber pasado por un cambio conformacional que colocara a los átomos involucrados en el rompimiento en una posición de Línea de Ataque. Al respecto también se han formulado varias teorías sobre la dirección en la cual se debe realizar el cambio. Una de estas establece que el nucleótido C170 debe girar sobre la cadena de fósforos

hacia fuera exponiendo al  $O2'C170$  a ser atacado por un  $Mg^{+2}$  y perder su hidrógeno y así ser atacado por el  $PA11$  para formar un enlace fosfodiéster. La otra teoría propone que es el fósforo  $PA11$  con sus dos oxígenos *pro - R* y *pro - S* el que debe girar hacia el interior de la cadena, favoreciendo el rompimiento del enlace  $PA11 - O5'A11$  y acercándose para atacar al oxígeno  $O2'C170$ . Cualquiera de estos dos mecanismos supone un cambio de energía muy grande.

Los cálculos realizados con DM clásica pretendían observar un cambio conformacional del "Hammerhead" que dejará a los átomos que intervienen en la catálisis en posiciones previas a la formación y rompimiento de enlaces en los átomos respectivos. Para tal efecto adoptamos el criterio establecido por Torres *et al* [6] para el desarrollo de una conformación de Línea de Ataque previa a la catálisis (CCA). Este criterio establece que la distancia promedio entre el  $PA11$  y el  $O2'C170$  debe ser menor a 3.5 Å y el ángulo entre  $O2'C170 - PA11 - O5'A11$  debe ser mayor a  $150^\circ$ . Conociendo las limitaciones de utilizar un método que otorga flexibilidad al sistema de átomos pero que no deja que éstos presenten rompimientos, creemos que el conjunto de simulaciones brindan al sistema las variables suficientes para observar cualquier cambio conformacional. Sin embargo, como se ha expuesto en el capítulo anterior no fue posible observar un cambio que favoreciera a alguna de las tendencias expuestas con anterioridad.

Con el fin de observar cualquier cambio conformacional en la ribozima "Hammerhead", las dinámicas fueron realizadas bajo diferentes condiciones. Las variables consideradas fueron el número y las posiciones de los átomos de magnesio, formación del cristal, temperatura, energía, flexibilidad y número de nucleótidos. No se observó este cambio en las posiciones de los átomos, y esto puede tener diversas causas, entre las cuales se pueden sugerir las siguientes:

- i) El programa computacional de DM fue CHARMM, pero el conjunto de parámetros para el Campo de Fuerza no es totalmente adecuado. Sabemos que existen nuevas versiones de este programa y tal vez en un futuro sea posible realizar nuevas simulaciones con parámetros diferentes.
- ii) Es probable que el cambio conformacional no pueda modelarse utilizando potenciales semiempíricos. Los mecanismos propuestos establecen un cambio conformacional previo a la catálisis. En las dinámicas

*ab initio* observamos rompimiento en la ribozima, lo cual podría sugerir que el mecanismo sea diferente; primero se efectúa el rompimiento y a continuación se da el cambio conformacional.

- iii) Por otra parte, es también posible que no se le hayan dado los periodos de relajación y equilibración adecuados, pero existen varias limitantes en el tiempo y espacio que no nos permitieron realizar las simulaciones por largos periodos de tiempo. Esto, sin embargo, nos llevó a buscar alternativas que nos permitieran acelerar el proceso de cambio conformacional, tales como el incremento de la temperatura o el número de magnesios. Sin embargo, no se observaron cambios evidentes.

Entre los resultados obtenidos y que nos permiten creer que las simulaciones arrojarían información correcta es la concordancia entre los resultados teóricos y algunos resultados experimentales. Uno de esos es el hecho de que el magnesio en la posición VI, se coordina de forma directa y consistente en todas las dinámicas al oxígeno *pro-R* del fosfato **A11**. La evidencia experimental indica que la catálisis no es posible si este oxígeno es sustituido por algún otro elemento, lo cual reafirma su participación. En referencia al magnesio en la posición III y su adherencia al oxígeno **O2'C170**, esta no fue tan marcada como el caso anterior. Mientras este proceso se observó fuertemente cuando se modeló un sólo metal, este efecto desapareció cuando se incluyeron más iones metálicos. Si consideramos que el sistema con un número mayor de magnesios favorece la catálisis, esto podría indicar que para observar el rompimiento del oxígeno **O2'C170**, este debe encontrarse en libertad de movimiento. Otra concordancia entre los cálculos computacionales y teóricos se refiere a los puentes de hidrógeno, los cuales se observaron en mayor o menor medida dependiendo de las condiciones, indicando que la conformación del "Hammerhead" con la cual estamos trabajando manifestaba una estructura estable que mantuvo su forma y tamaño de forma consistente en todas las dinámicas. Un caso en el cual se podría seguir trabajando sería el único puente de hidrógeno cercano a la región catalítica, el **N3C3 - H42C170**. En la hipótesis sobre la dirección del cambio conformacional se indica que la citosina **C170** debe girar. Sin embargo, este puente de hidrógeno permaneció inalterado en las dinámicas con pocos iones metálicos. Aunque en algunos casos este se rompió, no está directamente relacionado con el acercamiento entre los átomos ligantes **PA11** y **O2'C170**.

Los resultados *ab initio* mostraron que efectivamente se da la catálisis entre los nucleótidos A11 y C170 en los átomos PA11 – O5'A11 y sugiere que sólo es necesaria la intervención de un ion metálico, lo que conduce a pensar que el rompimiento surge de forma natural. Sin embargo, como los átomos no estaban colocados en una conformación adecuada, entonces no se formaron los enlaces correspondientes a un 5 – fosfodiester. Es posible que considerando un número mayor de átomos de agua y de las bases nucleicas, se obtengan mejores resultados, no obstante, estos cálculos requerirán un mayor número de recursos y por el momento no podrían efectuarse. Que sólo sea necesario un ión metálico puede ser debido a la gran repulsión electrostática que se presenta entre los dos iones  $Mg^{+2}$ . Esta repulsión afecta a la molécula y provoca separación entre los átomos a los cuales se encuentra adherida, por lo que su presencia no deja claro un incremento en la tasa de reacción.

Finalmente, aunque los resultados obtenidos son parciales, es posible decir que los métodos utilizados permitirán efectuar en un futuro un estudio más consistente sobre como es en realidad el mecanismo de rompimiento del "Hammerhead" y cuantificar el papel de los iones metálicos divalentes.

# Apéndice I. Archivo del programa de entrada de CHARMM

Los siguientes párrafos muestran el archivo de entrada para realizar una simulación de un sistema biológico utilizando CHARMM.

La primera parte del programa consiste en leer los parámetros adecuados al tipo de molécula que se quiere simular, en este caso un ácido nucleico (10).

A continuación viene la lectura de los archivos (PDB) que contienen las posiciones de los átomos dados por la cristalografía (20) y la adición de los hidrógenos faltantes (25).

Después se sumerge la molécula en un cubo de agua previamente fabricado. Las dimensiones del cubo son el tamaño original de la molécula más 7 Å de cada lado (30). Las moléculas de agua que se superponen con los átomos de la molécula son quitadas (35).

La molécula entra entonces en un periodo de relajación indicado por una minimización (40) de 30 pasos y se escriben las nuevas coordenadas atómicas (45).

Las condiciones CPT (presión y temperatura constantes) requieren la formación de un cristal con cargas imágenes que permita conservar el mismo número de partículas en la misma unidad de volumen. El tamaño del cristal se indica en (55) y se pide de nuevo que escriba las coordenadas atómicas (58).

La última parte comprende la dinámica (60) y los datos de esta como el tipo de algoritmo de integración, temperatura, presión, tamaño del paso, etc. (65).

La dinámica termina después de 250000 pasos de integración (70).





```
! for memory managment (see file nbonds.doc)
! allocate 60% more memory than CHARMM would usually guess
```

```
update nbscal 1.6
```

```
! the following a,b,c parameters are the |Xmas - Xmenos|,
! |Ymas - Ymenos|, etc of the solvated protein
```

```
55 crystal define orth 70.37 56.32 38.12 90.0 90.0 90.0
```

```
! inbfrq= frequency to update the non-bond terms (van der
!           Waals and electrostatic terms)
! ihbfrq= frequency to update the H-bond list
! shift = determines the long-range truncation option
!           in the electrostatics (see file nbonds.doc), shifted
!           potential acting to CTOFNB and zero beyond
! atom = electrostatics computed on an atom-atom pair basis,
!           also related with the shift option (shifted potential
!           acting to CTOFNB and zero beyond)
! vswitch=similar to the shift option
! vatom = Van Der Waal atom, related with the vswitch option
! cutnb = cutoff distance in the electrostatics
! ctonnb= initial value for switching function
! ctofnb= final value for switching function
! wmin = warning cutoff for minimum atom-atom distance
!           pairs are checked for close contact
```

```
update inbfrq 25 ihbfrq 0 shift atom vswitch vatom -
cutnb @nbcut ctonnb @nbon ctofnb @nboff wmin 1.5
```

```
! the cutoff distance "@imgcut" determines images to be
! included in the transformation list
```

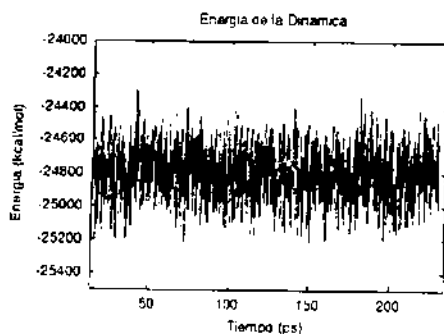
```
crystal build cutoff @imgcut
```

```
! imgfrq= frequency to update the image list
```

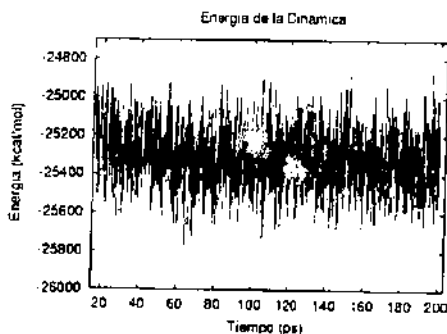




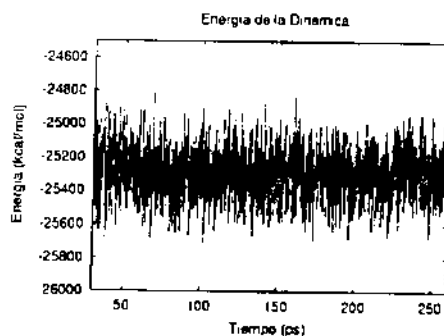
## Apéndice II. Energía de las dinámicas



(A)



(B)



(C)

Figura 1: (A) Energía de la dinámica sin  $Mg$ , (B) energía de la dinámica con un metal  $Mg_{VI}$ ; (C) energía de la dinámica con un metal  $Mg_{III}$ .

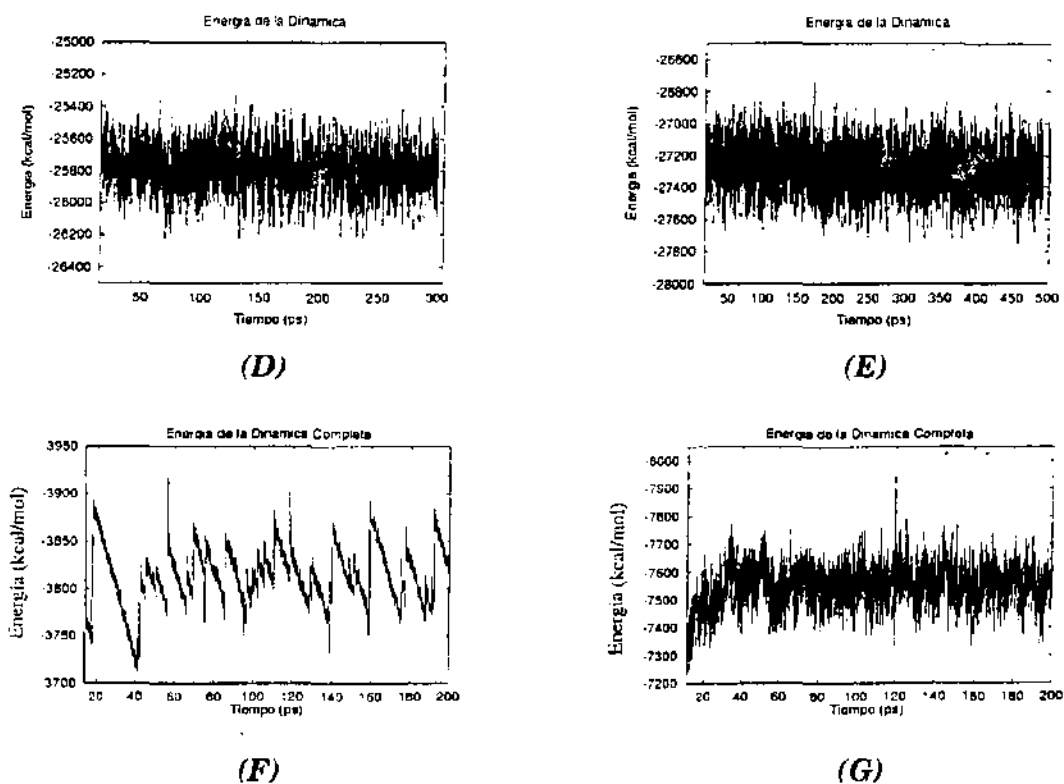


Figura 2: (D) Energía de la dinámica con dos metales  $Mg_{III}$  y  $Mg_{VI}$ , (E) energía de la dinámica con cinco metales  $Mg_{(I,II,III,IV,VI)}$ , (F) energía de la dinámica del modelo corto sin cristal, (G) energía de la dinámica del modelo corto con cristal.

## Referencias

- [1] Elliot, W. H. y Elliot, D. C.: "Biochemistry and molecular biology", Oxford University Press, EUA, (1997).
- [2] Saenger, W.: "Principles of nucleic acid structure", Springer Verlag, New York, (1984).
- [3] Garret, R., y Grisham, C.: "Biochemistry", Saunders College Publishing, EUA, (1995).
- [4] Mathews, C. K., van Holde, K. E.: "Biochemistry", The Benjamin/Cummings Publishing Company, EUA, (1990).
- [5] Naray-Szabo, G.: "Theoretical chemistry of biological systems", *Studies in physical and theoretical chemistry*, EUA, (1986).
- [6] Torres, R. A. y Bruice, T. C.: "Molecular dynamics study displays near in line attack conformations in the hammerhead ribozyme", *Proceedings of the National Academy of Science of the USA*, vol. **95**, 11077-11082, (1998).
- [7] Eckstein, F. y Lilley D.: "Catalytic RNA", *Nucleic acids and molecular biology*, vol. **10**, Editorial Springer, Alemania, (1997).
- [8] Landweber, L. F., Simon, P. J., Wagner, T. A.: "Ribozyme engineering and early evolution", American Institute of Biological Sciences, *Bioscience*, vol. **48**, (1998). [www.aibs.org/BioScienceLibrary](http://www.aibs.org/BioScienceLibrary)
- [9] Gesteland, R. R. y Atkins, J. F.: "The RNA World", Cold Spring Harbor Laboratory Press, EUA, (1993).
- [10] Pley, H. W., Flaherty, K. M. y McKay, D. B.: "Three-dimensional structure of a hammerhead ribozyme", *Nature*, vol. **372**, 68-74, (1994).

- [11] Zaug, A. J. y Cech, T. R.: "The intervening sequence RNA of *Tetrahymena* is an enzyme", *Science*, vol. **23**, 470-475, (1986).
- [12] Kruger, K., Grabowski, P., Zaug, A. J., Sands, J., Gottschling, D. E., Cech, T. R., "Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*", *Cell*, vol. **31**, 147-157, (1982).
- [13] Guerrier-Takada, C., Gardiner K., Marsh, T., Pace, N., Altman, S. "The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme", *Cell*, vol. **35**, 849-857, (1983).
- [14] Scott, W. G., Murray, J. B., et al: "Capturing the structure of a catalytic RNA intermediate: the Hammerhead ribozyme", *Science*, vol. **274**, 2065-2068, (1996).
- [15] Murray, J. B. y Scott, W. G.: Does a single metal ion bridge the A - 9 and scissile phosphate groups in the catalytically active Hammerhead ribozyme structure, *Journal of Molecular Biology*, vol. **296**, 33-41, (2000).
- [16] Barinaga, M.: "Ribozymes: killing the messenger", *Science*, vol. **262**, 1512-1514, (1993).
- [17] Wedekind, J. E. y McKay, D. B.: "Crystallographic structures of the Hammerhead ribozyme: relationship to ribozyme folding and catalysis", *Annu. Rev. Biophys. Biomol. Struct.*, vol. **27** 475-502, (1998).
- [18] Lott, W. B., et al: "A two-metal ion mechanism operates in the hammerhead ribozyme-mediated cleavage of an RNA substrate", *Biochemistry*, vol. **95**, 542-547, (1998).
- [19] De-Min Zhou, et al: "Explanation by the double-metal-ion mechanism of catalysis for the differential metal ion effects on the cleavage rates of 5'-oxy and 5'thio substrates by a hammerhead ribozyme", *Proceedings of the National Academy of Science of the USA*, vol. **94**, 14343-14348, (1997).
- [20] Peracchi, A.: "Involvement of a specific metal ion in the transition of the Hammerhead ribozyme to its catalytic conformation", *Journal of Biological Chemistry*, vol. **272**, 26822-26826, (1997).
- [21] Quan Du: "Ribozyme Enzimology", WWW: <http://academic.brooklyn.cuny.edu/chem/zhuang/QD/>

- [22] Pontius, B. W., Lott, W. B. y von Hippel, P. H.: "Observations on catalysis by hammerhead ribozymes are consistent with a two-divalent-metal-ion mechanism", *Proceedings of the National Academy of Science of the USA*, vol. **94**, 2290-2294, (1997).
- [23] Simorre, J., Legault, P., et al: "A conformational change in the catalytic core of the Hammerhead ribozyme upon cleavage of an RNA substrate", *Biochemistry*, vol. **36**, 518-525, (1997).
- [24] Bruice, T. C., Tsubouchi, A., et al: "One- and two-metal ion catalysis of the hydrolysis of adenosine 3'-Alkyl phosphate esters. Models for one- and two-metal ion catalysis of RNA hydrolysis", *Journal of the American Chemical Society*, vol. **118**, 9867-9875, (1996).
- [25] Mei, H., Kaaret, T. W., y Bruice, T. C.: "A computational approach to the mechanism of self-cleavage of hammerhead RNA", *Proceedings of the National Academy of Science of the USA*, vol. **86**, 9727-9731, (1989).
- [26] Pastor, N. C.: "Formación de complejos macromoleculares: un análisis microscópico", VII Escuela de Verano en Física, (2000).
- [27] Bernstein, F. C., et al: "The Protein Data Bank: a computer-based archival file for macromolecular structures", *Journal of Molecular Biology*, vol. **112**, 535-42 (1977).
- [28] Berman, H. M., et al.: "The nucleic acid database: A comprehensive relational database of three dimensional structures of nucleic acids", *Biophysics Journal*, vol. **63** 751-9 (1992).
- [29] Allen, M. P. y Tildesley, D. J.: "Computer simulation of liquids", Oxford University Press, New York, (1987).
- [30] Rapaport, D. C.: "The Art of molecular dynamics simulation", Cambridge University Press, (1995).
- [31] Karplus, M. y Petsko, A.: "Molecular Dynamics Simulation in Biology", *Nature*, vol. **347**, pag 631-639, (1990).
- [32] Goldstein, H.: "Classical mechanics", Addison Wesley Publishing Company, EUA, (1950).
- [33] Haile, J. M.: "Molecular dynamics simulation, elementary methods", John Wiley and Sons, (1992).

- [34] Gelin, B. R.: "Molecular modeling of polymer structures and properties". Hanser Publishers, (1994).
- [35] Colbourn, E. A.: "Computer simulation of polymers", (1994).
- [36] Marion, J. B. y Thornton, S. T.: "Classical dynamics of particles and systems", Harcourt Brace Jovanovich Publishers, (1988).
- [37] McCammon, J. A. y Harvey, S. C.: "Dynamics of proteins and nucleic acids", Cambridge University Press, Cambridge, (1987).
- [38] Pitt, W. G. y Weaver, D. R.: "Calculation of protein-polymer force fields using molecular dynamics", *Journal of Colloid and Interface Science*, vol. **185**, 258-264 (1997).
- [39] Ercolesi, F.: "A molecular dynamics primer",  
**WWW:** <http://www.sissa.it/furio/md/md/node1.html>,  
International School for Advances Studies (SISSA-ISSAS), Trieste, Italy.
- [40] Oren, B.: "Chapter 4, Molecular dynamics",  
**WWW:** <http://www.tau.ac.il/ becker/course.html>
- [41] Andrew, L. R.: "Molecular dynamics: principles and applications" Addison Wesley Longman, Inglaterra, (1996).
- [42] Kestin, J. y Dorfman, J. R.: "A course in statistical Thermodynamics", Academic Press, EUA, (1971).
- [43] Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. y Karplus, M. "CHARMM: A program for macromolecular energy, minimization and dynamics calculations", *Journal of Computational Chemistry*, vol. **4** 187-217 (1983).
- [44] Allinger, N., Burkert, U.: "Molecular mechanics", American Chemical Society, Washington, (1982).
- [45] Stillinger, F. H. y Rahman, A.: *Improved simulation of liquid water by molecular dynamics* *Journal of Chemical Physics*, **60** 1545-1557, (1974).
- [46] Weiner, S. J., et al.: "An all atom force field for simulations of protein and nucleic acids", *Journal of Comp Chemistry*, vol. **7** 230-252, (1986).
- [47] Mackerell Jr., A. D. et al.: "An all-atom empirical energy function for the simulation of nucleic acids", *Journal of the American Chemical Society*, vol. **117** 11946-11075, (1995).

- [48] Brunner, R. et al.: "NAMD programming guide, version 1.5"  
WWW: <http://sake.uchsc.edu/library/namd/ug/ug.html>  
Theoretical Biophysics Group, University of Illinois and Beckman Institute,  
1998.
- [49] *The NIH Guide to molecular Modeling*  
WWW: <http://www.biochem.vt.edu/courses/modeling/molecular-dynamics.html>
- [50] Arfken, G. B. y Weber, H. J.: "Mathematical Methods for Physicist", Fourth Edition, Academic Press, 1995.
- [51] Atkins, P. W. y Friedman, R. S.: "Molecular Quantum Mechanics", Third Edition, Oxford University Press, E. U. A., (1997).
- [52] Szabo, A. y Ostlund, N. S.: "Modern Quantum Chemistry", Maryland, (1989).
- [53] Kohn, W. "Nobel Lecture: Electronic structure of matter-wave functions and density functionals", *Reviews of modern physics*, vol. 71, (1999).
- [54] Labanowski, J.K., Andzelm, M. W.: "Density Functional methods in chemistry", Springer Verlag, E. U. A., (1991).