



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS
QUÍMICAS**

**SIMULACIÓN MOLECULAR DE LA HIPOXANTINA NEUTRA EN SOLUCIÓN ACUOSA CON
UN MODELO POLARIZABLE.**

TESIS
PARA OPTAR POR EL GRADO DE
MAESTRO EN CIENCIAS

PRESENTA

Q. FRANZ MICHEL MARTINEZ RIOS



TUTOR: DR. MILTON THADEU GARCIA MEDEIROS DE
OLIVEIRA

AÑO: 2011



Universidad Nacional
Autónoma de México



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.

Agradecimientos:

Mi más sincero agradecimiento al Dr. Milton Medeiros por todo el apoyo y confianza que me ha dado. A la Dra. Maria Eugenia Costas y al Dr. Rodolfo Acevedo por su amistad y consejos, también les agradezco profundamente.

A los miembros del jurado, por sus revisiones al manuscrito, comentarios y observaciones.

A la UNAM y a la DGSCA por el apoyo recibido mediante los recursos y servicios de supercómputo.

Al CONACYT, por el apoyo económico otorgado.

A mis compañeros de cubículo, Rodolfo Reynoso y Pablo Tellez. A mi compañero de la licenciatura Ángel Sánchez.

Y sobre todo a mi familia, muchas gracias por todo.

Tesis realizada en el cubiculario 106, departamento de Físicoquímica de la Facultad de Química de la Universidad Nacional Autónoma de México.

Contenido.

Introducción.....	4
1. Tautomerismo.....	4
2. La Hipoxantina	5
3. Tautomerismo de la Hipoxantina.	7
4. Justificación: tautomerismo de la hipoxantina y simulaciones moleculares.	10
5. Objetivos y metas.	12
6. Hipótesis.....	12
Marco Teórico.	13
7. Simulaciones moleculares.	13
8. Métodos de Monte Carlo.....	13
9. Modelado: cálculo de la energía potencial.	18
i. El caso de los sistemas acuosos.	20
10. Polarizabilidad y su modelado.....	20
i. Modelos no polarizables	21
ii. Modelos de dipolos puntuales polarizables.	22
iii. Modelos polarizables de capa o Shell.....	24
iv. Modelos de ecualización de electronegatividad.	26
v. Modelos polarizables semiempíricos.....	28
11. El Modelo MCHO.....	29
12. Modelado de los sitios electrónicos MCHO: ANES y Single-Update.....	30
13. Cálculo de la energía libre: la integración termodinámica.....	34
Metodología.....	39
14. Organización.....	39
15. Labores realizadas.	39
16. Detalles computacionales.	41
i. Ciclos y configuraciones.....	41
ii. Movimientos nucleares: traslaciones.....	42
iii. Movimientos nucleares: rotaciones.	44
iv. Movimientos electrónicos: el ciclo ANES.....	45
v. Listas de Verlet.....	45
vi. Sumas de Ewald.....	48
vii. Análisis por bloques.....	49

17. Tratamiento de los resultados obtenidos.	50
Resultados y discusión.	53
18. Validación del método Monte Carlo ANES y su diferenciación del método Single - Update: Agua Pura.	53
i. Energía: entalpia de vaporización.....	53
ii. Estructura.	54
iii. Validación del método de Monte Carlo ANES.	57
19. Tautómeros en fase gas.....	57
20. Tautómeros en solución.....	58
i. Energía: tautómeros en solución.....	59
ii. Estructura agua – hipoxantina.....	60
21. Energía libre y constante de equilibrio de la reacción de tautomerización.	62
Conclusiones.....	65
22. Tautomerismo de la hipoxantina MCHO.	65
Apéndices.	67
1. Parámetros MCHO para las interacciones agua-agua.	67
2. Parámetros MCHO para las interacciones hipoxantina-agua.	68
3. Análisis por bloques.	70
Bibliografía.	71

Índice de figuras.

Figura 1. Tautomerismos ceto-enol (arriba izq.), lactima-lactama (arriba der.) e imina-enamina (abajo).	4
Figura 2. Tautomerismo anular en el isoindol.	4
Figura 3. La glucosa se interconvierte de una cadena cerrada a otra abierta mediante una tautomerización.	5
Figura 4. Tautomerismo de valencia en el bulvaleno.	5
Figura 5. Algunos tautómeros de la guanina.	5
Figura 6. Hipoxantina en su forma oxo.	6
Figura 7. Inosina, precursora del AMP y GMP.	6
Figura 8. Numeración de los átomos de C, N y O de la hipoxantina.	6
Figura 9. Una de las reacciones de tautomerización de la hipoxantina.	7
Figura 10. Tautómero enol observado en espectroscopia IR en matriz de argón.	8
Figura 11. Tautomerismo de la hipoxantina N1-N9 a hipoxantina N1-N7.	9
Figura 12. La conexión entre experimento, teoría y las simulaciones moleculares.	11
Figura 13. Representación esquemática de los parámetros en el modelo MCHO.	29
Figura 14. Comparación método Single-Update vs ANES con $R_{el} = 2$, en un sistema de partículas (esferas) con un sitio electrónico (puntos).	34
Figura 15. Diagrama de flujo de un ciclo de Monte Carlo ANES.	42
Figura 16. Las relaciones de aceptación en el ciclo de Monte Carlo.	43
Figura 17. Ciclo ANES.	46
Figura 18. Definición de los radios de las esferas para generar las listas de Verlet.	47
Figura 19. El comportamiento de la energía potencial del agua MCHO ANES en función del número de ciclos Monte Carlo.	53
Figura 20. Funciones de distribución radial experimental [114] (negro) y la de la simulación MCHO ANES (azul).	55
Figura 21. Función de distribución radial O-O. SU (Single-Update) y ANES.	55
Figura 22. Función de distribución radial O-H. SU (Single-Update) y ANES.	56
Figura 23. Función de distribución radial H-H. SU (Single-Update) y ANES.	56
Figura 24. Definición de las zonas para la generación de funciones de distribución radial zona hipoxantina – Oagua.	60
Figura 25. Funciones de distribución radial zona hipoxantina – Oagua.	61
Figura 26. Resultado de la integración termodinámica.	63
Figura 27. Reacción $H19 \leftrightarrow H17$	64
Figura 28. Numeración en los átomos de la hipoxantina.	68

Introducción

1. Tautomerismo.

Los tautómeros son isómeros de ciertas moléculas que se intercambian entre sí a través de una reacción química llamada tautomerización. Durante la tautomerización una parte de la molécula migra a otra posición de la misma mientras se desarrolla un intercambio de enlaces. Este concepto es conocido generalmente como tautomerismo[1]. Como en muchas reacciones químicas, la tautomerización puede alcanzar un equilibrio químico entre los tautómeros, en el cual la población de estos depende de diversos factores, incluyendo la temperatura, pH, efectos del solvente, etc. [1].

Los tautomerismos más comunes son el ceto-enólico, lactama-lactima y enamina-imina (Figura 1); todos estos pertenecientes a un tipo de tautomerismo: el tautomerismo prototrópico, en el cual el átomo que migra durante la reacción es un protón. Otras clasificaciones del tautomerismo, pueden basarse en la forma en cómo ocurre la migración: el tautomerismo anular, donde el átomo o átomos que migran pueden ocupar dos o más lugares en un sistema heterocíclico (Figura 2); el tautomerismo anillo-cadena, el cual ocurre cuando la migración viene acompañada con un cambio en la estructura de la molécula de una cadena abierta a un anillo y viceversa (Figura 3); y finalmente el tautomerismo de valencia, que involucra una rápida reorganización de los electrones enlazantes (Figura 4) [2].

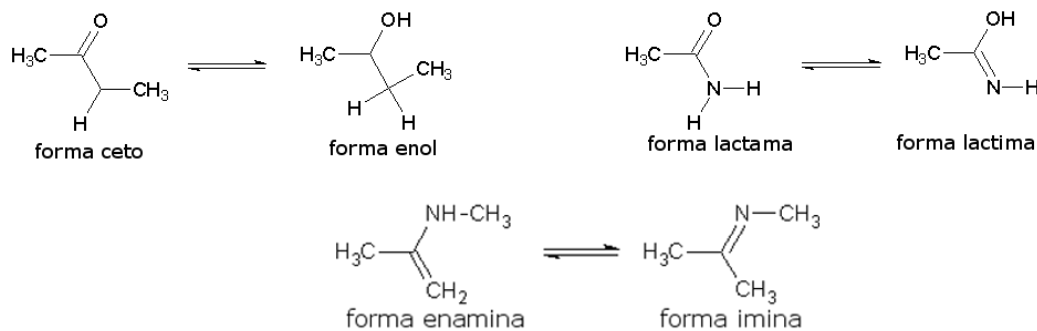


Figura 1. Tautomerismos ceto-enol (arriba izq.), lactima-lactama (arriba der.) e imina-enamina (abajo).

El tautomerismo es un concepto muy importante en química, bioquímica y farmacología debido a las interacciones específicas que ocurren durante una reacción química. Así, las diferencias estructurales entre tautómeros y el equilibrio establecido entre ellos es un factor determinante en la dirección a la cual una reacción química ocurre cuando hay un tautómero involucrado [3].

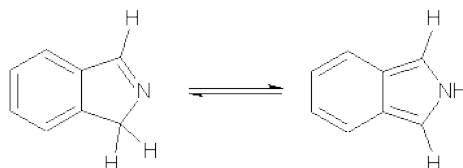


Figura 2. Tautomerismo anular en el isoindol.

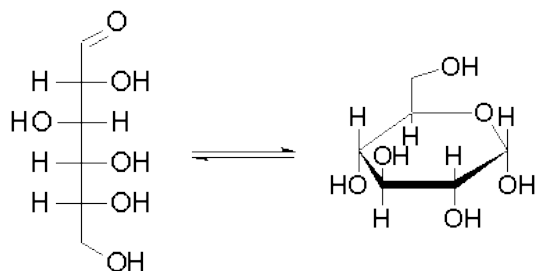


Figura 3. La glucosa se interconvierte de una cadena cerrada a otra abierta mediante una tautomerización.

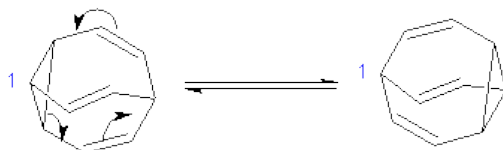


Figura 4. Tautomerismo de valencia en el bulvaleno.

2. La Hipoxantina

Los heterociclos orgánicos cuya estructura está formada por anillos de pirimidina unidos a estructuras imidazólicas son llamados **purinas**. Las purinas son moléculas de gran interés en diversas áreas de la ciencia ya que este tipo de heterociclos nitrogenados representan una de las especies más abundantes en la naturaleza[4], por ejemplo, los ácidos nucleicos tienen la estructura de una purina. Los compuestos mencionados presentan varios tautómeros ya que el intercambio de un átomo de H a diferentes partes del heterociclo es posible, generando así varios equilibrios químicos que le confieren cierta complejidad a la molécula (Figura 5).

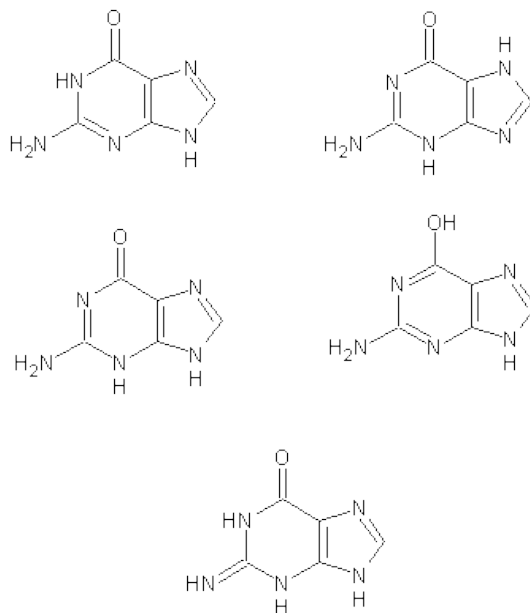


Figura 5. Algunos tautómeros de la guanina.

La hipoxantina (Figura 6), es un heterociclo derivado natural de la oxidación de la purina y se le puede encontrar en diversas estructuras bioquímicas junto a los ácidos nucleicos. La forma oxo de la hipoxantina es de considerable importancia debido a que se encuentra en pequeñas cantidades en el RNA de transferencia[5]. Además participa como intermediaria varios procesos metabólicos, formando Inosina (Figura 7), precursora del AMP (adenosín monofosfato) y GMP (guanosín monofosfato) [5]. Por otro lado, la adenina presente en el DNA puede ser oxidada por ácido nitroso y producir hipoxantina, la cual dentro de las cadenas de DNA es potencialmente mutagénica, ya que codifica como guanina y forma un par con la citosina durante la replicación del DNA, dando lugar a transiciones A-T → G-C [6].

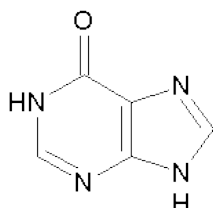


Figura 6. Hipoxantina en su forma oxo.

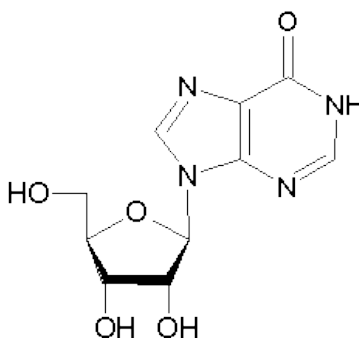


Figura 7. Inosina, precursora del AMP y GMP.

Nomenclatura de la hipoxantina.

Para hacer referencia a la hipoxantina sin ambigüedad se introduce su nomenclatura IUPAC [7]. La hipoxantina tiene como base a la purina, siendo ésta un heterociclo con anillos fusionados se numera como se muestra en la Figura 8. Generalmente los hidrógenos se localizan refiriéndose al número del átomo al cual están unidos.

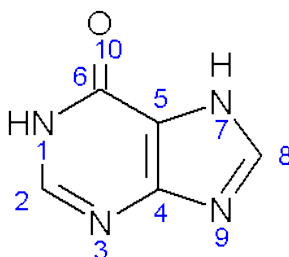


Figura 8. Numeración de los átomos de C, N y O de la hipoxantina.

Así, el nombre oficial de la hipoxantina es purin-6-ona en forma oxo (Figura 6, Figura 8). Sin embargo, debido a la variedad de tautómeros que puede formar, estos se diferencian de acuerdo a la posición de los hidrógenos; por consiguiente, el tautómero presentado en la Figura 8, es la *7H*-purin-6(1*H*)-ona.

Para simplificar la nomenclatura se hace uso solamente de los números a los cuales están unidos los hidrógenos que participen en alguna tautomerización. De esta manera, la hipoxantina N1-N9 hace referencia a aquella con átomos de hidrógeno en los átomos 1 y 9, los cuales son nitrógenos y corresponde a la forma mostrada en la Figura 6. Mientras que por otro lado la hipoxantina N1-N7 (Figura 8) señala a la molécula con los átomos de hidrógeno en las posiciones 1 y 7, que corresponden también a nitrógenos.

Para simplificar en expresiones matemáticas, fórmulas o subíndices, también se hace referencia a estas especies mediante los números, por ejemplo H17 o Hp17.

3. Tautomerismo de la Hipoxantina.

Los ácidos nucleicos pertenecen a un sistema de estudio de bastante interés. La intrincada red de interacciones de puente de hidrógeno que modula el reconocimiento entre las bases del DNA, así como su misma estructura, se basa en el hecho de que las bases de los ácidos nucleicos existen en determinados estados tautoméricos; de hecho, la importancia del equilibrio tautomérico en las bases de los ácidos nucleicos ha sido ampliamente enfatizada desde que se propuso la estructura del DNA por Watson y Crick [8]. Inclusive, diversos modelos que intentan explicar la mutación del DNA están basados en la idea general de que la existencia de formas tautoméricas no predominantes de las bases inducen la mutación al cambiar la estructura de la red de interacciones del DNA [9; 10]. Debido a esto, el estudio del tautomerismo en los ácidos nucleicos, tanto experimental como teórico, ha sido hecho de manera extensiva, enfocándose principalmente en la guanina y citosina [11- 21].

De esta muestra de estudios, se puede sintetizar que aún existe cierta incertidumbre en el tautomerismo de este tipo de compuestos, debido a:

- Limitantes experimentales: el estudio de una sola especie aislada o poco poblada, es muy difícil.
- Limitantes teóricas: complejidad en los cálculos en fase acuosa, principal medio de interés.

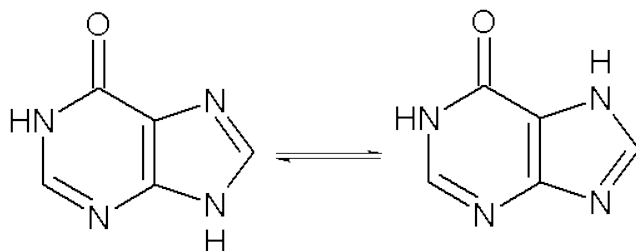


Figura 9. Una de las reacciones de tautomerización de la hipoxantina.

El caso de la hipoxantina no es diferente, ésta ha sido extensamente estudiada los últimos años [22-50;56-58], debido a sus características biológicas ya mencionadas en el apartado anterior y el tautomerismo que involucra (Figura 9). Su estudio como molécula prototipo a los ácidos nucleicos la hace candidata perfecta como punto de partida hacia un mejor entendimiento de los sistemas más complejos así como para probar nuevas metodologías de estudio experimentales o teóricas que se puedan extrapolar a cadenas de ácidos nucleicos.

Estudios experimentales que varían entre la espectroscopia UV [11; 12], espectroscopia foto electrónica UV [13; 14], espectroscopia IR en matrices de gas inerte [15] y rayos X [16], muestran que la tendencia en general es que la hipoxantina neutra se encuentre en forma oxo (cetónica) [17; 18] y que el tautómero N1-N9 sea el más abundante en solución así como en estado sólido, mientras que en fase gas es el tautómero N1-N7 el predominante [19; 20].

Diversos estudios teóricos han sido usados para poder establecer la diferencia energética relativa entre los tautómeros de la hipoxantina. En fase gaseosa, métodos semi empíricos han reportado diferencias de aproximadamente 4 kJ mol^{-1} , mostrando al tautómero N1-N9 como el más estable [21; 22]. Sin embargo al aumentar el nivel de cálculo, a métodos ab initio y DFT (Density Functional Theory, Teoría de funcionales de densidad), se ha encontrado que el tautómero N1-N7 es el más estable en fase gas [23; 24], lo que sugiere que, para métodos cuánticos, este tipo de moléculas deben ser estudiadas a un nivel de teoría que presente correlación electrónica confiable. De hecho una de las metodologías sugeridas para el cálculo teórico de energías en este tipo de compuestos en fase gaseosa es el uso de una optimización de geometría mediante DFT y a partir de éstas cálculos single point por medio de DFT o MP2 [19]. DFT B3LYP [25; 26] ha demostrado proporcionar resultados satisfactorios para el cálculo de la geometría de moléculas aisladas de los ácidos nucleicos usando al menos una base como 6-31G**, ya que ésta ha mostrado suficiente precisión para cálculos en monómeros del DNA [27; 28; 29].

Por otro lado, aunque teóricamente y experimentalmente se ha observado un predominio de los tautómeros N1-N7 y N1-N9, estudios de espectroscopia IR en matriz de Argón sugieren la presencia de una especie más en fase gaseosa, la hipoxantina N9 en forma enol (Figura 10), con una población del 5% frente a las otras especies predominantes [15].

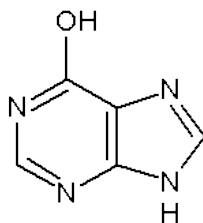


Figura 10. Tautómero enol observado en espectroscopia IR en matriz de argón.

Sin embargo, son las especies N1-N7 y N1-N9, debido a una reacción de tautomerización (Figura 11), las consideradas más abundantes en las soluciones acuosas y de DMSO [11; 30; 31].

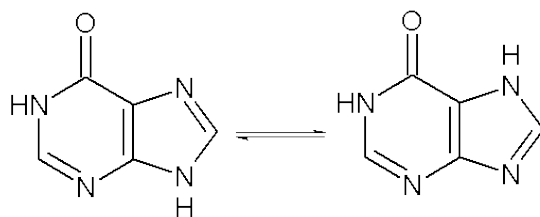


Figura 11. Tautomerismo de la hipoxantina N1-N9 a hipoxantina N1-N7.

Desafortunadamente, la información experimental que se tiene acerca del comportamiento de la hipoxantina aún no es cien por ciento concluyente: se desconoce concretamente el mecanismo por el cual puede suceder el tautomerismo, la población exacta de los tautómeros en solución, así como su dependencia a otros factores como la temperatura.

Sin embargo, se ha encontrado que el tautomerismo de la hipoxantina en solución, con respecto a las especies N1-N7 y N1-N9, parece ser fuertemente dependiente de la constante dieléctrica del medio [32; 33; 34]. En solución acuosa se ha visto una proporción mayor del tautómero N1-N9 [11], mientras que en solución de DMSO la proporción del tautómero N1-N7 es sólo ligeramente mayor [35]. Varios estudios muestran una tendencia: *aquella especie tautomérica con el mayor momento dipolar en fase gaseosa es la predominante en un medio polar* [32; 33; 34; 36; 37]. Mediante diferentes cálculos teóricos se ha podido establecer que la hipoxantina N1-N7 tiene un momento dipolar mucho menor que cualquiera de sus otros tautómeros [19; 24]. Dicha situación sugiere entonces un posible cambio en la estabilidad de los tautómeros dependiendo del medio: del predominio de la hipoxantina N1-N7 en fase gas al predominio de la hipoxantina N1-N9 en soluciones polares. Lo anterior es esperado ya que esta tendencia se ha observado en otros sistemas similares más simples, como en el tautomerismo de la 2-hidroxipiridina [38; 39].

En síntesis, de todos los estudios anteriormente mencionados, pueden obtenerse datos cuantitativos representativos:

- Las diferencias de energía relativa en fase gas de los dos tautómeros más estables se encuentra en el orden de 2 a 4 kJ mol⁻¹, siendo el tautómero N1-N7 favorecido [19].
- Se ha encontrado también que entre otras formas tautoméricas, N3 y enol, la diferencia de energía con respecto al tautómero N1-N7 es de alrededor de 33 kJ mol⁻¹. Estas especies son poco estables y contribuyen muy poco a la población total [19; 40].
- De acuerdo a varios estudios teóricos y experimentales, la población de tautómeros N1-N7 y N1-N9 parece ser en promedio cercana al 64% y 36% respectivamente en fase gas. En solución acuosa, lo que se ha observado es que la población de los tautómeros N1-N7 y N1-N9 tiende a ser de entre 40% y 25% a 60% y 75% respectivamente [19; 13; 15; 20; 40].
- El momento dipolar calculado teóricamente en fase gas de los tautómeros varía de entre 1.4 a 1.8 D en el caso del tautómero N1-N7 y de entre 5.2 a 5.8 D en el caso del tautómero N1-N9 [24].

4. Justificación: tautomerismo de la hipoxantina y simulaciones moleculares.

La idea del presente trabajo es estudiar la reacción de tautomerismo de la hipoxantina, con ello se abordan dos temáticas de interés actual:

- i. La hipoxantina. Esta molécula participa de manera activa en varios procesos bioquímicos importantes. Además, ya que su estructura le permite poseer un amplio repertorio de tautómeros, la identidad de las estructuras presentes en solución y los factores que regulan su población son de particular interés. Tal situación ha sido motivo de estudio por muchos años.
- ii. La reacción de tautomerización. Esta reacción química es bastante significativa, particularmente en el medio biológico, por lo que su estudio en ácidos nucleicos se ha hecho extensivo los últimos años. Y el interés no termina sólo en su relación con las biomoléculas, también es una reacción cuyo estudio ha permitido ampliar el conocimiento que se tiene de las reacciones químicas: transferencia de protones, rearrreglos, propiedades estructurales, etc. [1; 2]

En la sección anterior se han señalado trabajos teóricos con diversos estudios al tautomerismo de la hipoxantina principalmente utilizando métodos cuánticos y semiempíricos.

En lo que respecta a los trabajos en solución acuosa, muchos de ellos utilizan modelos continuos para representar al disolvente. Algunos trabajos con métodos cuánticos también estudian las interacciones agua – hipoxantina incluyendo una o más moléculas de agua, pero el costo computacional de un método cuántico hace muy complicado que pueda agregarse un gran número de moléculas de disolvente.

Debido a que en química, el medio físico en el cual se encuentra una reacción es de suma importancia para ésta, **se considera entonces la idea de utilizar un método para modelar la reacción de tautomerización tratando de involucrar explícitamente al disolvente** y así poder obtener nuevos datos que contribuyan a un mejor entendimiento del tautomerismo.

Particularmente, los métodos de simulación molecular se han utilizado extensivamente para el estudio de fluidos y pueden ampliarse para el estudio de ciertos fenómenos que pueden ocurrir en éstos, como la asociación y la reacción química [41; 42; 43; 44]. La diferencia con los métodos cuánticos recae en que los métodos de simulación molecular utilizan un enfoque diferente a la mecánica cuántica: el de la mecánica estadística. Una característica interesante, producto de esta diferencia, es que los métodos de simulación molecular suelen ser menos demandantes computacionalmente pudiendo a su vez llegar a estudiar sistemas con cientos de moléculas interactuando al mismo tiempo.

Lo anterior resulta *atractivo* ya que puede hacerse el estudio de una reacción en donde quieran incluirse cientos de moléculas de disolvente y ver la forma en cómo estas moléculas interaccionan con las moléculas reaccionantes.

Es conveniente mencionar que los métodos cuánticos, en donde la gran mayoría de los estudios del tautomerismo de la hipoxantina se han realizado, proporcionan datos rigurosos sobre la energética de átomos y moléculas así como un estimado de sus propiedades termodinámicas, por lo que la simulación molecular debe verse como un método para complementar estos y aportar así más referencias que lleven a un mejor entendimiento de la reacción química. Ambas metodologías también son complementarias a los datos experimentales, los cuales además proporcionan un sentido a todo el aparato teórico utilizado.

Una simulación molecular por ende ocupa un lugar sólido en el estudio de los sistemas reales (Figura 12), cuyo uso va desde la predicción de las propiedades de algunos sistemas así como para la exploración y estudio de estos.

De esta manera, con el uso de las simulaciones moleculares, debe ser posible calcular la población de los tautómeros de la hipoxantina en solución, incluyendo al disolvente de manera explícita, aportando así nuevos detalles al estudio de este tautomerismo, el cual actualmente continúa siendo objeto de estudio tanto experimental como teórico [12; 45; 46; 47].

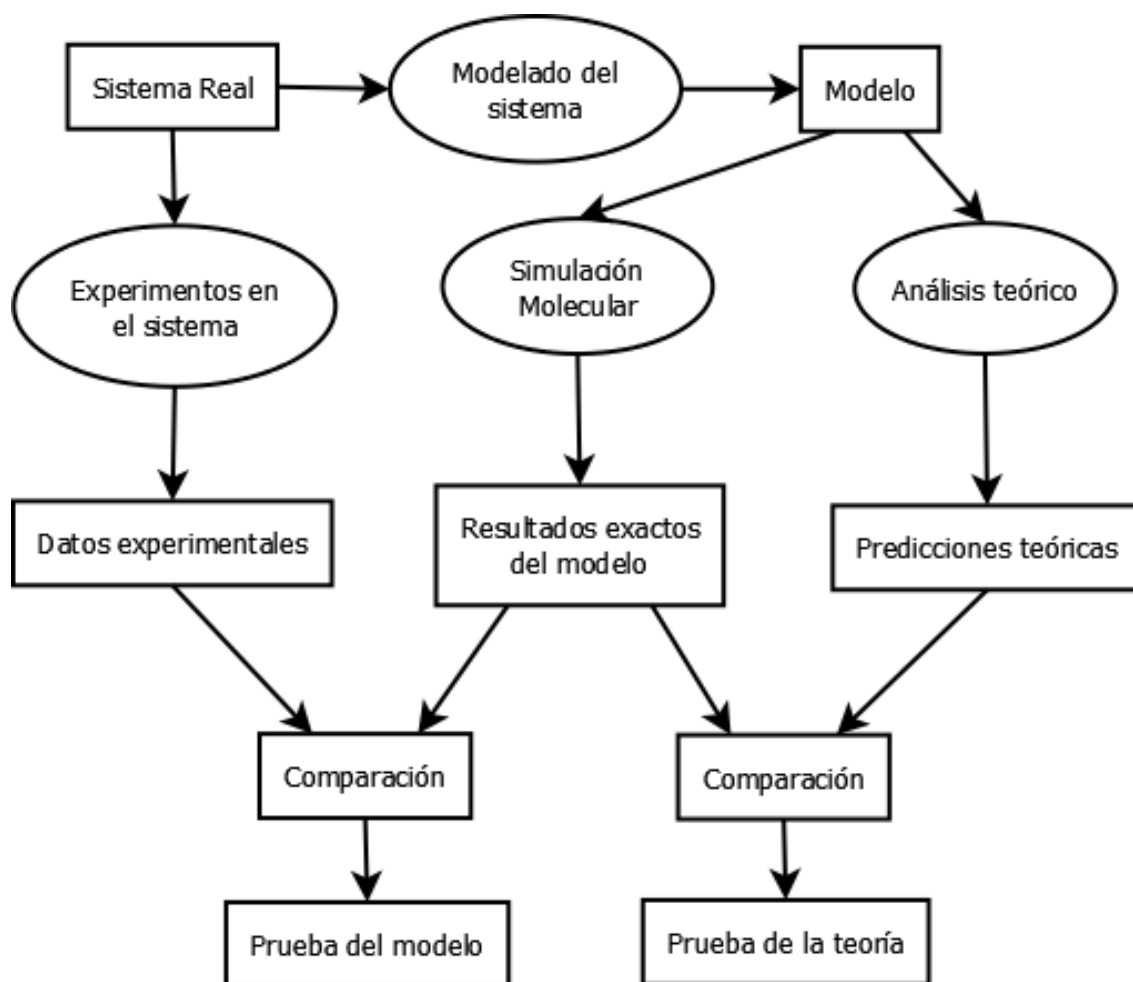


Figura 12. La conexión entre experimento, teoría y las simulaciones moleculares.

Debe mencionarse que no es la primera vez que un trabajo similar se realiza, pero tampoco existen muchos estudios con este enfoque. En la referencia [48], única que se encontró con un acercamiento similar, se hace un trabajo extensivo en la determinación de los parámetros para el modelado de las interacciones hipoxantina – agua mediante un modelo particular. Ahí se realizan simulaciones con el método de Monte Carlo de tautómeros de la hipoxantina en agua, pero el alcance dado en la referencia a la simulación molecular sólo es para el cálculo de la diferencia de energía entre tautómeros y a la estructura del agua alrededor de los tautómeros. El presente trabajo busca ir más adelante y realizar el cálculo de la población de los tautómeros en fase acuosa.

5. Objetivos y metas.

En el presente trabajo el estudio del tautomerismo de la hipoxantina se centra principalmente en cumplir los siguientes objetivos:

- La estimación de la energía libre de una reacción en solución acuosa mediante simulación molecular
- El cálculo de la constante de equilibrio del tautomerismo en solución acuosa.

Otras metas de interés incluyen:

- El uso de algún método que incluya el tratamiento de la respuesta del medio al momento dipolar de los tautómeros en las simulaciones moleculares.
- La posible comparación de dichos métodos si es que existieran más.
- Cálculos cuánticos en fase gas de las moléculas involucradas.
- Análisis de las propiedades estructurales del soluto y solvente de los tautómeros.

En los siguientes capítulos se cimentan todas las bases teóricas para poder alcanzar estos objetivos y también se ahonda sobre los términos usados en los objetivos y lo que implican.

Después del apartado teórico, se explica la metodología usada en profundidad y con ello se retoman los puntos dados como objetivos y metas para especificarlos en relación a los métodos de trabajo que se usarán.

6. Hipótesis

Las hipótesis planteadas en el trabajo son:

- De entre el par de tautómeros de la hipoxantina N1-N7 y N1-N9, uno de ellos estará mayormente favorecido en solución acuosa. Su proporción no será 50:50.
- La diferencia en la magnitud de los momentos dipolares de los tautómeros afecta la población de éstos. Por tal motivo, el momento dipolar debe ser una propiedad que debe modelarse lo mejor posible durante la simulación molecular.

A partir de las hipótesis es por lo que se elige tratar el problema del cálculo de la constante de equilibrio de la reacción, mediante una simulación molecular, utilizando un modelo polarizable.

Marco Teórico.

7. Simulaciones moleculares.

Para describir a un sistema macroscópico utilizamos propiedades que podemos medir en el laboratorio: Volumen (V), presión (p), temperatura (T), etc. De ahí deriva el uso de las ecuaciones fundamentales, pues otras propiedades pueden ser expresadas conociendo la ecuación fundamental y por ende, estas propiedades son función de aquellas que podemos medir. Así, estas pocas variables son capaces de describir el estado del sistema.

En una simulación molecular, se observan sistemas microscópicos compuestos por partículas cuyas propiedades se describen mediante la posición (\vec{r}) y el momento (\vec{p}) de cada una de éstas.

La conexión entre los resultados que puede tener un experimento y una simulación molecular la provee la mecánica estadística, la cual proporciona las relaciones entre el sistema macroscópico (p, N, V, T) y el microscópico (\vec{r}, \vec{p}). Tal conexión no es cien por ciento directa y depende mucho del modelado de las partículas.

Los dos tipos de simulaciones moleculares más usados son: los métodos de Monte Carlo y los métodos de dinámica molecular. Ambos métodos son usados extensivamente en el modelado del comportamiento de moléculas en diversos campos de estudio en las ciencias de materiales, química computacional y biología computacional [49; 50; 51; 52].

La diferencia principal entre ambos métodos estriba en que los métodos de dinámica molecular recrean el comportamiento del sistema en función del tiempo, resolviendo las ecuaciones del movimiento con el agregado intrínseco de los efectos de la temperatura. Los métodos de Monte Carlo por otro lado son usados para evaluar expresiones matemáticas complejas, las cuales pueden ser aquellas que modelen el estado del sistema en función de las interacciones energéticas entre partículas, encontrando el estado de mínima energía del sistema a determinadas condiciones en un momento fijo.

Dado que en el presente trabajo no se requiere el cálculo de ninguna propiedad dinámica o dependiente del tiempo, se elige el uso de las simulaciones por los métodos de Monte Carlo, cuyo desarrollo teórico se expone a continuación.

8. Métodos de Monte Carlo.

Antes de poder abordar una síntesis detallada sobre los métodos de Monte Carlo hay que definir ciertos conceptos:

El **espacio fase** (Γ) de una partícula. Por definición, *es un espacio $6N$ dimensional, consistente en un sistema de N partículas puntuales, que provee toda la información del*

estado mecánico del sistema: las tres componentes del momento y las tres componentes de la posición de cada partícula [49].

Un **colectivo** es una colección de puntos del espacio fase distribuidos de acuerdo a una densidad de probabilidad $\rho(\Gamma)$ [50; 49].

La densidad de probabilidad $\rho(\Gamma)$ es una función que depende de los parámetros macroscópicos representativos del sistema, esto es N, V, p, T , etc. De acuerdo a estos parámetros los colectivos se pueden dividir de acuerdo a las propiedades macroscópicas que mantienen constantes. Los colectivos más representativos son:

- Colectivo canónico. Volumen, temperatura y número de partículas constante.
- Colectivo microcanónico. Volumen, energía y número de partículas constante.
- Colectivo isobárico – isotérmico o NPT. Presión, temperatura y número de partículas constante.

El problema principal que se busca resolver es obtener una propiedad observable del sistema a partir de las propiedades microscópicas del mismo. Definidos los conceptos de espacio fase y colectivos, se puede decir entonces que toda propiedad instantánea A_{inst} del sistema debe ser función de algún punto del espacio fase (Γ):

$$A_{inst} = A(\Gamma)$$

Si el sistema evoluciona con el tiempo, el espacio fase cambia. Por lo tanto A_{inst} también cambia.

El postulado en el que se basa la mecánica estadística es que *una propiedad macroscópica observable A_{obs} es el promedio de A_{inst} a través de un enorme periodo de tiempo en el cual el sistema evoluciona [53].*

$$A_{obs} = \langle A(\Gamma(t)) \rangle_{tiempo} = \lim_{t_{obs} \rightarrow \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} A(\Gamma(t)) dt$$

Debido a la gran complejidad que tiene resolver una integral como la anterior cuando el número de partículas que se tiene es grande, en 1878 Gibbs sugirió que en vez de manejar un promedio del tiempo, se manejara un promedio de colectivo. Es decir:

$$A_{obs} = \langle A(\Gamma(t)) \rangle_{colectivo} = \sum_{\Gamma} A(\Gamma) \rho(\Gamma)$$

Entonces, el tratamiento del problema implica recorrer una gran cantidad de puntos del espacio fase representativos del colectivo, para que así, al realizar el cálculo, se tenga una A_{obs} consistente con un sistema real.

Para abordar el tratamiento de una forma más conveniente, se introduce el concepto de función de partición. La función de partición Q conecta la termodinámica clásica con el

sistema microscópico; se define un potencial ψ relacionado a la función de partición de acuerdo al colectivo (c) en el que se desea trabajar. Así:

$$\psi_c = -\ln Q_c$$

De hecho para el colectivo NVT se encuentra que $\psi_{NVT} = A/kT$, donde A es la energía libre de Helmholtz, o para el colectivo NPT la expresión es $\psi_{NPT} = G/kT$, donde G es la energía libre de Gibbs [53].

Luego definimos una función de “peso” del colectivo, $w_c(\Gamma)$, como una densidad de probabilidad $\rho(\Gamma)$ no normalizada, en donde la normalización se alcanza con la función de partición:

$$\rho(\Gamma) = \frac{w_c(\Gamma)}{Q_c}$$

tal que:

$$Q_c = \sum_{\Gamma} w_c(\Gamma)$$

Dadas estas condiciones se tiene que:

$$\langle A \rangle_c = \frac{\sum_{\Gamma} A(\Gamma) w_c(\Gamma)}{\sum_{\Gamma} w_c(\Gamma)} \quad (8.1)$$

Y consecuentemente puede conocerse A_{obs} .

Supongamos ahora un sistema de partículas en el colectivo canónico (o NVT). El sistema presenta las siguientes ecuaciones [53]:

$$w_{NVT} = \exp\left(\frac{-\mathcal{H}(\Gamma)}{kT}\right) \quad (8.2)$$

Y

$$Q_{NVT} = \sum_{\Gamma} \exp\left(\frac{-\mathcal{H}(\Gamma)}{kT}\right) \quad (8.3)$$

Donde k es la constante de Boltzmann, T la temperatura y \mathcal{H} el hamiltoniano del sistema. La función de partición del sistema en una aproximación clásica viene dada por:

$$Q_{NVT} = \frac{1}{h^{3N} N!} \int \exp\left(\frac{-\mathcal{H}(\vec{p}, \vec{r})}{kT}\right) d\vec{p} d\vec{r}$$

Con h la constante de planck y N el número de partículas.

Q_{NVT} puede ser dividida:

$$\begin{aligned} Q_{NVT} &= \frac{1}{h^{3N} N!} \int \exp\left(\frac{-\mathcal{K}(\vec{p})}{kT}\right) d\vec{p} \int \exp\left(\frac{-\mathcal{V}(\vec{r})}{kT}\right) d\vec{r} \\ &= Q_{NVT}^{\mathcal{K}} Q_{NVT}^{\mathcal{V}} \end{aligned} \quad (8.4)$$

Considerando que por definición, el Hamiltoniano es la suma de la energía cinética y potencial del sistema: $\mathcal{H} = \mathcal{K} + \mathcal{V}$.

La parte que comprende a la energía cinética por lo general suele ser posible evaluarla, quedando solo la parte de la energía potencial, conocida como *integral de configuración* Z:

$$Z = \int \exp\left(\frac{-\mathcal{V}(\vec{r})}{kT}\right) d\vec{r} \quad (8.5)$$

El tratamiento expuesto es de suma utilidad para las simulaciones moleculares, porque en una simulación molecular estas ecuaciones suelen ser el punto de partida.

Nótese que en el colectivo canónico, combinando las ecuaciones 8.1 y 8.2, en la combinación clásica:

$$\langle A \rangle_{NVT} = \frac{\int A(\vec{p}, \vec{r}) \exp(-\beta\mathcal{H}(\vec{p}, \vec{r})) d\vec{r}d\vec{p}}{\int \exp(-\beta\mathcal{H}(\vec{p}, \vec{r})) d\vec{r}d\vec{p}} \quad (8.6)$$

Donde se introduce β , relacionada con la temperatura y definida como $\beta = \frac{1}{kT}$

Un intento de obtener valores promedio de un colectivo es resolver la ecuación 8.6 numéricamente.

Como ya se ha mostrado (ecuaciones 8.3 y 8.4), el hamiltoniano contenido en la función de partición, puede separarse en una contribución de energía cinética y otra de energía potencial. La dependencia de la energía cinética, \mathcal{K} , es cuadrática en los momentos y se puede efectuar su integración; el problema radica en evaluar la parte de la energía potencial, \mathcal{V} . En general no es posible [49; 50]. Se tiene un problema, el cual consiste en evaluar una relación de la forma:

$$\frac{\int A(\vec{r}) \exp(-\beta\mathcal{V}(\vec{r})) d\vec{r}}{\int \exp(-\beta\mathcal{V}(\vec{r})) d\vec{r}}$$

Para lidiar con esto, sea $N(\vec{r})$ una densidad de probabilidad definida por:

$$N(\vec{r}) = \frac{\exp(-\beta\mathcal{V}(\vec{r}))}{Z} \quad (8.7)$$

Con Z, la integral de configuración, definida en 8.5.

Entonces, suponemos que somos capaces de generar diferentes configuraciones de acuerdo a esta densidad de probabilidad $N(\vec{r})$. Inmediatamente se tiene que:

$$\langle A \rangle = \langle A \rangle_{generadas}$$

El problema entonces se reescribe y ahora consiste en generar puntos en el espacio fase tal que se apeguen a la densidad de probabilidad 8.7.

Una cadena de Markov tiene los elementos necesarios para resolver el problema ya que se basa en generar puntos independientes de manera aleatoria, restringidos por una densidad de probabilidad [54; 55].

Debido a que la mejor forma de describir las transiciones entre puntos en una cadena de Markov es usando una matriz de transición, la explicación siguiente usa términos del álgebra matricial [54].

Definimos una matriz estocástica o de transición como una matriz $T \in M_{n \times n}$ tal que:

- Cada elemento de la matriz del renglón m y la columna n son no negativos.
- La suma de los elementos de cada renglón o columna es uno.

Sea la matriz $T \in M_{n \times n}$ tal que T^n tiene todos sus elementos positivos para toda n . Las siguientes afirmaciones son verdaderas:

- El valor propio de T es 1.
- Existe $\lim_{n \rightarrow \infty} T^n$
- $L = \lim_{n \rightarrow \infty} T^n$ es también una matriz de transición.
- Para todo vector \vec{x} , $\lim_{n \rightarrow \infty} (T^n \vec{x}) = \vec{v}$

La última de las afirmaciones es la de mayor interés. La idea es proponer una matriz de transición de manera que nuestro vector final \vec{v} es la densidad de probabilidad dada por la ecuación 8.7. No importa por cual vector de probabilidad \vec{x} se empiece, pues las propiedades de la cadena de Markov siempre hacen que, después de un gran número de pasos, la operación converja a \vec{v} , siempre y cuando la matriz de transición este bien definida de acuerdo al problema.

Una manera de obtener las matrices de transición es la siguiente: supongamos un sistema de Markov que pasa del estado x al y , dos puntos del espacio fase. Sea T una matriz de transición y N una densidad de probabilidad, entonces:

$$N_x T_{x \rightarrow y} = N_y T_{y \rightarrow x} \quad (8.8)$$

Esto es, la probabilidad de que el sistema pase del estado x al y es la misma que si pasase del estado y al x . Esta condición es suficiente para que la cadena de Markov sea convergente aunque no necesaria [49]. La ecuación 8.8 es interesante porque describe una propiedad existente en los sistemas químicos, *reversibilidad microscópica*.

En 1953, Metropolis y colaboradores [56] propusieron que la matriz T se construyera de manera que:

$$T_{x \rightarrow y} = \alpha_{x \rightarrow y} \times acc_{x \rightarrow y} \quad (8.9)$$

Donde α es una matriz de transición que nos da la probabilidad de realizar o no el cambio de x a y en un movimiento de prueba. α Es generalmente llamada matriz fundamental de la cadena de Markov y en el método original de Metrópolis, se le considera simétrica ($\alpha_{x \rightarrow y} = \alpha_{y \rightarrow x}$). acc Es una matriz que determina la probabilidad de aceptar o rechazar el movimiento prueba de x a y .

Luego de combinar la ecuación 8.8 y 8.9 se tiene que:

$$N_x acc_{x \rightarrow y} = N_y acc_{y \rightarrow z}$$

E inmediatamente:

$$\frac{acc_{x \rightarrow y}}{acc_{y \rightarrow z}} = \frac{N_y}{N_x}$$

Y de la definición de la ecuación 8.7:

$$\frac{N_y}{N_x} = \frac{\exp(-\beta\mathcal{V}_y)}{\exp(-\beta\mathcal{V}_x)} = \exp(-\beta\Delta\mathcal{V}) \quad (8.10)$$

Nuevamente, se deben proponer formas para las matrices acc . Metropolis y colaboradores propusieron finalmente que:

$$acc_{x \rightarrow y} = \begin{cases} \frac{N_y}{N_x} & \text{si } N_y \geq N_x \\ 1 & \text{si } N_y < N_x \end{cases} \quad (8.11)$$

De esta manera, todo el aparato matemático nos muestra que para calcular una propiedad promedio A en el colectivo canónico, se debe recorrer una trayectoria en el espacio fase que corresponda al colectivo canónico (NVT) y promediar A a lo largo de esa trayectoria.

El método de Metropolis Monte Carlo se basa entonces en recorrer una trayectoria adecuada en el espacio fase, de acuerdo al esquema propuesto por Metropolis, utilizando masivamente generadores de números aleatorios con el fin de crear puntos en el espacio fase tales que la probabilidad de que el sistema evolucione de un punto a otro este dada por $\exp(-\beta\Delta\mathcal{V})$ de acuerdo a las ecuaciones 8.10 y 8.11.

9. Modelado: cálculo de la energía potencial.

Se ha hablado del formalismo en el que están basadas las simulaciones moleculares y en particular el método de Metropolis Monte Carlo en el colectivo canónico, pero como se mencionó en la sección 7, el enlace entre el formalismo de la mecánica estadística y la medición de observables macroscópicos se completa con la elección del modelo con el cual se van a representar las interacciones entre las partículas que componen al sistema.

En el método de Monte Carlo, la mayor importancia recae en el cálculo de la energía potencial del sistema. Como se mostró la probabilidad de aceptar un movimiento de un estado a otro depende de la diferencia de energía potencial entre estos dos estados:

$$\frac{N_y}{N_x} = \exp(-\beta\Delta\mathcal{V}) \quad (9.1)$$

El enfoque utilizado casi de manera general consiste en descomponer \mathcal{V} en un conjunto de funciones usadas para describir todas las interacciones en un sistema [49]. Por lo general estas funciones son términos que involucran interacciones que se dan en la misma partícula, entre pares de partículas, tripletes, etc.

Sin embargo tomar en cuenta todas las interacciones que se presentan es generalmente imposible en un sistema de N partículas. La aproximación por pares suele ser suficiente para el cálculo de la energía potencial. Esta aproximación consiste en calcular la contribución a la energía potencial de las interacciones que se den hasta entre dos cuerpos. Esto es:

$$\mathcal{V} \approx \sum_i U_i(\vec{r}_i) + \sum_i \sum_{j>i} U_{ij}(\vec{r}_{ij})$$

Donde U_i representa la contribución a la energía potencial dada por los efectos sobre cada partícula de manera individual; U_{ij} representa la contribución debida a las interacciones entre los pares de partículas i y j. A las formas matemáticas de U_{ij} se le suelen llamar potenciales entre pares.

Ejemplos de potenciales entre pares comunes son:

- De esferas duras [57], de pozo cuadrado [58] y de esferas blandas [59] para describir materiales ideales.
- De Lennard Jones, para describir materiales como el argón que sólo tienen interacciones de repulsión y atracción débiles (fuerzas de Van der Waals) [60].
- Electrostático, el cual describe materiales con carga.

En una simulación molecular comúnmente se trabaja con *campos de fuerza*. El concepto de campo de fuerza está relacionado con lo planteado arriba. Un campo de fuerza son funciones y parámetros derivados teórica o experimentalmente que tienen como objeto describir la energía potencial de un sistema real. Generalmente un campo de fuerza se construye acoplando potenciales, como el de Lennard Jones con el potencial electrostático (ecuación 9.2), de manera que las constantes puedan ser parametrizadas mediante el ajuste a alguna propiedad teórica o experimental.

$$U_{ij}(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + C \frac{q_i q_j}{r_{ij}} \quad (9.2)$$

A partir de esta expresión si se quieren añadir efectos adicionales a las fuerzas de Van der Waals e interacciones electrostáticas, se van adicionando más términos a la ecuación.

La selección de un campo de fuerza adecuado es primordial en una simulación molecular ya que de este dependen todos los cálculos y resultados que pueda arrojar. Por lo tanto es de suma importancia que el campo de fuerza describa correctamente las interacciones del sistema en estudio.

i. El caso de los sistemas acuosos.

El sistema más modelado desde el advenimiento de las simulaciones moleculares es el del agua debido a su presencia en la mayor parte de los procesos químicos, biológicos y físicos [61]. Sin embargo, uno de los problemas observados en el modelado del agua es que el cálculo del momento dipolar o de la constante dieléctrica siempre suele ser mal estimado por los campos de fuerza construidos solamente a partir de la ecuación 9.2, como el SPC [62] o TIP3P [63]. Se descubrió entonces que para estimar mejor estas cantidades era necesario incluir una contribución más en este sistema: las interacciones dipolo – dipolo inducidos [64]. La manera más eficiente de hacer esto es introducir los efectos de *polarización* en el modelo [65; 64].

Por lo tanto una simulación molecular de un sistema acuoso que tenga como fin el cálculo de una propiedad dependiente del momento dipolar, tiene que introducir la polarizabilidad del sistema para poder tener un modelado correcto.

Por consiguiente, si en la introducción se habló acerca de la dependencia del momento dipolar en la población relativa de los tautómeros de la hipoxantina en solución acuosa, lo más correcto debería ser, por lo tanto, usar un modelo para el agua y los tautómeros que incluya los efectos de la polarizabilidad. Tal concepto y sus modelos se discuten a continuación.

10. Polarizabilidad y su modelado.

La polarización se refiere a la redistribución de la densidad electrónica de una partícula debido a la acción de un campo eléctrico. En el caso de un conjunto de moléculas, estas tenderán a cambiar su distribución electrónica si se ven afectadas por un campo eléctrico, afectando por lo tanto propiedades que dependan de la distribución electrónica.

Así, en términos de interacciones moleculares la polarización conduce a la *no-aditividad*, esto es, la interacción molecular existente entre dos moléculas cambia, debido al rearrreglo de la densidad electrónica, cuando una de las moléculas es polarizada por una tercera. Es por ello que los sistemas polarizables modelados a partir de interacciones aditivas no producen resultados satisfactorios cuando las propiedades que se quieren calcular dependen fuertemente de la polarización del sistema.

El cambio en la densidad electrónica puede ser caracterizado por cambios en las cargas monopolo, el momento dipolar o en los momentos de mayor orden de la molécula. Es por ello que en general, los métodos para modelar la polarización en un sistema polarizable introducen momentos dipolares inducibles o cargas fluctuantes para cumplir con la *no-aditividad*.

De acuerdo al tratamiento de la polarizabilidad, los modelos se pueden clasificar de la siguiente manera:

- Modelos no polarizables.
- Modelos polarizables.
 - Modelos de dipolos puntuales polarizables.
 - Modelos de capa o *Shell*.
 - Modelos de ecualización de electronegatividad.
 - Modelos semiempíricos.

i. Modelos no polarizables

Antes de discutir los modelos que incluyen polarización, es útil considerar como punto de partida los modelos no polarizables y el porqué es necesario el desarrollo de modelos polarizables.

Un modelo típico usado en simulación molecular es uno del estilo[49; 50]:

$$V(r) = \sum_{i < j} 4\epsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{r_{ij}} \quad (10.1)$$

El cual permite cuantificar la energía de las interacciones intermoleculares, V , como una suma de una contribución de atracción – dispersión tipo Lennard-Jones (LJ) y una contribución electrostática. En la formula, ϵ y σ son parámetros que describen la energía y su influencia respecto a la distancia, respectivamente; r_{ij} es la distancia entre los átomos i y j no enlazados. Finalmente q_i es la carga respectiva de cada átomo i .

A primera instancia, pareciera que estos modelos no tienen ningún término que refleje la polarizabilidad del sistema. No obstante, la contribución de LJ contiene una parte repulsiva a corto alcance que decrece en el orden de r^{12} y una parte atractiva de largo alcance que decrece en el orden de r^6 [66]. La parte atractiva tiene una dependencia similar a la energía de dispersión de London dipolo-dipolo, que es expresada como $-\frac{\alpha^2}{r^6}$ para dos partículas con polarizabilidades α . Sin embargo a pesar del parecido en las expresiones los parámetros ϵ y σ generalmente no son asignados de acuerdo al valor de α . A pesar de esto, al ser escogidos de tal manera que puedan reproducir curvas de potencial, se dice que los parámetros de LJ “incluyen” de manera implícita y promedio la polarizabilidad del sistema [67].

Otra manera de introducir efectos de polarizabilidad es mediante las cargas en la contribución electrostática. Generalmente, estas q_i son asignadas de tal manera que el momento dipolar en fase gas sea consistente, o también se eligen de manera que éstas reproduzcan de la mejor manera posible el potencial electrostático (ESP) en fase gas, obtenido de cálculos *ab-initio*. Por ello, a estas q_i se les conoce también como cargas electrostáticas *mejoradas* y son estrictamente obtenidas por parametrización [68] o por cálculos cuánticos [69]. Aunque las cargas mejoradas tratan la polarización de manera efectiva, estas:

- no pueden mostrar la dependencia correcta de las distribuciones de carga de acuerdo al estado del sistema;

- tampoco pueden responder dinámicamente a las fluctuaciones del campo eléctrico debido al movimiento de las moléculas en una simulación molecular.

El campo eléctrico, así como el momento dipolar y la distribución de carga, dependen fuertemente del estado y composición del sistema. También dependen de las fluctuaciones en la distribución electrostática resultado de los movimientos moleculares, los cuales ocurren en el orden de picosegundos [70]. Es por ello que el tratamiento de estos efectos no puede ser cuantificado rigurosamente con modelos no polarizables, los cuales promedian o simplemente no pueden tomar en cuenta todos estos factores correctamente. Por esto mismo se requiere el desarrollo de modelos polarizables para el tratamiento de este tipo de propiedades.

ii. Modelos de dipolos puntuales polarizables.

Uno de los métodos para incluir la polarización se basa en la adición de dipolos puntuales inducidos en uno o todos los átomos de la molécula a modelar. Lo cual provee de un nuevo término al campo de fuerza que se vaya a utilizar (por ejemplo la ec. 10.1). Este método ha sido aplicado con éxito a una gran variedad de sistemas que van desde los gases nobles hasta las proteínas [68; 70].

Si $\vec{\mu}_i$ es el momento dipolar inducido en un sitio i , éste es proporcional al campo eléctrico \vec{E}_i en ese sitio. La constante de proporcionalidad viene dada por el tensor de polarizabilidad α_i . La expresión para $\vec{\mu}_i$ es entonces:

$$\vec{\mu}_i = \alpha_i \vec{E}_i = \alpha_i \left[\vec{E}_i^0 - \sum_{j \neq i} T_{ij} \vec{\mu}_j \right] \quad (10.2)$$

Donde \vec{E}_i^0 es el campo eléctrico debido a las cargas permanentes. Los dipolos inducidos interactúan a través del tensor de campo dipolar, T_{ij} :

$$T_{ij} = \frac{1}{r^3} I + \frac{3}{r^5} \begin{pmatrix} x^2 & xy & xz \\ yx & y^2 & yz \\ zx & zy & z^2 \end{pmatrix}$$

Siendo I la matriz identidad, r la distancia entre dos sitios i y j ; y finalmente x, y, z son los componentes cartesianos del vector que une i con j .

Así, la energía que proveen los dipolos inducidos, U_{ind} , puede ser expresada de esta manera:

$$U_{ind} = U_{stat} + U_{\mu\mu} + U_{pol}$$

Los tres términos en los que puede ser separada U_{ind} son:

- La contribución U_{stat} es la energía de interacción de los N dipolos inducidos del sistema con el campo permanente:

$$U_{stat} = - \sum_{i=0}^N \vec{\mu}_i \cdot \vec{E}_i^0 \quad (10.3)$$

- La energía debida a la interacción dipolo inducido – dipolo inducido, $U_{\mu\mu}$:

$$U_{\mu\mu} = \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N \vec{\mu}_i \cdot T_{ij} \cdot \vec{\mu}_j \quad (10.4)$$

- La energía de polarización U_{pol} :

$$U_{pol} = \frac{1}{2} \sum_{i=1}^N \vec{\mu}_i \cdot \vec{E}_i \quad (10.5)$$

Al sumar los términos se obtiene:

$$U_{ind} = \sum_{i=1}^N \vec{\mu}_i \cdot \left[-\vec{E}_i^0 + \frac{1}{2} \sum_{j \neq i} T_{ij} \cdot \vec{\mu}_j + \frac{1}{2} \vec{E}_i \right]$$

La cual se reduce a:

$$U_{ind} = -\frac{1}{2} \sum_{i=1}^N \vec{\mu}_i \cdot \vec{E}_i^0 \quad (10.6)$$

Usando:

$$\vec{E}_i = \vec{E}_i^0 - \sum_{j \neq i} T_{ij} \vec{\mu}_j$$

La ecuación 10.6 revela que la contribución a la energía de los dipolos inducidos depende del producto punto de los dipolos inducidos y el campo permanente debido a las cargas, no del campo total [71].

Finalmente, el campo permanente en el sitio i viene dado por:

$$\vec{E}_i^0 = \sum_{j \neq i} \frac{q_j \vec{r}_{ij}}{r_{ij}^3}$$

Este método resulta bastante atractivo ya que su implementación es muy directa. Al agregar la contribución de los dipolos inducidos solo una nueva variable se presenta, la polarizabilidad $\vec{\mu}_i$. Esta cantidad puede ser tomada de la literatura en algunos casos y en otros puede ser calculado por métodos cuánticos [72].

iii. Modelos polarizables de capa o Shell.

El modelo discutido anteriormente se basa en tratar la respuesta a la polarización usando dipolos puntuales en uno o más centros situados en una molécula. Un enfoque alternativo es modelar la respuesta a la polarización de la molécula utilizando dipolos de longitud finita representados con un par de cargas puntuales.

Se describe entonces a cada ión o átomo polarizable como un par de cargas puntuales, separadas cierta distancia la cual es variable. Estas cargas son llamadas *core*¹ y *shell*², puesto que la carga core esta situada en el centro atómico (núcleo) y la carga shell fuera de éste. Ambas cargas están conectadas entre si por medio de un oscilador armónico en el más simple de los casos. La magnitud de las cargas, en general, es un valor ajustable en el modelo y una vez escogidas son constantes en una simulación [73; 74].

El efecto de la polarización es entonces mostrado con el desplazamiento relativo de las cargas core y shell. De esta manera el momento dipolar $\vec{\mu}_i$, en el sitio i , viene dado por:

$$\vec{\mu}_i = -q_i \vec{d}_i$$

Donde q_i es la carga del átomo i y \vec{d}_i la distancia que separa las cargas. En un sitio, generalmente, las cargas core y shell son iguales pero de signo opuesto. Así, la energía que provee el modelo shell, U_{ind} , puede ser expresada de esta manera:

$$U_{ind} = U_{stat} + U_{\mu\mu} + U_{pol}$$

Y la contribución U_{pol} es obtenida usando el “resorte armónico” que separa las cargas:

$$U_{pol} = \frac{1}{2} \sum_{i=1}^N k_i d_i^2$$

La interacción entre sitios polarizables independientes, $U_{\mu\mu}$, es la suma de las interacciones carga - carga entre los sitios:

$$U_{\mu\mu} = \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N q_i q_j \left[\frac{1}{|\vec{r}_{ij}|} - \frac{1}{|\vec{r}_{ij} - \vec{d}_j|} + \frac{1}{|\vec{r}_{ij} + \vec{d}_i|} - \frac{1}{|\vec{r}_{ij} - \vec{d}_j + \vec{d}_i|} \right] \quad (10.7)$$

Las interacciones de Coulomb entre las cargas core y shell de un mismo sitio no se toman en cuenta.

¹ Del inglés núcleo.

² Del inglés capa, cascara, concha. Hace esa referencia por estar *sobre* el núcleo.

Se puede observar que si se utilizan sólo términos carga – carga esta contribución entre los dipolos, ya está presente en ecuaciones de campo de fuerza con potenciales electrostáticos, (como en la ec. 10.1), por lo que en estas condiciones no hay nuevos tipos de interacciones como en el caso del modelo de dipolos puntuales polarizables.

Luego la interacción de los dipolos con el campo permanente es la suma de los términos de cada sitio cargado:

$$U_{stat} = - \sum_{i=1}^N q_i \left[\vec{r}_i \cdot \vec{E}_i^0 - (\vec{r}_i + \vec{d}_i) \cdot \vec{E}_i^{0'} \right] \quad (10.8)$$

Donde \vec{E}_i^0 es el campo permanente en \vec{r}_i lugar donde está ubicada la carga core y $\vec{E}_i^{0'}$ es el campo permanente ubicado en $\vec{r}_i + \vec{d}_i$ que es donde está la carga shell. Por lo general $\vec{E}_i^0 \neq \vec{E}_i^{0'}$.

Las ecuaciones 10.8 y 10.7 corresponden directamente a 10.3 y 10.4 respectivamente en el caso de dipolos separados a una distancia finita. Cuando la separación de las cargas tiende a cero, el modelo shell es equivalente al modelo de dipolo puntual polarizable.

A diferencia del método anterior, existe una nueva variable que debe de ser tomada como un grado de libertad más en una simulación: la distancia entre las cargas shell y core. La forma de lidiar con esto es utilizar un enfoque adiabático o dinámico [72].

El enfoque adiabático tiene como fundamento la aproximación adiabática. En una simulación, después de que una molécula se ha movido, las cargas de todos los sitios que representan los dipolos son actualizadas inmediatamente de tal manera que estas ocupan nuevas posiciones que minimizan la energía total del sistema. El enfoque dinámico, permite que las cargas se muevan a diferentes lugares tratándolas como un sistema de partículas independientes, explorando así posiciones que no necesariamente minimizan la energía total del sistema. Sobre estos conceptos se abordara una discusión más amplia en la sección 12.

Finalmente, la contribución a la energía potencial usando el modelo shell es:

$$U_{ind}(\{\vec{r}_i\}, \{\vec{d}_i\}) = \sum_{i=1}^N \left\{ \frac{1}{2} k_i d_i^2 + q_i \left[\vec{r}_i \cdot \vec{E}_i^0 - (\vec{r}_i + \vec{d}_i) \cdot \vec{E}_i^{0'} \right] \right. \\ \left. + \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i} \left[\frac{1}{|\vec{r}_{ij}|} - \frac{1}{|\vec{r}_{ij} - \vec{d}_j|} \right] \right. \\ \left. + \frac{1}{|\vec{r}_{ij} + \vec{d}_i| |\vec{r}_{ij} - \vec{d}_j + \vec{d}_i|} \right\} \quad (10.9)$$

Los modelos shell han sido usados satisfactoriamente en una gran variedad de sistemas y su mayor aplicación la han encontrado en la simulación de materiales iónicos, sales fundidas, especies poliméricas y moleculares[75; 76].

iv. Modelos de ecualización de electronegatividad.

La polarizabilidad también puede introducirse en los modelos no polarizables (ecuación. 10.1) haciendo que la magnitud de las cargas electrostáticas cambie dependiendo del campo eléctrico del entorno. La ventaja de este enfoque es que la polarizabilidad se introduce sin necesidad de aumentar el número de interacciones carga – carga. Además se tiene la ventaja conceptual de que las interacciones electrostáticas polarizables y permanentes son tratadas con las mismas variables. Una manera de lograr que las cargas se acoplen a los cambios que presente el medio es mediante la ecualización de la electronegatividad [72].

La energía requerida para crear una carga q , en un átomo puede expresarse como una expansión en serie de Taylor:

$$U(q) = E^0 + \chi^0 q + \frac{1}{2} J q^2 + \dots$$

Si consideramos que la serie truncada a segundo orden es válida para cargas de magnitud a la carga elemental³, entonces, de las definiciones de potencial de ionización ($IP = U(1) - U(0)$) y afinidad electrónica ($EA = U(-1) - U(0)$), los coeficientes de la serie de Taylor son:

$$\chi^0 = \frac{IP + EA}{2}$$

$$J = IP - EA$$

Hay que señalar que χ^0 es la definición de la electronegatividad de Mulliken [77] y que $\frac{J}{2}$ es la dureza del átomo, η [78].

Cuando dos átomos se unen de tal manera que forman una molécula, la energía de las cargas en este modelo es descrita como:

$$U(\vec{q}) = \sum_i (E_i^0 + \chi_i^0 q_i + \frac{1}{2} J_{ii} q_i^2) + \sum_i \sum_{j>i} J_{ij}(r_{ij}) q_i q_j$$

Donde el vector \vec{q} representa el conjunto de las cargas q_i en la molécula formada. El coeficiente J_{ij} depende de la distancia entre i y j , que a largas distancias debería ser igual a $\frac{1}{r_{ij}}$; a distancias cortas, debido al apantallamiento de las cargas, el último término se convierte en:

$$\int \frac{\rho_i(\vec{r}_i) \rho_j(\vec{r}_j)}{|\vec{r}_i - \vec{r}_j - \vec{r}|} d\vec{r}_i d\vec{r}_j$$

Donde las cargas se encuentran deslocalizadas en distribuciones de carga $\rho(\vec{r})$ [65].

³ $e = -1.602564 \times 10^{-19}$ C

Las cargas parciales en cada átomo se encuentran minimizando la energía total de la molécula formada, con la condición de que la energía total del sistema se conserve.

$$\sum_i q_i = q_{tot}$$

La conservación de la carga puede ser reforzada con el uso de un multiplicador indeterminado λ ,

$$U(\vec{q}) = U(\vec{q}) - \lambda \left(\sum_i q_i - q_{tot} \right)$$

Y puede minimizarse la energía para cada q_i bajo la condición de que $q_{tot} = 0$ (la molécula es neutra), resultando:

$$\left(\frac{\partial U}{\partial q_i} \right) - \lambda = 0$$

Dado que $\left(\frac{\partial U}{\partial q_i} \right)$ resulta en el mismo multiplicador indeterminado para cada átomo, esta cantidad debe ser idéntica para todos los átomos en la molécula[79]:

$$\left(\frac{\partial U}{\partial q_i} \right) = \left(\frac{\partial U}{\partial q_j} \right)$$

Se identifica a $\left(\frac{\partial U}{\partial q} \right)$ como la definición de la electronegatividad según Mulliken [77], por lo que:

$$\chi_i \equiv \left(\frac{\partial U}{\partial q_i} \right) = \chi_i^0 + J_{ii}q_i + \sum_{j \neq i} J_{ij}(r_{ij})q_j$$

Nótese que la electronegatividad del átomo i dentro de una molécula, χ_i , es diferente a aquella del átomo aislado, χ_i^0 y además depende de las cargas de los otros átomos, de la dureza del átomo y de la interacción con otros átomos por medio del coeficiente $J_{ij}(r_{ij})$.

Así, la contribución a la energía dada por este modelo es:

$$U(\{q\}, \{r\}) = \sum_{\alpha} \left(\sum_{i \in \alpha} \chi_i^0 q_i + \frac{1}{2} \sum_{i \in \alpha} \sum_{j \in \beta} q_i q_j J_{ij}(r_{ij}) - E_{\alpha}^{gas} \right) + \sum_{\alpha} \sum_{\beta > \alpha} \sum_{i \in \alpha} \sum_{j \in \beta} q_i q_j J_{ij}(r_{ij})$$

Donde α y β representan dos moléculas e i y j los sitios o átomos en estas. E_{α}^{gas} es la energía de la molécula en fase gas; esta contribución a la energía también incluye el término coulombico, así que esta expresión sustituye al termino coulombico en campos de fuerza como el presentado en la ecuación 10.1. De esta manera, las cargas q_i son variables

independientes y la respuesta a la polarización es reflejada como la variación en la magnitud de estas cargas.

El modelo fue primeramente propuesto por Sanderson [80] y se ha visto que con una parametrización apropiada el modelo ha probado ser exitoso en una gran variedad de casos [81].

v. Modelos polarizables semiempíricos.

Este tratamiento de la polarización es un paso adelante hacia métodos completamente *ab initio*. Generalmente los modelos citados anteriormente no toman en cuenta los efectos de correlación e intercambio electrónico totalmente pues estos volverían el cálculo poco práctico [82].

En general como ha sido expresado por Gao [83], la idea es calcular la energía total del sistema, E , como:

$$E = E_{pol} + \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N E_{ij}$$

Suponiendo que la densidad electrónica no se transfiere entre moléculas, se proponen energías de interacción, E_{ij} parecidas a esta:

$$E_{ij} = \frac{1}{2} (\langle \Psi_i | \hat{H}_{ij} | \Psi_i \rangle + \langle \Psi_j | \hat{H}_{ji} | \Psi_j \rangle) + E_{LJ}$$

Siendo

$$\hat{H}_{ij}(\Psi_j) = - \sum_{\alpha=1}^M V_{\alpha}(\Psi_j) + \sum_{\alpha=1}^A Z_{\alpha}(i) V_{\alpha}(\Psi_j)$$

Donde $V(\Psi_j)$ es el potencial electrostático ejercido por una molécula j en el electrón a o en el núcleo de la molécula i .

De ahí que:

$$E_{pol} = \sum_{i=1}^N (\langle \Psi_i | \hat{H}_i^0 | \Psi_i \rangle - \langle \Psi_i^0 | \hat{H}_i^0 | \Psi_i^0 \rangle)$$

Da la energía de polarización del sistema. Aquí, $\langle \Psi_i^0 | \hat{H}_i^0 | \Psi_i^0 \rangle$ es la energía de la molécula en estado aislado. Por ende, \hat{H}_i^0 es la parte del Hamiltoniano que contiene solo las interacciones electrón - electrón, la interacción electrón - núcleo y la energía cinética dentro una misma molécula. El tratamiento de la polarizabilidad se efectúa variando la magnitud de las cargas a lo largo de la simulación [83].

11. El Modelo MCHO.

En 1990, Saint Martin, Medina Llanos y Ortega Blake propusieron el potencial de cargas móviles y oscilador armónico (MCHO, Mobile Charges and Harmonic Oscillator) [84] como una propuesta de un campo de fuerza que maneja el modelado de la polarización.

El modelo MCHO, tipo átomo – átomo, propone representar a la molécula como una colección de centros con dos tipos de sitios: sitios nucleares y sitios de cargas (o electrónicos) de tal manera que los sitios electrónicos son móviles alrededor del sitio nuclear de acuerdo al modelo de oscilador armónico (Figura 13). Este enfoque permite incluir los efectos de polarización y de no aditividad, ya que la posición de las cargas es variable de acuerdo al campo eléctrico que experimenten. Por otro lado, los sitios nucleares se encuentran repartidos de tal manera que puedan representar la estructura geométrica experimental de la molécula en cuestión. El modelo MCHO tiene características de un modelo polarizable Shell (sección 10.iii).

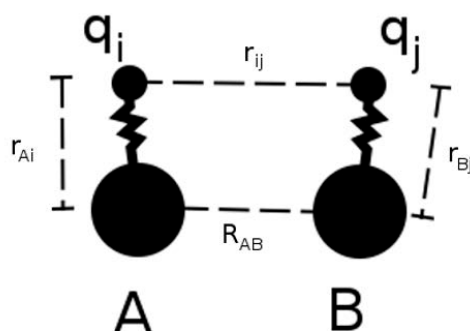


Figura 13. Representación esquemática de los parámetros en el modelo MCHO.

La ecuación del potencial MCHO viene dada por:

$$V_{AB} = \sum_{A < B} \left[\sum_{i \in A} \sum_{j \in B} \left(\Lambda_{AB} e^{-n_1 R_{AB}} + \Theta_{AB} e^{-n_2 R_{AB}} + C \frac{q_i q_j}{r_{ij}} + \frac{k}{r_{ij}^{n_3}} \right) \right] + \sum_A \sum_{i \in A} K_A r_{A_i}^2 \quad (11.1)$$

Donde R_{AB} es la distancia que separa los núcleos de los átomos A y B; r_{ij} es la distancia entre la carga del sitio i del átomo A y la carga del sitio j del átomo B; r_{A_i} es la distancia a la cual está alejada la carga del sitio i de su respectivo núcleo A.

Λ_{AB} , n_1 , Θ_{AB} y n_2 son parámetros que dependen de la forma en cómo los núcleos A y B interactúan. A la parte de la ecuación que contiene estas constantes se le llama término de dispersión - atracción y corresponde a interacciones del tipo Lennard-Jones.

La parte electrostática está claramente definida por las cargas q_i y q_j , así como por la constante C . El término siguiente, con las constantes k y n_3 corresponde a una interacción repulsiva que ayuda a prevenir el colapso de las cargas.

Finalmente, el último término de la ecuación es el de los osciladores armónicos, con una constante de fuerza K_A que depende exclusivamente del tipo átomo al cual este acoplado el oscilador.

Hay que hacer énfasis en que el modelo es tipo átomo – átomo, por lo que algunas interacciones son del tipo intermolecular y otras de tipo intramolecular. Así, el término de dispersión – atracción no existe cuando A y B están en la misma molécula. Además las constantes C , k , n_3 cambian para una misma interacción dependiendo de si los átomos que interactúan estén dentro de la misma molécula o en otra molécula diferente.

Los parámetros de esta ecuación se obtienen ajustándolos de tal manera que el potencial MCHO reproduzca la superficie de potencial de una interacción entre moléculas de interés.

El modelo MCHO es presentado ya que se perfila como un buen candidato para el desarrollo del presente trabajo. Originalmente el modelo MCHO fue desarrollado para el agua por lo que los parámetros para la interacción agua – agua existen y están listados en el apéndice. Dichos parámetros han sido obtenidos de [84] y están parametrizados de tal manera que pueden reproducir simultáneamente la superficie de potencial de la interacción del dímero de agua obtenida por Matsuoka y colaboradores [85], la contribución no aditiva de tres cuerpos de la superficie obtenida por Clementi y colaboradores [86], así como la no aditividad de cuatro cuerpos calculada por Gil Adalid y Ortega Blake [84].

Para el caso de las interacciones agua – hipoxantina también existen datos (los cuales de igual manera están listados en el apéndice) y han sido obtenidos de [48]. Estos parámetros han sido desarrollados de tal manera que reproducen la superficie de potencial generada por las interacciones agua – hipoxantina N1-N7 y agua – hipoxantina N1-N9.

Las ventajas del modelo MCHO original frente a otros modelos de polarización más conocidos como el PTIP4P[87] o el SPC/FQ[65] son (entre las más importantes) [61]:

- Una exactitud más alta en el cálculo de energías.
- Una muy buena reproducción de las propiedades estructurales líquidas del agua, sin que el modelo haya sido ajustado para esto *a priori*.
- Una buena estimación del momento dipolar del agua.

Tales ventajas también son mencionadas para las interacciones agua – agua [84] e hipoxantina – agua [48].

12. Modelado de los sitios electrónicos MCHO: ANES y Single-Update.

El modelo MCHO tiene características predominantes de un modelo polarizable Shell y se ha mencionado (sección 10.iii) que se tiene que lidiar con un grado de libertad más en el desarrollo de una simulación. Esto significa que durante el transcurso de una simulación el espacio fase también contendrá puntos en donde los estados nucleares (posición y

orientación de la molécula) sean los mismos pero los estados electrónicos difieran (posición de las cargas móviles).

Esta separación entre las cargas móviles y los núcleos permite que durante el paso de una configuración a otra de las posiciones nucleares, los sitios electrónicos respondan al nuevo ambiente eléctrico, cambiando sus respectivas posiciones.

Originalmente en el modelo MCHO [84], al efectuarse un movimiento nuclear, inmediatamente se resolvían las ecuaciones de fuerza cero para obtener el nuevo conjunto de posiciones de los sitios electrónicos. Sin embargo, la resolución de las ecuaciones de fuerza cero para todas las moléculas de un sistema es impráctico, así que generalmente lo que se hace es resolver sólo las ecuaciones de fuerza cero para la molécula que ha cambiado de posición. Físicamente esto significa que la nueva molécula se ambienta: permite una respuesta de sus sitios electrónicos al nuevo campo eléctrico al cual se ha movido, mientras las demás moléculas permanecen inalteradas. A este caso se le conoce como *Single-Update*⁴.

Sin embargo en el presente trabajo la validez del método Single-Update se pone a discusión. Esta idea se desprendió originalmente del trabajo, ya que de la observación de los conceptos de polarización, se notó que la respuesta de un sistema a la polarización después de un movimiento nuclear no debería ser exclusiva de la molécula que ha cambiado de posición, sino de todo el entorno molecular que le rodea.

Dentro del contexto del método de Monte Carlo, lo anterior se muestra como una posible discrepancia en la reversibilidad microscópica, hecho que compromete la convergencia de la cadena de Markov, puesto que se usa el enfoque de Metropolis para aceptar los movimientos y que por consiguiente puede invalidar el rigor de la simulación.

Para ver esto, supongamos que tenemos un sistema de dos moléculas (A y B) con coordenadas nucleares \vec{r} y electrónicas \vec{q} ; al inicio, llamado configuración 1 por conveniencia, supongamos que partimos de un sistema se encuentra en equilibrio y se tiene que:

$$\left(\frac{\partial \mathcal{V}(r^{A,B}, q^{A,B})}{\partial q^{A,B}} \right)_{r^A, r^B, q^A, q^B} = 0$$

La energía está minimizada en función de las posiciones electrónicas de A y de B. Si suponemos ahora que movemos a la molécula A aleatoriamente, digamos a la posición r_1^A se genera una nueva posición electrónica de los sitios en A q_1^A aplicando el método single-update. En esta situación, configuración 2, se tiene que:

$$\left(\frac{\partial \mathcal{V}(r^{A,B}, q^{A,B})}{\partial q^A} \right)_{r_1^A, r^B, q_1^A, q^B} = 0 \quad (12.1)$$

⁴ Single-update o actualización individual, ya que solo un objeto (la molécula movida) se actualiza.

$$\left(\frac{\partial \mathcal{V}(r^{A,B}, q^{A,B})}{\partial q^B} \right)_{r_1^A, r_2^B, q_1^A, q_2^B} \neq 0$$

Se ha minimizado la energía de la molécula que se ha movido solamente de acuerdo al método.

Ahora supongamos que la molécula B se desplaza aleatoriamente y las posiciones cambian a r_2^B y q_2^B . En este estado se aplica entonces el single-update a la molécula B y se tiene que en la configuración 3:

$$\begin{aligned} \left(\frac{\partial \mathcal{V}(r^{A,B}, q^{A,B})}{\partial q^A} \right)_{r_1^A, r_2^B, q_1^A, q_2^B} &\neq 0 \\ \left(\frac{\partial \mathcal{V}(r^{A,B}, q^{A,B})}{\partial q^B} \right)_{r_1^A, r_2^B, q_1^A, q_2^B} &= 0 \end{aligned}$$

Ahora, supongamos que se regresa la molécula B a su estado anterior, el cual es el proceso de reversibilidad microscópica; se tiene entonces que, en la configuración 4:

$$\begin{aligned} \left(\frac{\partial \mathcal{V}(r^{A,B}, q^{A,B})}{\partial q^A} \right)_{r_1^A, r_2^B, q_1^A, q_2^B} &\neq 0 \\ \left(\frac{\partial \mathcal{V}(r^{A,B}, q^{A,B})}{\partial q^B} \right)_{r_1^A, r_2^B, q_1^A, q_2^B} &= 0 \end{aligned} \tag{12.2}$$

Y el problema viene cuando se observa que en las ecuaciones 12.1 y 12.2, el término para la molécula B no es el mismo. Durante el proceso de reversibilidad microscópica de la molécula B, en el transcurso de una simulación, las posiciones electrónicas en B *podrían no ser iguales* en la configuración 2 y 4 las cuales esencialmente *deben ser las mismas*.

Se presenta un posible problema en el modelado de los sitios electrónicos por lo que se opta por una alternativa al modelado de los sitios electrónicos. En 1998, Martin, Chen y Siepmann [88], propusieron un método con el que se podía muestrear el espacio fase del colectivo canónico sin necesidad de resolver las ecuaciones de fuerza cero, denominado muestreo adiabático nuclear y electrónico (ANES⁵). El principio básico detrás del método ANES es similar al formalismo del Lagrangiano extendido usado en las simulaciones de dinámica molecular, donde los grados de libertad electrónicos son tratados como grados de libertad clásicos sujetos a un termostato a baja temperatura; de esta manera los grados de libertad electrónicos son optimizados *al vuelo*⁶ durante cada movimiento nuclear.

El fundamento teórico del método ANES es la aproximación adiabática (aproximación de Born-Oppenheimer), la cual nos justifica dividir los grados de libertad electrónicos y

⁵ ANES Adiabatic Nuclear and Electronic Sampling

⁶ Del inglés "on the fly", implica que la optimización es realizada inmediatamente entre cada movimiento nuclear

nucleares. De esta manera puede tomarse una temperatura electrónica T_{el} y se piensa en la densidad electrónica como un conjunto de grados de libertad clásicos, generando una distribución continua de estados electrónicos para cada configuración nuclear.

A partir de estas aseveraciones es necesario entonces incluir explícitamente a los grados de libertad electrónicos en la función de partición del colectivo canónico. Chen demostró que [89]:

$$Q(N, V, T, T_{el}) = Q_{nu}(N, V, T)Q_{el}(N, T_{el})$$

$$Q_{nu}(N, V, T) = \frac{V^N}{\Lambda^{3N} N!} \int \exp(-\beta E_{gs}(s, V)) ds$$

$$Q_{el}(N, T_{el}) = \frac{1}{\Lambda_{el}^{cN}} \int \exp\left(-\frac{\varepsilon(p)}{kT_{el}}\right) dp$$

Donde Q es la función de partición total, y los subíndices indican los grados de libertad nucleares (nu) y electrónicos (el). $T_{el}' = T^{-1} + T_{el}^{-1}$ Tal que $T \gg T_{el}$. Λ es la función de onda térmica de De Broglie a las temperaturas T para los grados de libertad nucleares y T_{el}' para los electrónicos; c es el número de grados de libertad electrónicos por molécula. $E_{gs}(s)$ es la energía basal del sistema que corresponde a las coordenadas nucleares s en el límite de la aproximación de Bohr-Oppenheimer; $\varepsilon(p)$ es la diferencia entre la energía potencial del sistema y E_{gs} y depende sólo del conjunto de coordenadas electrónicas del sistema.

De esta manera, el método ANES permite tratar los movimientos nucleares normales de una simulación y los movimientos electrónicos de manera independiente.

El sentido físico del método ANES involucra al número R_{el} : la cantidad de movimientos de los grados de libertad electrónicos usados para optimizar el ambiente eléctrico entre cada movimiento nuclear. R_{el} se escoge de tal manera que entre cada movimiento nuclear hayan suficientes movimientos electrónicos para que el sistema cambie adecuadamente. Se ha comprobado que en una sucesión pequeña y finita de movimientos nucleares la elección una R_{el} pequeña (del orden de 10^1) es la más efectiva tanto en términos computacionales como en términos de optimización del ambiente eléctrico [89].

Para terminar con la presentación del método ANES es de señalar lo que sucede en los casos extremos: $R_{el} = 0$ y $R_{el} \rightarrow \infty$. En el primer caso las ecuaciones se reducen a un sistema sin muestreo electrónico, mientras que en el segundo caso se ha demostrado que corresponde a una minimización de la energía completa para todas las moléculas a partir de las ecuaciones de fuerza cero [89; 88].

El método ANES resuelve la posible problemática en el método Single-Update y dado que esta derivado a partir del formalismo del método de Monte Carlo en el colectivo canónico, se asegura el criterio de reversibilidad microscópica, el cual es una condición suficiente para asegurar la convergencia de la cadena de Markov.

De hecho, de acuerdo a Siepmann y colaboradores [89], la mayor ventaja del método ANES es su rápida implementación a los métodos de simulación molecular, a diferencia de

los métodos de minimización como el Single-Update que son más complicados de incluir en un programa de simulación molecular. Adicionalmente, conforme a los datos aportados en [89; 88] la ganancia en tiempo computacional escala mejor en el método ANES, en el orden de N^2 , mientras que en los métodos de minimización, el tiempo de computo escala en el orden de N^4 .

En la práctica, los métodos Single-Update y ANES se diferencian como lo presenta el muy sencillo sistema de la Figura 14. Con respecto a la figura, después de un movimiento de rotación de la partícula A, si el tratamiento de los grados de libertad electrónicos es hecho mediante el método Single-Update, sólo este grado de libertad de esta partícula A cambia para responder al nuevo campo eléctrico. Por otro lado, en el método ANES, si R_{el} es 2, dos partículas son elegidas al azar para que sus grados de libertad electrónicos cambien, en este caso B y C. Esto significa que aunque la partícula A se haya movido, no necesariamente implica que sus grados de libertad electrónicos se vayan a mover también.

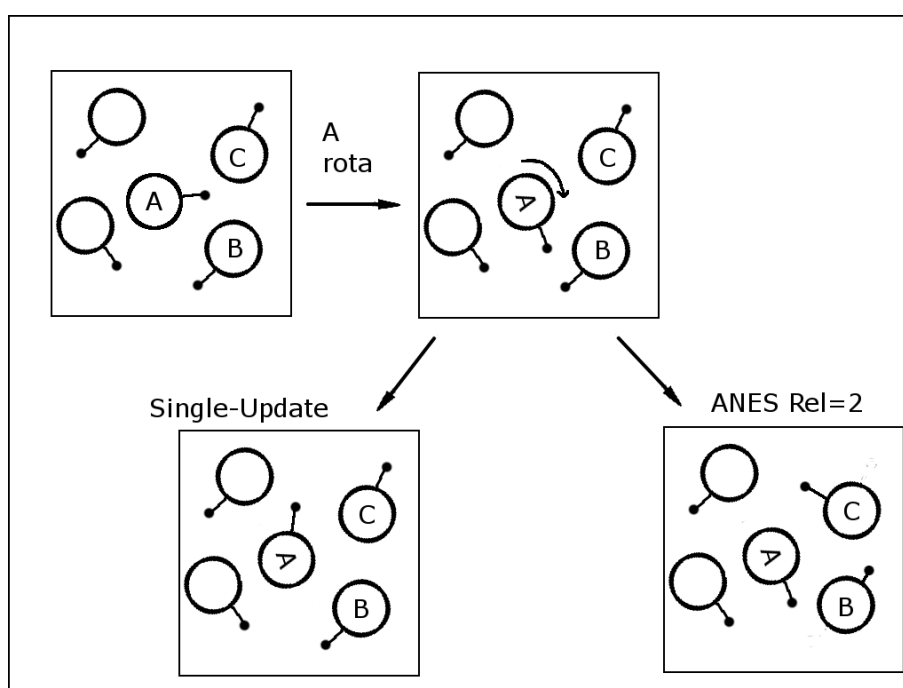


Figura 14. Comparación método Single-Update vs ANES con $R_{el} = 2$, en un sistema de partículas (esferas) con un sitio electrónico (puntos).

13. Cálculo de la energía libre: la integración termodinámica.

Teniendo ya el modelo y el método de simulación molecular, la última pieza fundamental del trabajo, englobada en el marco teórico, es el método para calcular la energía libre.

La energía libre es uno de los conceptos más importantes en fisicoquímica ya que ésta puede describir la tendencia de un sistema de moléculas a asociarse o reaccionar. Lo anterior resulta sumamente atractivo, ya que una metodología capaz de calcular esta cantidad en un sistema de interés tendría una enorme importancia. Desafortunadamente, el

cálculo de la energía libre no puede ser obtenido de los resultados de una simulación molecular de manera inmediata, como el caso de las capacidades caloríficas, por dar un ejemplo [49], ya que para evaluar la energía libre es necesario conocer la totalidad del espacio fase.

Experimentalmente, para la obtención de una energía libre, lo que se hace es calcular derivadas de la energía libre (A) con respecto al volumen (V):

$$\left(\frac{\partial A}{\partial V}\right)_{N,T} = -P$$

O respecto a la temperatura (T):

$$\left(\frac{\partial A/T}{\partial 1/T}\right)_{V,T} = -E$$

Ya que tanto la presión (P) como la energía (E) son propiedades mecánicas del sistema pueden ser medidas a partir de una simulación molecular. Entonces para evaluar la energía libre de un sistema a determinadas densidad y temperatura, se debería encontrar una trayectoria reversible a través del plano V-T que une al sistema en consideración con alguno de energía libre conocida y entonces integrar las ecuaciones anteriores. No obstante el problema recae en que existen muy pocos sistemas con energías libres conocidas.

Bajo este punto de vista, una simulación molecular tiene una gran ventaja ya que no se tiene la restricción de avanzar sólo por una *trayectoria* de integración termodinámica física. Con los datos que provee una simulación molecular se pueden usar parámetros relacionados a la energía potencial como si fueran termodinámicos e integrar a través de una trayectoria entre esos parámetros.

El concepto anterior es la base del método de *integración termodinámica* [90; 91]. Para mostrar cómo funciona el método de integración termodinámica propuesto, supongamos que existe un sistema de N partículas con energía potencial \mathcal{V} ; asumimos que \mathcal{V} depende linealmente de un parámetro de acoplamiento λ tal que $\lambda = 0$ corresponde a la energía potencial del sistema A, mientras que cuando $\lambda = 1$, se obtiene la energía potencial del sistema B:

$$\mathcal{V}(\lambda) = (1 - \lambda)\mathcal{V}_A + \lambda\mathcal{V}_B$$

A este potencial (λ) se le llama *potencial compuesto*.

Para un sistema a N, V y T constantes, la energía libre de Helmholtz A es:

$$A = -kT \ln Q$$

Donde Q es la función de partición del colectivo NVT , k la constante de Boltzmann y T la temperatura. De las relaciones obtenidas de la mecánica estadística, utilizando nuestra relación para la energía potencial, se tiene que:

$$Q(N, V, T, \lambda) = \frac{1}{\Lambda^{3N} N!} \int \exp(-\beta \mathcal{V}(\lambda)) dr$$

Entonces se puede demostrar que [91]:

$$\left(\frac{\partial A(\lambda)}{\partial \lambda} \right)_{NVT} = \left\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \quad (13.1)$$

Donde $\left\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \right\rangle_{\lambda}$ indica el promedio de la energía potencial en el colectivo. Es así que la diferencia en energía libre entre los sistemas A y B se obtiene integrando la ecuación anterior:

$$A_{\lambda=1} - A_{\lambda=0} = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (13.2)$$

Este resultado es de gran importancia, pues permite calcular diferencias de energía libre en términos del promedio en un colectivo, que a diferencia de la energía libre, puede ser calculada directamente de la simulación.

Hay que hacer notar también la implicación más sobresaliente de la relación mostrada en la ecuación 13.2: ésta relación permite obtener los datos de energía libre, comparables a los experimentales, a partir de las simulaciones moleculares, métodos usados desde computadoras, mostrando la sinergia existente entre el modelado y la experimentación.

La evaluación de la integral 13.2 puede ser hecha mediante cuadratura Gaussiana:

$$\int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda = \sum_{i=1}^n w_i \left\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i}$$

El cálculo de la diferencia de energías libres ahora requiere la selección de diferentes puntos λ_i con su respectiva w_i . Tales valores pueden encontrarse en diferentes manuales de matemáticas [92; 93].

Nótese que, entre mayor sea el número de puntos elegidos para evaluar la suma, mayor es la precisión que se obtiene del resultado de la integral. De hecho, una cuadratura Gaussiana es exacta en el intervalo evaluado para un polinomio de grado $2n - 1$, con $n > 0$, donde n es el número de puntos elegidos [94]. No obstante, cada punto λ_i representa una simulación molecular completa y diferente por lo que hay que tomar en cuenta el tiempo de cómputo a la hora de seleccionar el número de puntos a evaluar.

La implementación de la integración termodinámica es muy sencilla ya que $\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \rangle_{\lambda_i}$ es fácil de calcular durante en una simulación:

$$\left\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \right\rangle = \left\langle \frac{\partial}{\partial \lambda} (1 - \lambda) \mathcal{V}_A + \lambda \mathcal{V}_B \right\rangle = \langle \mathcal{V}_B - \mathcal{V}_A \rangle \quad (13.3)$$

La simulación se realiza de tal manera que ésta evoluciona de acuerdo al potencial compuesto:

$$\mathcal{V}(\lambda) = (1 - \lambda) \mathcal{V}_A + \lambda \mathcal{V}_B$$

Donde λ ha sido elegido de la cuadratura Gaussiana. Y entonces, se calcula la cantidad indicada en la ecuación 13.3 en el transcurso de la simulación. Así, el cálculo de la energía libre consiste en sumar el producto del promedio de 13.3 con el peso w_i de la correspondiente λ_i usada en todos los puntos elegidos de la cuadratura Gaussiana.

Existen otros métodos para el cálculo de la energía libre, los más conocidos a parte de la integración termodinámica son: la teoría de perturbación estadística y el método de crecimiento lento⁷.

En el método de perturbación estadística [95], la energía libre se calcula también a partir de un potencial compuesto mediante:

$$\Delta G = \sum_{\lambda=0}^1 -RT \text{Ln} \langle e^{-\Delta \mathcal{V}_c / RT} \rangle_{\lambda}$$

Donde $\Delta \mathcal{V}_c = \mathcal{V}_{\lambda+d\lambda} - \mathcal{V}_{\lambda}$. Lo que se hace es dividir la suma en pequeños bloques relacionados con λ y se calcula la energía potencial en los respectivos estados.

En el método de crecimiento lento [96], lo que se hace es evaluar la integral 13.2 como si fuera una suma:

$$\Delta A = \sum_{\lambda=0}^1 \langle \mathcal{V}(\lambda + 1) - \mathcal{V}(\lambda) \rangle$$

La desventaja de la integración termodinámica es que el cálculo de la energía libre requiere de resolver numéricamente una integral, mientras que en los otros casos no. Sin embargo, la ventaja que tiene la integración termodinámica sobre los otros métodos es que si el muestreo, la suma sobre el parámetro de acoplamiento λ , es insuficiente, tanto el método de perturbación estadística como el método de crecimiento lento tendrían que repetir el procedimiento con un nuevo conjunto de parámetros λ , mientras que en la integración termodinámica es suficiente con agregar nuevos puntos.

⁷ Del inglés Slow growth method.

De entre los tres métodos, se ha encontrado que la integración termodinámica es el método que se implementa más fácilmente en cuanto al cálculo de la energía libre [97; 98]. No obstante, los tres métodos son igualmente viables si se les provee de suficientes datos para evaluar la energía libre [99]. De hecho, las variables más críticas para que el cálculo de la energía libre sea satisfactorio son: la elección de un modelo correcto para las interacciones del sistema y el muestreo del sistema (suma sobre λ) [44].

Las aplicaciones de la integración termodinámica para el cálculo de energías libres son vastas e incluyen, por ejemplo, el estudio de la solvatación en diferentes disolventes [100], el tratamiento de confórmeros [101], las reacciones de tautomerismo [102], las reacciones redox [103], las reacciones ácido base [104], la asociación molecular [105] entre otros. De esta manera, la integración termodinámica, se considera como uno de los métodos más importantes para el cálculo de la energía libre mediante simulaciones moleculares.

Metodología.

14. Organización.

El objetivo del presente trabajo es calcular la energía libre de reacción del tautomerismo de la hipoxantina N1-N9/N1-N7 en solución acuosa ($\Delta G_{H19 \leftrightarrow H17}$), así como su constante de equilibrio (K_{eq}).

La termodinámica nos da la relación entre $\Delta G_{hp19-hp17}$ y K_{eq} por lo que el problema principal recae en calcular $\Delta G_{hp19-hp17}$.

Se pretenden usar los métodos de simulación molecular de Metropolis Monte Carlo para modelar la reacción en un medio que incluya al disolvente explícitamente.

Se elige el método de integración termodinámica para el cálculo de $\Delta G_{H19 \leftrightarrow H17}$ ya que éste se deriva fácilmente de los métodos de simulación molecular.

El modelo seleccionado para representar a los tautómeros de la hipoxantina y al agua es el MCHO ya que introduce el manejo de la polarización en el sistema, el cual de acuerdo a la hipótesis y lo visto en el marco teórico es importante incluir.

El modelo MCHO introduce grados de libertad extra a la simulación que son necesarios manejar. El método original Single-Update para manejar estos grados de libertad parece presentar una posible discrepancia al usarse junto con el método de Metropolis Monte Carlo por lo que se elige un método alternativo, el método ANES.

Los puntos de partida son un estudio en fase gas de los tautómeros y la simulación del agua pura. En el primer caso se usa DFT para obtener las geometrías de los tautómeros y sus momentos dipolares así como las poblaciones relativas en fase gas. En la simulación de agua pura se prepara el camino para tener una configuración de disolvente para los tautómeros así como para probar y comparar los métodos ANES y Single-Update.

15. Labores realizadas.

Se hicieron cálculos cuánticos en fase gas para cada uno de los dos tautómeros más estables de la hipoxantina: N1-N9 y N1-N7. Para esto se utilizó la metodología propuesta por Ramaekers y colaboradores [19]. Utilizando la colección de programas GAUSSIAN 03[106], se realizaron optimizaciones de geometría de los dos tautómeros utilizando teoría de los funcionales de la densidad (DFT) con el funcional B3LYP y la base 6-311**G++. El funcional B3LYP es generalmente aceptado para realizar cálculos de este tipo de moléculas [19], mientras que para la base usada Ramaekers y colaboradores, sugieren el uso de bases mayores a 6-31**G+, para obtener resultados más exactos. Lo anterior es importante ya que el resultado final depende de los valores obtenidos de los cálculos en fase gas, por lo que añadir más funciones a la base resultaría benéfico para la exactitud de los resultados.

Obtenidas las geometrías de cada tautómero, se procedió a realizar un cálculo de frecuencias para obtener otros datos de interés: energías libres y momentos dipolares.

La simulación molecular de agua MCHO fue efectuada con el método de Monte Carlo en el colectivo canónico junto con los métodos ANES y Single-Update, utilizando 512 moléculas de agua, colocadas en una celda primitiva cúbica de 24.866 \AA de dimensión en la arista, que corresponde a la densidad del agua de 0.997 g cm^{-3} . Las posiciones iniciales de los centros de masa y orientación de cada una de las moléculas de agua se generaron completamente al azar obedeciendo la geometría propuesta del modelo MCHO[84] (listada también en la pág. 67, en el apéndice). Las posiciones iniciales de los grados de libertad electrónicos de cada molécula de agua corresponden al estado de mínima energía de la molécula aislada. Las temperaturas utilizadas fueron de $T = 298 \text{ K}$ y $T_{el} = 5 \text{ K}$. Así, se realizaron 200×10^6 configuraciones nucleares y en cada una se realizaba un ciclo ANES de $R_{el} = 20$; las últimas 100 millones de configuraciones nucleares con sus respectivas configuraciones electrónicas fueron utilizadas para obtener las propiedades promedio. Se verificó, también, que la parte promediada de la simulación estuviera en equilibrio mediante el método de análisis por bloques [107]. Los parámetros que se utilizaron para las sumas de Ewald fueron de $k_{max} = 400$ y $\alpha = 0.2 \text{ \AA}^{-1}$. Por último se utilizaron las relaciones de aceptación de 0.5, 0.5 y 0.7 para la aceptación de los movimientos nucleares, de rotación y electrónicos.

Los parámetros para el modelo MCHO correspondientes al agua están listados en el apéndice y fueron tomados de [84].

Se realizó la simulación molecular de cada tautómero en solución. Utilizando la geometría de los tautómeros obtenidas del estudio cuántico en fase gas. Cada una de las moléculas de hipoxantina se introdujeron a una celda cúbica con la configuración final de moléculas de agua obtenida de la simulación de agua pura realizada (MCHO ANES), generando así dos nuevas configuraciones: una celda con la HP19 y otra celda con la HP17, cada una con la misma configuración de moléculas de disolvente alrededor de ella. Cada molécula de tautómero se fijó al centro del cubo y las moléculas traslapadas o muy cercanas ($< 1 \text{ \AA}$) a alguno de los tautómeros fueron removidas. La densidad se reajustó modificando el tamaño del cubo. Se cuidó que cada molécula removida de una celda fuera igualmente removida de la otra, dejando así una configuración y número de moléculas de disolvente constante en ambos cubos. Esto se hace debido a que, al utilizar la integración termodinámica el sistema tiene al hidrógeno tanto en N7 como en N9 al mismo tiempo y como se quiere empezar de una misma configuración de disolvente en todas las simulaciones no se desean traslapes moleculares. Así cada molécula evoluciona, en una simulación molecular.

Los parámetros de la simulación corresponden a los del agua realizada anteriormente. Adicionalmente, los parámetros del modelo MCHO correspondientes a la hipoxantina y a las interacciones hipoxantina — agua se muestran también en el apéndice (página 68) y fueron obtenidos de [48].

Nueve simulaciones más se realizaron para poder hacer la integración termodinámica. En estas se utilizó la configuración de moléculas de agua inicial construida para las simulaciones de los tautómeros anteriores. También se utilizaron los mismos parámetros

para cada simulación por el método de Monte Carlo. La diferencia entre las simulaciones para la integración termodinámica y la de los tautómeros normales es el uso de una molécula ficticia y del potencial compuesto para el cálculo de las interacciones entre las moléculas.

Todos los cálculos fueron realizados en la súper computadora KanBalam (HP CP 4000) de la UNAM con apoyo de la Dirección General de Servicios de Cómputo Académico.

16. Detalles computacionales.

Todas las simulaciones fueron escritas en Fortran bajo el estándar 90/95. Dado que gran parte del programa para efectuar las simulaciones se desarrollo como parte del mismo trabajo de tesis, escribiendo partes originales y adaptando rutinas existentes, a continuación se explican los tecnicismos computacionales de las simulaciones escritas. Esta exposición se enfoca principalmente a la descripción de las variables mencionadas en la sección anterior. Las rutinas mencionadas y la adaptación de estas se basó en lo encontrado en los libros de Allen y Frenkel [49; 50] y para el método ANES del trabajo de B. Chen y J. I. Siepmann [89].

i. Ciclos y configuraciones

Los métodos de Monte Carlo se basan en lo generación de diferentes **configuraciones** de manera aleatoria de acuerdo a una distribución de probabilidad. Donde una configuración corresponde a un punto en el espacio fase.

De acuerdo a la ecuación 8.10, la probabilidad con la que el cambio de una configuración A a una B es aceptado depende enteramente de la diferencia de energía potencial entre estas configuraciones: $\exp(-\beta\Delta\mathcal{V})$

Hay tres maneras en la que el sistema puede evolucionar de un estado A a uno B :

- Traslación de una molécula.
- Rotación de una molécula.
- Traslación de cargas.

Los primeros dos movimientos son forzosamente elegidos a lo largo de toda la simulación de manera aleatoria para generar las configuraciones aleatoriamente. La traslación de cargas se hace mediante un *ciclo ANES* o un cálculo Single-Update, dependiendo de la que se use, pero siempre es hecha después de que cada movimiento molecular es realizado.

Por conveniencia, un número finito de configuraciones se agrupa en un **ciclo** de Monte Carlo. De esta manera en cada ciclo se pueden hacer ajustes a las variables que controlan el flujo del programa y acomodar los datos de la simulación molecular en bloques que faciliten su análisis. En el presente trabajo un ciclo corresponde a 512 configuraciones, elegidas de acuerdo al número de moléculas que se trabajan en la simulación del agua.

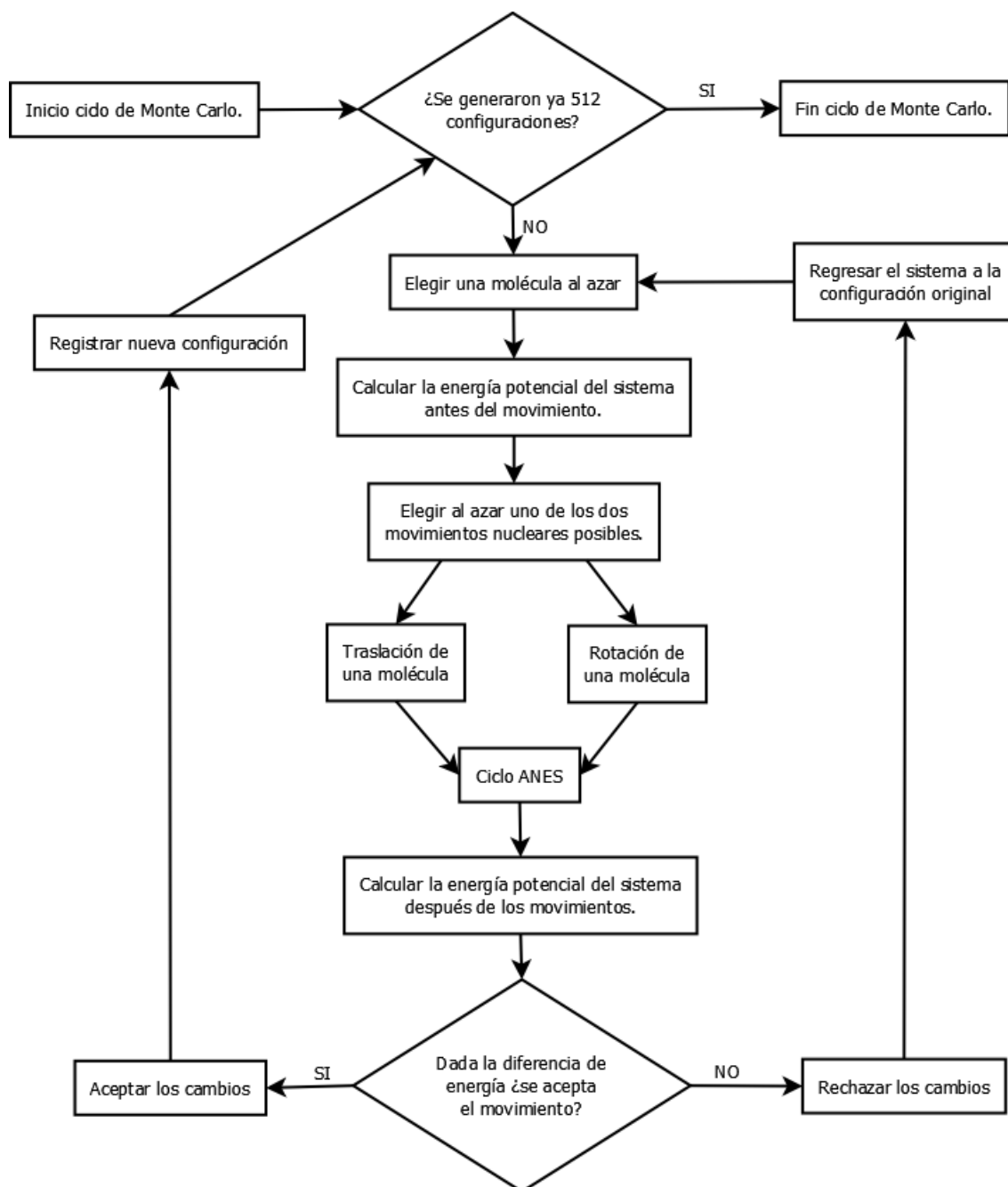


Figura 15. Diagrama de flujo de un ciclo de Monte Carlo ANES.

Todo lo que se ha comentado hasta este punto se sintetiza en la Figura 15 para el método de Monte Carlo ANES.

ii. Movimientos nucleares: traslaciones.

Para realizar la traslación de una molécula simplemente se genera un desplazamiento al azar. Una variable, que llamamos Δr_{max} controla la máxima distancia que una molécula

puede moverse a través de sus coordenadas x, y, z . Tal variable existe para controlar el muestreo del espacio fase durante las traslaciones: de generarse desplazamientos muy grandes la posibilidad de que la molécula se acerque demasiado a otra es mayor y por lo tanto la cantidad de movimientos rechazados sería mayor; por el contrario si los desplazamientos son muy pequeños el sistema no cambiaría mucho entre configuraciones y muchos serían los movimientos aceptados. Para mantener un equilibrio entre el número de movimientos rechazados y aceptados, se utilizan las **relaciones de aceptación**.

Las relaciones de aceptación son una necesidad en todo movimiento presente en una simulación molecular. El objetivo de estas cantidades es controlar que todos los movimientos entre configuraciones se efectúen de un modo controlado, de tal manera que la evolución del sistema no se estanque.

Las relaciones de aceptación se fijan al inicio de la simulación como un porcentaje de movimientos aceptados de entre todos los que se efectuaron del mismo tipo en un ciclo de Monte Carlo.

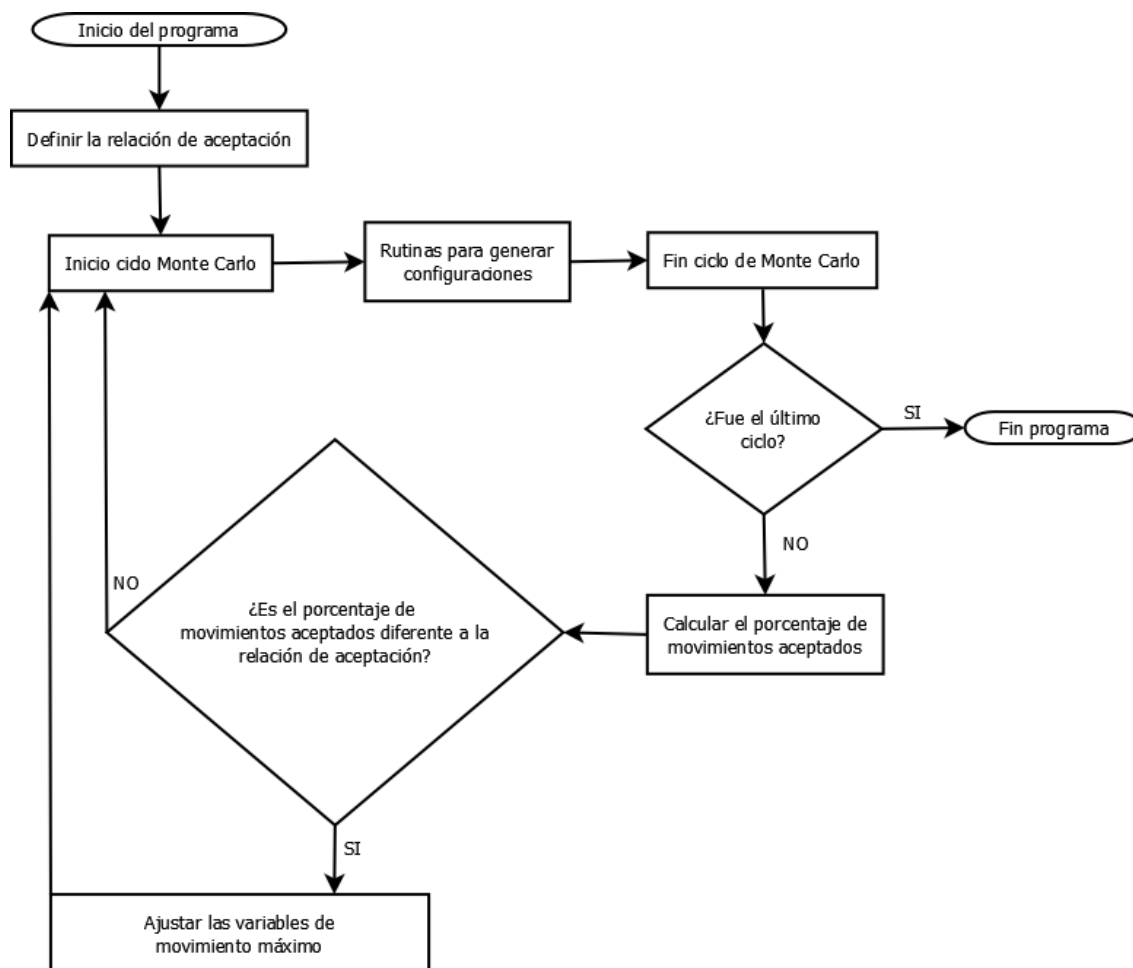


Figura 16. Las relaciones de aceptación en el ciclo de Monte Carlo.

Para el caso de la traslación, durante el transcurso de la simulación molecular cada ciclo de Monte Carlo se cuantifica el porcentaje de traslaciones aceptadas. Si este porcentaje se desvía de la *relación de aceptación de traslación*, la variable Δr_{max} se ajusta de tal manera que el porcentaje de traslaciones aceptadas en el siguiente ciclo de Monte Carlo se acerque nuevamente a la relación de aceptación.

Así si una molécula se encuentra en la posición $\vec{r}_1 = (x_1, y_1, z_1)$ la traslación se generará creando una nueva posición $\vec{r}_2 = (x_2, y_2, z_2)$ tal que:

$$\begin{aligned}x_2 &= x_1 + (2 \times \text{random}[0,1] - 1) \times \Delta r_{max} \\y_2 &= y_1 + (2 \times \text{random}[0,1] - 1) \times \Delta r_{max} \\z_2 &= z_1 + (2 \times \text{random}[0,1] - 1) \times \Delta r_{max}\end{aligned}$$

Es decir, la nueva posición de cada coordenada estará entre $[-\Delta r_{max}, \Delta r_{max}]$.

iii. Movimientos nucleares: rotaciones.

Cuando se efectúa una rotación el tratamiento no cambia mucho. Existe una relación de aceptación de rotación que controla que el giro de una molécula sea controlado mediante la variable $\Delta \alpha_{max}$.

Debido a que la implementación computacional más práctica y eficiente de la rotación de un cuerpo son los cuaterniones [108], $\Delta \alpha_{max}$ se agrega como se expone enseguida.

Sea \hat{r} un cuaternión, definido como

$$\hat{r} = r_0 + r_x i + r_y j + r_z k$$

Donde el conjunto base (i, j, k) son números imaginarios. El conjugado de este cuaternión es definido como:

$$\hat{r}^* = r_0 - r_x i - r_y j - r_z k$$

Si se cumple que

$$r_0^2 + r_x^2 + r_y^2 + r_z^2 = 1$$

Se dice que el cuaternión es unitario.

Ahora, si \vec{s} es un vector con componentes (s_x, s_y, s_z) , este puede ser representado como un cuaternión \hat{s} con componentes (s_0, s_x, s_y, s_z) . De esta manera una rotación es llevada a cabo usando al cuaternión unitario \hat{r} mediante la operación:

$$\hat{s}_2 = \hat{r} \hat{s}_1 \hat{r}^*$$

Entonces la rotación se vuelve un problema de sumas, multiplicaciones y generación de cuaterniones unitarios.

Para generar un cuaternión unitario, si se quieren introducir ángulos, es suficiente con hacer al componente real $\cos \alpha$ y a los componentes imaginarios una suma cuyo resultado sea $\sin \alpha$, pues así la suma de los cuadrados de los componentes sería 1.

Así, el ángulo de rotación es controlado mediante la relación de aceptación de rotación que manipula la variable $\Delta\alpha_{max}$ de esta manera:

$$\alpha = (2 \times \text{random}[0,1] - 1)\Delta\alpha_{max}$$

Y el vector de orientación de la molécula rotará entre el ángulo $[-\Delta\alpha_{max}, \Delta\alpha_{max}]$.

iv. Movimientos electrónicos: el ciclo ANES.

Después de cada movimiento nuclear es necesario que se realicen movimientos de los grados de libertad electrónicos. Esto se logra mediante los ciclos ANES.

Un ciclo ANES corresponde a un número R_{el} de *intentos de traslación* de las cargas móviles del modelo MCHO.

Debido a la aproximación adiabática, las traslaciones electrónicas difieren de las nucleares, ya que las cargas electrónicas, a una temperatura T_{el} , están completamente separadas de las nucleares, que están a una temperatura T . La consecuencia principal es que la probabilidad de aceptación de un movimiento electrónico es diferente e independiente a la de un movimiento nuclear. Así, cuando se ha efectuado un movimiento nuclear, se tiene que elegir una molécula al azar y de ésta una carga al azar cuya posición se intentara cambiar con cierta probabilidad.

La posición de una carga se cambia de la misma manera que una molécula se mueve en una traslación. Se define una variable Δs_{max} que cambia de acuerdo a la relación de aceptación de traslación electrónica para que las coordenadas cambien entre $[-\Delta s_{max}, \Delta s_{max}]$:

$$\begin{aligned}x_2 &= x_1 + (2 \times \text{random}[0,1] - 1) \times \Delta s_{max} \\y_2 &= y_1 + (2 \times \text{random}[0,1] - 1) \times \Delta s_{max} \\z_2 &= z_1 + (2 \times \text{random}[0,1] - 1) \times \Delta s_{max}\end{aligned}$$

Y entonces el nuevo vector de posición se acepta o rechaza de acuerdo a $\exp\left(-\frac{\Delta V}{kT_{el}}\right)$.

El ciclo ANES se ilustra en la Figura 17.

v. Listas de Verlet.

Las listas de Verlet se implementaron en el trabajo para aumentar la velocidad de cómputo. Las primeras simulaciones que no contaban con este método eran visiblemente más lentas que las que ocupaban listas de Verlet, aunque el aumento de velocidad no se cuantificó rigurosamente, la implementación es sencilla y por eso se decidió usar el método.

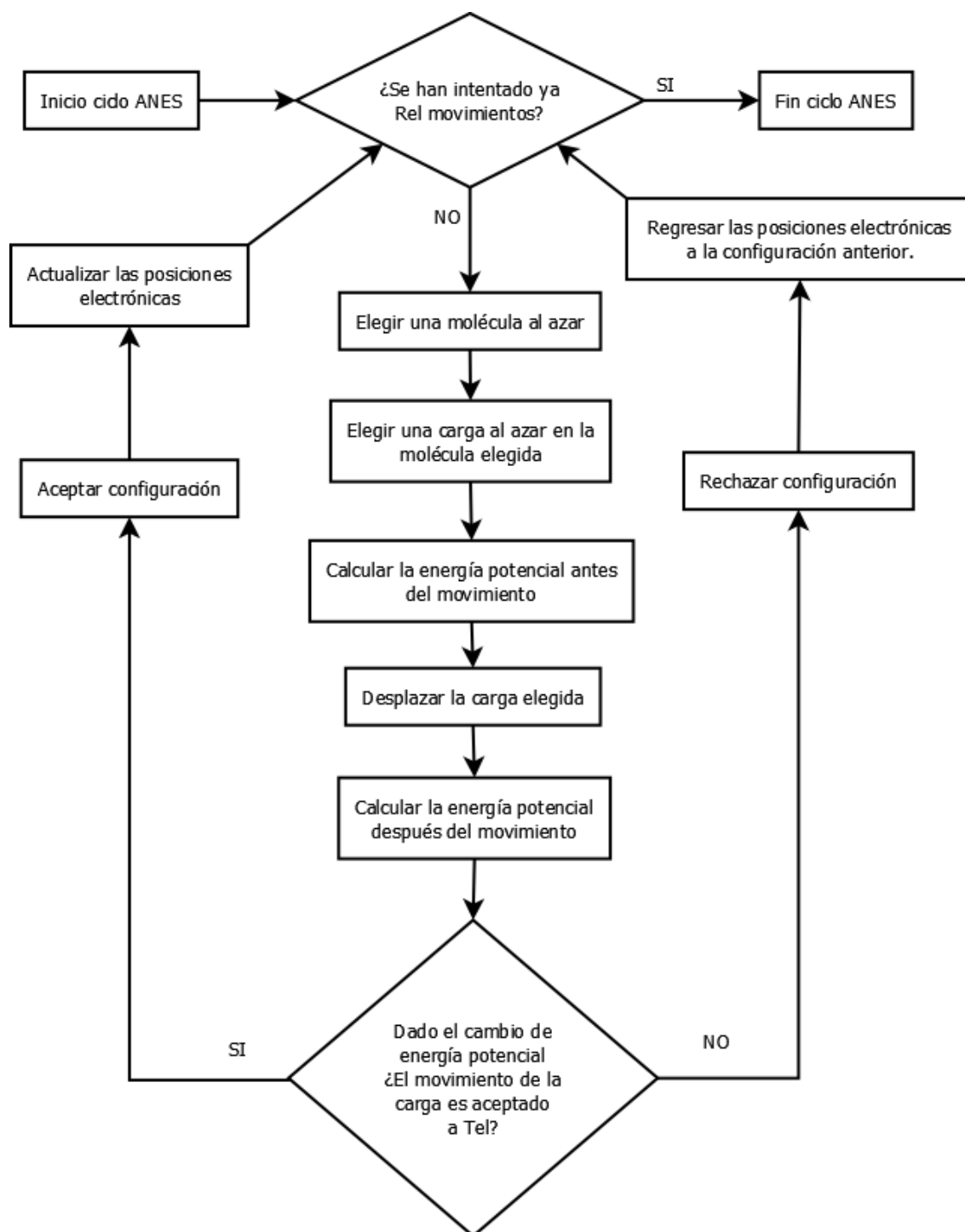


Figura 17. Ciclo ANES.

Uno de los cuellos de botella en una simulación molecular es el recurrente cálculo de la energía potencial en el sistema. Si suponemos que el cálculo es efectuado por pares en N partículas, hay que tomar en cuenta $N(N-1)/2$ interacciones para el cálculo de la energía potencial. Esto implica que sin usar ninguna técnica de simulación el tiempo necesario para el cálculo de la energía se escala en el orden de N^2 .

Verlet [109] desarrolló una técnica, comúnmente llamada *listas de Verlet*, que previene el gasto de tiempo que involucra el recurrente cálculo de la energía potencial tomando sólo en cuenta las moléculas contenidas dentro del volumen generado por un radio de corte. Esta técnica es particularmente útil para el tratamiento de las contribuciones a la energía potencial como las de atracción – repulsión, término incluido en el modelo MCHO.

La idea consiste en generar una lista de moléculas contenida en una esfera Π_v de radio r_v alrededor de la molécula i . Entonces las interacciones sólo son contabilizadas entre las moléculas dentro de la esfera Π_v .

Una nueva esfera Π_c de radio r_c tal que $r_v > r_c$ es generada. Es evidente que esta nueva esfera Π_c está contenida en la esfera Π_v . De esta manera, cada que la molécula i se mueve si alguna parte de la esfera Π_c queda fuera de la esfera Π_v la lista de moléculas alrededor de la molécula i se tiene que actualizar. La Figura 18 ilustra el concepto de las esferas.

Por lo general la magnitud de los radios se elige dependiendo del tamaño que se quiera de la lista. De hecho si la lista contiene menos de 100 interacciones no existe ganancia de tiempo computacional [50].

La ganancia en tiempo computacional viene dada por el hecho de que no es necesario calcular todas las interacciones entre la molécula i y las contenidas en el cubo, sino sólo con aquellas presentes en la lista.

Para el trabajo fue suficiente agregar todas las interacciones dadas en esferas de radio $r_v = 0.48l$ y $r_c = 0.45l$, donde l es el tamaño del lado de la celda cúbica, para observar ganancias considerables en el tiempo de cómputo.

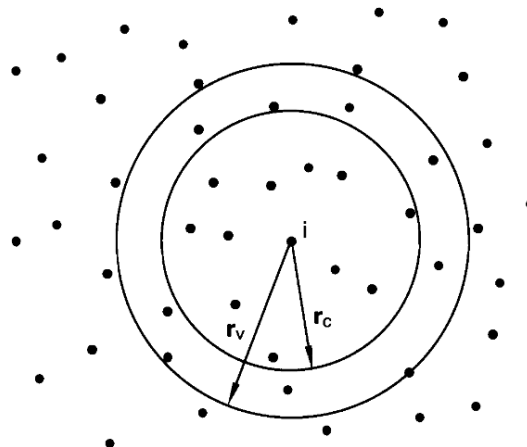


Figura 18. Definición de los radios de las esferas para generar las listas de Verlet.

vi. Sumas de Ewald.

Las sumas de Ewald [110; 111; 112] son un concepto muy técnico relacionados con el cálculo de las interacciones de largo alcance, como las electrostáticas contenidas en el modelo MCHO. El planteamiento teórico de la técnica extenso y para fines prácticos, fuera del alcance del trabajo, se resume en unos comentarios acerca de la rutina implementada en el trabajo (adaptada de [49]).

Las sumas de Ewald consisten en transformar la suma de las interacciones electrostáticas en una suma de interacciones en un espacio recíproco y un espacio real de forma a garantizar la convergencia de la suma. El espacio recíproco está formado precisamente por la celda unitaria, con las moléculas presentes, replicada (en teoría) infinitamente en el espacio. La transformación de las sumas se hace mediante una transformación de Fourier en el espacio recíproco, resultando en una suma rápidamente convergente, a diferencia de original.

Los dos parámetros más importantes de la sumatoria son k_{max} y α . En las sumas de Ewald k_{max} tiene como función contabilizar el número de celdas replicadas que se suman en el espacio recíproco. Dado que la suma de Ewald converge rápidamente, no es necesario utilizar una cantidad de celdas excesiva. Por otro lado, α es una cantidad que describe cuán rápido decrece la contribución de los términos en la sumatoria, entre más grande sea esta cantidad más interacciones se toman en cuenta dentro de una celda. La elección de k_{max} y α es de acuerdo a la exactitud que quiera alcanzarse en el cálculo de las interacciones electrostáticas y el tiempo que quiera invertirse en ello.

Las relaciones que involucran estos detalles son [50]:

$$k_{max} = \frac{4\pi}{3} n_c^3$$

$$\alpha = \left(\frac{r_t \pi^3 N}{L^6} \right)^{\frac{1}{6}}$$

Donde n_c es el número de veces que se replica la celda en el espacio recíproco sobre un eje, r_t es la relación entre el tiempo que se desea evaluar el espacio recíproco y el espacio real. Además, N es el número de moléculas y L es lado del cubo. Con $N = 512$ y $L \approx 25 \text{ \AA}$, si elegimos $n_c = 5$ y $r_t = 1$, es decir que la celda en el espacio recíproco se replica 5 veces sobre cada eje cartesiano y que el tiempo que tomemos en el cálculo de las interacciones en el espacio recíproco y real sea el mismo, nos queda $k_{max} = 520$ y $\alpha = 0.2 \text{ \AA}^{-1}$. Estos valores teóricos pueden optimizarse con simulaciones cortas para obtener los mejores resultados. De diversas simulaciones propias realizadas en el pasado en sistemas acuosos a partir de los datos en [89; 88], se ha encontrado un buen balance entre exactitud y tiempo de cálculo con los parámetros $k_{max} = 400$ y $\alpha = 0.2 \text{ \AA}^{-1}$, por lo que estos son los usados en el trabajo.

vii. Análisis por bloques.

Flyvbjerg y Petersen [113] desarrollaron el método de análisis promedio por bloques. Uno de sus posibles usos es saber si una simulación molecular es lo suficientemente larga como para empezar a promediar las cantidades de interés, es decir si el sistema está equilibrado.

Supongamos que el conjunto de datos A_1, A_2, \dots, A_L son L muestras consecutivas de una cantidad fluctuante A que queremos promediar en el colectivo, si asumimos que el sistema está equilibrado, la cantidad promedio de A se obtiene como:

$$\langle A \rangle \approx \bar{A} \equiv \frac{1}{L} \sum_{i=1}^L A_i$$

Y su varianza como:

$$\langle A^2 \rangle - \langle A \rangle^2 \approx \frac{1}{L} \sum_{i=1}^L [A_i - \bar{A}]^2$$

Si el sistema está equilibrado, todas las L muestras de la cantidad A no estarían correlacionadas y estas formulas serían validas. Sin embargo en una simulación, las muestras obtenidas están correlacionadas hasta cierto punto y hay que tomar en cuenta este detalle para saber cuándo los datos no están correlacionados y poder promediar.

De acuerdo a Flyvbjerg y Petersen, si transformamos el conjunto de datos A_1, A_2, \dots, A_L a un nuevo conjunto A'_1, A'_2, \dots, A'_L que es la mitad de tamaño que el original, tal que:

$$A'_i = \frac{1}{2} (A_{2i-1} + A_{2i})$$

Con

$$L' = \frac{1}{2} L$$

La varianza estaría dada por:

$$\langle A'^2 \rangle - \langle A' \rangle^2 = \frac{1}{L'} \sum_{i=1}^{L'} A'^2_i - \bar{A}'^2$$

Y el promedio \bar{A}' sería el mismo que el del conjunto de datos original. Nótese que lo único que se ha hecho es dividir los datos en bloques con pares de datos. Ahora, al seguir con este procedimiento de hacer bloques con los datos, si la simulación es lo suficientemente larga, las cantidades A'_i empezaran a perder correlación y entonces

$$\frac{\langle A'^2 \rangle - \langle A' \rangle^2}{L' - 1} \approx \text{constante}$$

Si se hace una gráfica de $\langle A'^2 \rangle - \langle A' \rangle^2$ vs $L' - 1$, cuando se empiece a observar una curva constante, se puede decir que la simulación es suficientemente larga para empezar a obtener promedios.

Este procedimiento se realizó a todas las simulaciones para verificar que los datos promediados estuvieran en una zona equilibrada de la simulación.

17. Tratamiento de los resultados obtenidos.

Los resultados obtenidos en fase gas son la geometría de los tautómeros, las diferencias energéticas y los momentos dipolares en fase gas. Las geometrías de los tautómeros sirven como datos entrantes a las simulaciones realizadas de los tautómeros en solución acuosa.

De las simulaciones se obtienen principalmente los datos energéticos y estructurales. En este trabajo, el primero viene dado en la forma de la energía configuracional y el segundo en forma de funciones de distribución radial.

Es importante señalar que a los resultados de energía se les efectúa un análisis por bloques para asegurar que los resultados obtenidos son de la parte de la simulación ya estabilizada. Para ello, el valor de la energía configuracional se guarda en archivos por cada ciclo de Monte Carlo y se acomoda en bloques para hacer las graficas de la varianza contra el número de bloques. La presencia de una meseta indica que la simulación es lo suficientemente larga. Dado que cumplir este requisito es absolutamente necesario en todas las simulaciones, no se presentan como resultados, pero se muestran las graficas en el apéndice.

Sobre la energía configuracional: la energía configuracional es la energía potencial promedio que se obtiene durante el transcurso de una simulación. De hecho, la entalpia de vaporización está íntimamente relacionada con el valor de la energía configuracional y está expresada de la siguiente manera:

$$\Delta H_{vap} = (\mathcal{V}_g^{intra} - \mathcal{V}_l^{intra}) + (\mathcal{V}_g^{inter} - \mathcal{V}_l^{inter}) + P(V_g - V_l)$$

Si suponemos que la fase vapor es un gas ideal y se tiene un modelo que no presenta movimientos de torsión en sus enlaces, la relación se simplifica ya que $(\mathcal{V}_g^{intra} - \mathcal{V}_l^{intra})$ se vuelve casi cero en el caso del modelo MCHO (los núcleos son rígidos), así como \mathcal{V}_g^{inter} . También $P(V_g - V_l) \approx RT$, quedando:

$$\Delta H_{vap} = -\mathcal{V}_l^{inter} + RT$$

Donde \mathcal{V}_l^{inter} es la energía que obtenemos de la simulación. De esta manera es posible relacionar una cantidad que puede ser medida experimentalmente con una que es directamente obtenida de cálculos numéricos.

Sobre las funciones de distribución radial: las funciones de distribución radial describe la manera en que la densidad de átomos varía como función de la distancia de un punto en particular. Dichas funciones se obtienen de las simulaciones y pueden ser comparadas inmediatamente con las experimentales. Un solapamiento mayor entre las curvas

experimentales y las obtenidas de la simulación indicara una mejor descripción de la estructura por parte del modelo.

Para el caso del agua, se utilizan para comparar las curvas experimentales obtenidas por rayos X por Sorensen y colaboradores [114] y de difracción de neutrones por Bellissent-Funel y colaboradores [115], ya que son dos de los estudios más recientes.

Para el caso de los tautómeros de la hipoxantina se hace una discusión de lo que estas curvas representan en la estructura del disolvente alrededor de los tautómeros, ya que no hay datos experimentales sobre estas.

Sobre la integración termodinámica: En el caso de la integración termodinámica, se eligió efectuar una cuadratura Gaussiana de 11 puntos. Para tal número de puntos, los siguientes datos son necesarios para la evaluación de la integral mediante este método numérico [92]:

λ_i	0.00000	0.01592	0.08198	0.19331	0.33787	0.50000
	1.0000	0.98402	0.91802	0.80669	0.66213	
w_i	0.00000	0.04064	0.09032	0.13031	0.15617	0.16512

La elección de puntos dependió de la exactitud que se buscó y del tiempo computacional con que se contó.

Por un lado, la cuadratura Gaussiana es exacta en el intervalo evaluado para un polinomio de grado $2n - 1$, con $n > 0$, siempre y cuando los pesos w_i tengan bastantes cifras decimales. Generalmente para comprobar si la cantidad de puntos elegidos es suficiente basta graficar $\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \rangle$ contra λ , entonces se compara la curva obtenida y si ésta tiene la forma de un polinomio mayor a $2n - 1$, el número de puntos es insuficiente. Sin embargo, en este caso, el criterio más relevante es el tiempo computacional disponible. Añadir más puntos de los necesarios haría el cálculo más exacto, ya que compensa la inexactitud que puedan acarrear los pesos w_i , pero como cada punto representa una simulación más, la carga computacional se eleva bastante. Contemplando lo anterior, se hizo un estimado entre el tiempo computacional disponible y se resolvió que 11 puntos eran suficientes para realizar las simulaciones correspondientes.

De esta manera, la gráfica de $\langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \rangle$ contra λ de los resultados que se obtuvieron servirá para justificar el uso de 11 puntos y la evaluación de

$$\sum_{i=1}^n w_i \langle \frac{\partial \mathcal{V}(\lambda)}{\partial \lambda} \rangle_{\lambda_i}$$

dará como resultado la diferencia de energía libre entre los tautómeros.

Finalmente hay que señalar que las magnitudes de energía obtenida de las simulaciones son residuales, ya que en la simulación solo se van cuantificando las energías debido a las

interacciones de las partículas. Por lo que para obtener las energías totales habría que sumarle a las energías residuales las energías ideales, es decir, las del sistema sin interacciones, esto es:

$$\Delta G_{H19 \leftrightarrow H17}^{real} = \Delta G_{H19 \leftrightarrow H17}^{ideal} + \Delta G_{H19 \leftrightarrow H17}^{residual}$$

Para el caso de la energía libre de Gibbs. Afortunadamente, la energía ideal es aquella que se obtiene de los cálculos en fase gas, por lo que con los datos obtenidos de las simulaciones se obtendrán las magnitudes reales resultado del modelado.

Resultados y discusión.

18. Validación del método Monte Carlo ANES y su diferenciación del método Single - Update: Agua Pura.

Los resultados de la simulación del agua pura se presentan a continuación. Nótese que se usa esta sección para también comparar el método ANES con el Single-Update, ya que de este último se tenían datos presentes.

Las simulaciones efectuadas en el colectivo canónico utilizando el modelo MCHO ANES en el agua arrojaron resultados muy satisfactorios en cuanto a energía y estructura del agua líquida, inclusive mejores de los reportados en [84], que es la fuente original.

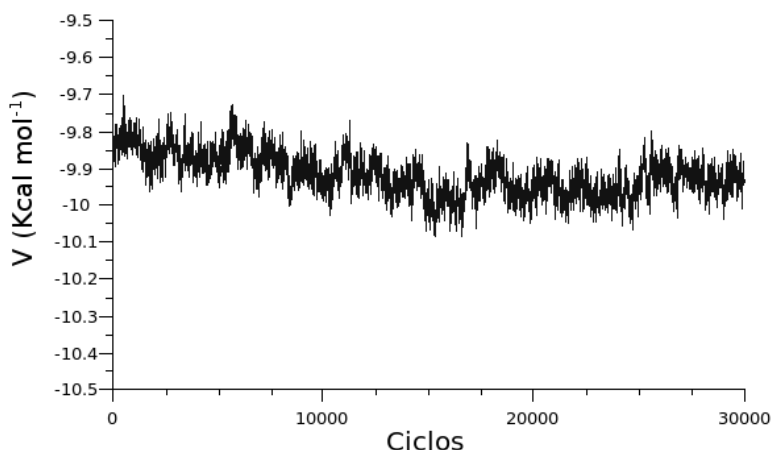


Figura 19. El comportamiento de la energía potencial del agua MCHO ANES en función del número de ciclos Monte Carlo.

La Figura 19 muestra el comportamiento de la energía potencial durante el transcurso de los primeros 30,000 ciclos Monte Carlo ($\sim 15 \times 10^6$ configuraciones) de la simulación; puede verse que al avanzar la simulación, la energía potencial oscila entre ciertos valores, sin dar saltos muy abruptos. Cuando el sistema llegó a equilibrarse ya pasadas las 100×10^6 configuraciones, se procedió a cuantificar las propiedades promedio en particular la energía configuracional promedio del agua.

Para asegurar que la sección utilizada para promediar ya estaba equilibrada se utilizó un procedimiento de análisis por bloques.

i. Energía: entalpía de vaporización.

La energía configuracional promedio del agua obtenida en la simulación MCHO ANES es de: -9.9836 ± 0.0191 Kcal mol⁻¹.

Así el valor de referencia aceptado de \mathcal{V}_l^{inter} se obtiene de la ΔH_{vap} experimental del agua [116]. Este valor de \mathcal{V}_l^{inter} corresponde a $-9.93 \text{ Kcal mol}^{-1}$. Otro valor de \mathcal{V}_l^{inter} que se maneja mucho como referencia es $-10.41 \text{ Kcal mol}^{-1}$, el cual se obtiene de cálculos teóricos con un modelo clásico (no incluye ninguna consideración cuántica) de agua rígida [117]. Puede observarse que el valor de \mathcal{V}_l^{inter} obtenido con datos experimentales concuerda bastante bien con el obtenido en la simulación, mientras que el otro valor teórico está considerablemente más desviado, lo cual es entendible porque aunque el potencial MCHO es rígido en los grados de libertad nucleares, no lo es en los grados de libertad electrónicos y se aleja mucho del modelado clásico.

Para el método Single-Update, la referencia original reporta un valor de energía de $-8.1 \text{ Kcal mol}^{-1}$. En la simulación MCHO Single-Update con sumas de Ewald se obtuvo un valor más alto de $-9.9618 \pm 0.0190 \text{ Kcal mol}^{-1}$. La discrepancia entre el valor original y el MCHO Single-Update con sumas de Ewald, se cree, es debido al tratamiento de las fuerzas de largo alcance.

De igual manera, la magnitud de la energía MCHO Single-Update con sumas de Ewald es diferente a la obtenida por el método ANES. La pequeña diferencia en la energía se atribuye a la diferencia entre los métodos ANES y Single-Update. Tal diferencia era esperada, no obstante la estimación de la energía por el método Single-Update parece ser mejor que la del método ANES.

Hay que notar que la introducción de parámetros que incluyen polarización y no aditividad aumentan significativamente la exactitud del modelo MCHO con el valor experimental de la entalpia de vaporización; sin embargo esto no representa un avance de gran magnitud puesto que modelos como el SCP o el TIP4P, rígidos y que no incluyen polarización, predicen valores de $-9.79 \text{ Kcal mol}^{-1}$ y de $-9.98 \text{ Kcal mol}^{-1}$ respectivamente para la energía configuracional, a sabiendas que este tipo de modelos han usado como parámetro de ajuste la propia entalpia de vaporización.

ii. Estructura.

Sin embargo, el éxito del modelo MCHO recae en que logra una muy buena reproducción de la estructura líquida del agua (Figura 20), sin que el modelo haya sido ajustado *a priori* con ese fin [84; 61].

Las funciones de distribución radial obtenidas de la simulación se comparan con los datos obtenidos experimentalmente por rayos X por Sorensen y colaboradores [114] y de difracción de neutrones por Bellissent-Funel y colaboradores [115].

Para el caso de la función de distribución radial O-O, se obtuvo el primer máximo de $g_{OO}(r)$ alrededor de los 2.75 \AA , así como el segundo “pico” alrededor de 4.25 \AA y 5 \AA que representa la presencia de puentes de hidrógeno (Figura 21).

Los datos experimentales de rayos X muestran que el primer pico se presenta a 2.75 Å con un valor de aproximadamente 2.8. La reproducción en el modelo es bastante buena. El segundo “pico” debería presentarse a 4.5 Å, aunque el obtenido en la simulación no es muy pronunciado, es significativo. Este segundo pico sirve como prueba de la calidad de reproducción de la estructura del agua debido a que hay modelos que no pueden reproducirlo [61].

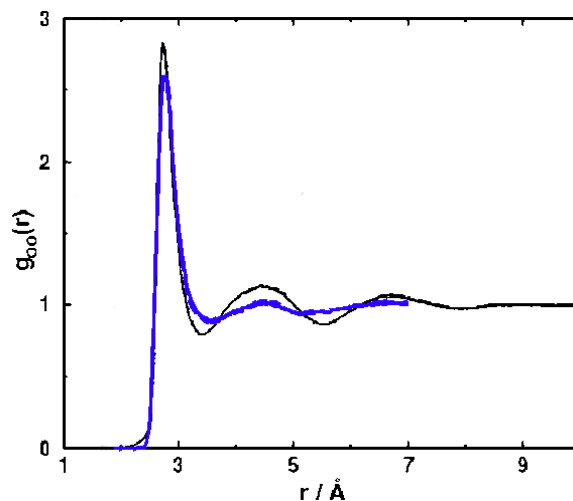


Figura 20. Funciones de distribución radial experimental [114] (negro) y la de la simulación MCHO ANES (azul).

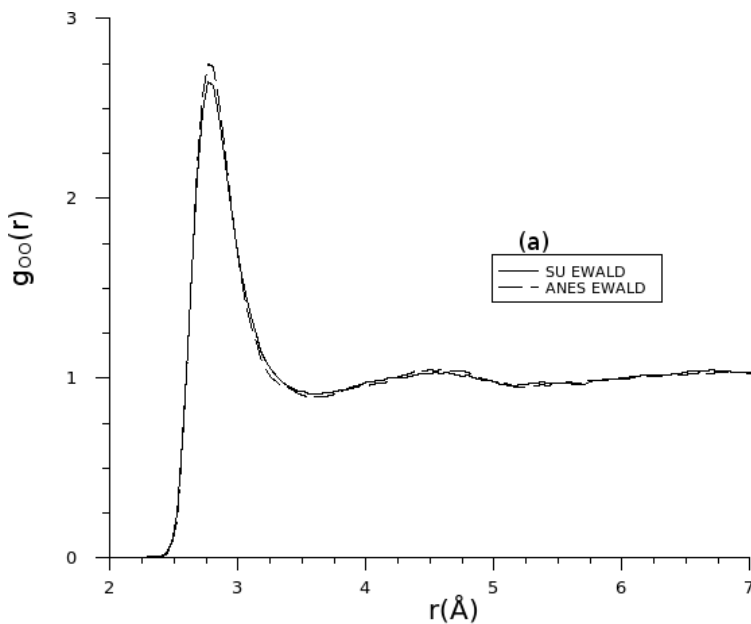


Figura 21. Función de distribución radial O-O. SU (Single-Update) y ANES.

De acuerdo a los datos de difracción de neutrones, la única diferencia es que el primer pico debería aparecer en 2.3 Å, lo cual está desviado de la medición de rayos X. La discrepancia recae en que los estudios experimentales ignoran los efectos cuánticos que tienen lugar en la estructura cuando se supone que el agua es solamente H₂O (agua ligera) y se ignora la

presencia, no despreciable, del D₂O (agua pesada) [118; 119]. Aún con los problemas experimentales el modelo puede reproducir efectivamente $g_{OO}(r)$.

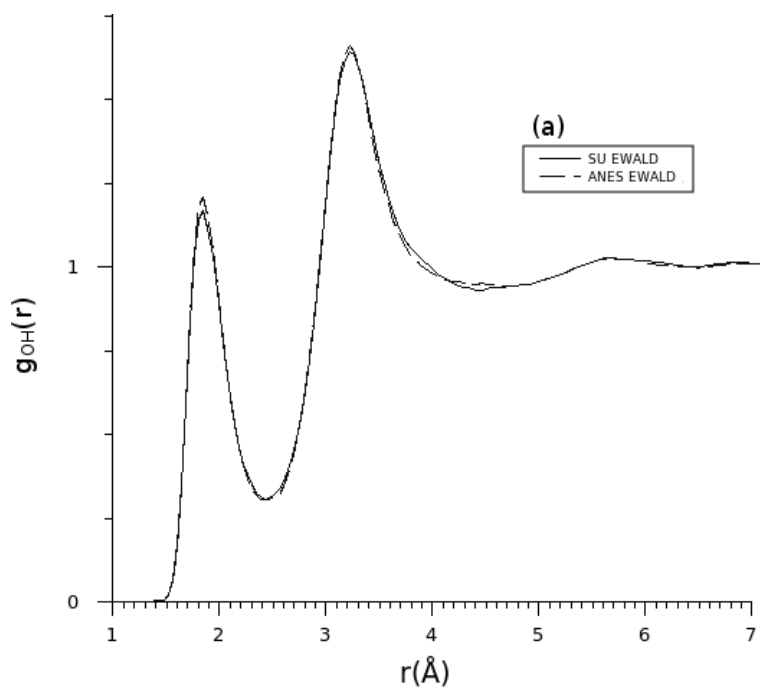


Figura 22. Función de distribución radial O-H. SU (Single-Update) y ANES.

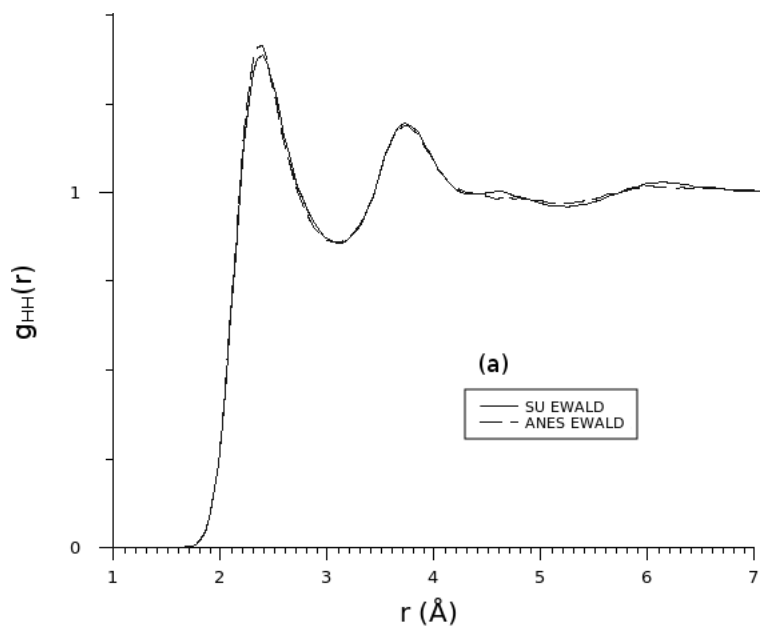


Figura 23. Función de distribución radial H-H. SU (Single-Update) y ANES.

Para $g_{OH}(r)$ y $g_{HH}(r)$, la simulación molecular reproduce la función de distribución radial O-H, con máximos a 1.8 Å y a 3.2 Å en $g_{OH}(r)$ (Figura 22) así como los picos a 2.3 Å y 3.6 Å en $g_{HH}(r)$ (Figura 23). Experimentalmente tales picos deberían estar localizados en 1.8 Å, 3.4 Å, 2.8 Å y 3.8 Å respectivamente. La tendencia es reproducida bastante bien aunque los puntos están un poco desviados, sin embargo esto puede deberse a lo comentado sobre los efectos del agua pesada, en particular porque en el entorno del H la afectación por el deuterio en el experimento es mayor.

En las figuras de las funciones de distribución radial se presentan también los gráficos de la función de distribución radial de la simulación utilizando Single-Update. La diferencia es mínima pero perceptible ya que el traslape no es al cien por ciento, particularmente en el primer pico de cada una de las funciones de distribución, donde el método ANES produce picos más acentuados que el método Single-Update.

Dado que la estructura que se obtiene en una simulación está íntimamente relacionada con la forma en cómo va evolucionando la energía, es de esperarse que si una evolución energética tiende a un valor que asegura la convergencia de la cadena de Markov, la estructura entonces será diferente a aquella obtenida de acuerdo a otro método, en el que la convergencia no está asegurada. Esta diferencia, aunque aparentemente pequeña en magnitud en las curvas de función de distribución radial obtenidas puede asegurarnos que al menos, los métodos Single-Update y ANES no nos darán el mismo resultado para sistemas iguales y por ende la convergencia del método Single-Update *sigue en duda*.

iii. Validación del método de Monte Carlo ANES.

El método de Monte Carlo ANES de agua MCHO mostro muy buenos resultados al compararlo con los experimentales. Además de los buenos resultados, el método ANES se caracteriza por tener un alto desempeño computacional, por lo que queda justificado su uso para las simulaciones hipoxantina – agua.

Por otro lado, con lo que se ha planteado matemáticamente, cualitativamente y de los resultados, el método ANES tiene argumentos que le dan una mayor aprobación para su uso que el método Single-Update. Aunque cabe mencionar que para realmente confirmar que el método Single-Update no es válido, en simulaciones de Metropolis Monte Carlo, se requerirá de un tratamiento de mayor rigor que el del presente trabajo.

19. Tautómeros en fase gas.

Mediante cálculos DFT B3LYP/6-311**G++ Se obtuvieron las geometrías en fase gas de los tautómeros de la hipoxantina. Estas geometrías son importantes debido a que la geometría en fase gas es la utilizada para la parametrización del potencial MCHO [48].

Además de la geometría de ambos tautómeros se calcularon también la diferencia de energía y energía libre entre estos, así como el momento dipolar de cada tautómero. Los resultados obtenidos son:

- $E_{Hp17} - E_{Hp19} = -1.254 \text{ Kcal mol}^{-1}$.
- $G_{Hp17} - G_{Hp19} = -0.633 \text{ Kcal mol}^{-1}$.
- $\mu_{Hp19} = 5.11 \text{ D}$
- $\mu_{Hp17} = 1.87 \text{ D}$

De los cálculos puede observarse que en fase gas el tautómero energéticamente más favorecido es la hipoxantina N1-N7. Lo cual concuerda con lo encontrado experimentalmente [19; 20] y lo estimado en otros estudios teóricos [23; 24].

Los momentos dipolares son muy parecidos a los reportados de 1.8 D para el tautómero N1-N7 y de 4.9 D para el tautómero N1-N9 [48]. Aunque el valor del momento dipolar para la hipoxantina N1-N9 parece estar sobreestimado, se acerca al intervalo aceptado de magnitudes que es de 5.2 a 5.8 D [24]. Estos resultados, de acuerdo a la hipótesis, nos mostrarán un predominio de la forma de la hipoxantina N1-N9, en el medio acuoso.

20. Tautómeros en solución.

En la introducción se presentaron hechos sobre el tautomerismo de la hipoxantina, proceso cuyo mecanismo no es del todo conocido ya que hay poca información referente a las condiciones en las que esta es favorecida en solución.

La hipoxantina es relativamente soluble en agua, aunque no mucho: 0.010 g mL^{-1} [120; 121; 122]; de hecho la mayoría de las veces en las que la hipoxantina se lista en solución, es en mezclas de agua – base, alcohol o en agua caliente. Dado que se puede efectuar una simulación en cualquier tipo de condiciones, se prestó atención en manejar una concentración capaz de alcanzarse en el laboratorio. Como se mencionó en el apartado de metodología, al terminar la simulación del agua pura se quitaron moléculas de agua y se reajustó la densidad cambiando el tamaño de la arista de la celda con el fin de proveer una cavidad que alojara a los tautómeros y a las especies “artificiales” que el método de integración termodinámica requiriera. El fin era evitar tener traslapes que condujeran a problemas con la energía. El resultado fue un balance de 1 molécula de soluto (tautómero) rodeada de 509 moléculas de disolvente (agua); lo que se traduce en una concentración de 0.1088 M en la celda, cerca de la solubilidad de la hipoxantina en agua observada a $25 \text{ }^\circ\text{C}$, la cual es de $\sim 0.073 \text{ M}$. A pesar de ello, curiosamente las condiciones a las cuales fueron llevadas las simulaciones confieren un carácter de dilución infinita a la solución, ya que no se cuantifican explícitamente las interacciones soluto – soluto.

Aunado a lo anterior hay una particularidad resaltada por algunos autores [123; 124; 125]. Las sumas de Ewald, aunque son el método más ampliamente aceptado para tratar las interacciones electrostáticas, confiere al sistema una condición de periodicidad, al replicar la celda primitiva de la simulación. Si bien esta situación puede resultar sutil en sistemas pequeños, en sistemas de tamaño importante (por ejemplo, simulación de cadenas

poliméricas) puede tener efectos muy importantes en la estructura del sistema, dando lugar a resultados desviados de los observados experimentalmente. De hecho debido a que no se cuentan las interacciones soluto – soluto explícitamente, el efecto de la periodicidad del sistema es algo no deseable, ya que las réplicas de la celda hacen parecer en el cálculo de las interacciones electrostáticas que el soluto no está a dilución infinita. Lo cual es paradójico con lo comentado al inicio.

Computacionalmente aunque se puede pensar que estas situaciones pueden dar lugar a problemas, en la práctica, no es tan preocupante. Como el desarrollo de la simulación se va dando por medio de diferencias de energía potencial, los errores que se presentan se van cancelando, en gran parte, al efectuar el cálculo de la diferencia de energía en dos configuraciones.

i. Energía: tautómeros en solución.

De acuerdo a las condiciones planteadas en la metodología, las energías configuracionales promedio obtenidas fueron:

- $\mathcal{V}_{HP19} = -10.151 \pm 0.016 \text{ Kcal mol}^{-1}$.
- $\mathcal{V}_{HP17} = -10.031 \pm 0.015 \text{ Kcal mol}^{-1}$.

Como es de esperarse, no existe gran diferencia en la energía entre ambos tautómeros:

- $\mathcal{V}_{HP17} - \mathcal{V}_{HP19} = 0.120 \text{ Kcal mol}^{-1}$.

Comparado a la diferencia de energías en fase gas ($-1.254 \text{ Kcal mol}^{-1}$), hay un cambio en el signo. Esto significa que al pasar de fase gas a fase acuosa el tautómero Hp19 es estabilizado en mayor medida por el medio que el tautómero Hp17. Si negáramos los efectos entrópicos esto sugeriría que en la reacción de tautomerización, el tautómero Hp19 sería el favorecido en el entorno acuoso. No obstante la diferencia energética no es tan grande por lo que efectos entrópicos pequeños podrían cambiar la magnitud de este número. No será hasta que pueda obtenerse la constante de equilibrio cuando se pueda decir cuantitativamente cual tautómero es el predominante.

Como punto de comparación existe el trabajo realizado por M. E. Costas y colaboradores [48], en él se realizaron simulaciones similares (sólo especies puras), con el potencial MCHO y el método Single-Update; el estudio determina energías configuracionales de $-10.46 \text{ Kcal mol}^{-1}$ y $-10.64 \text{ Kcal mol}^{-1}$ para los tautómeros Hp17 y Hp19, respectivamente.

La magnitud de las energías son ligeramente mayores a las obtenidas en este trabajo; sin embargo los resultados son congruentes en el sentido de que los dos dan al Hp19 como el tautómero energéticamente más estable. Por otro lado la magnitud de la diferencia en las energías potenciales que se reportan en [48] también es bastante similar $\mathcal{V}_{HP17} - \mathcal{V}_{HP19} = 0.18 \text{ Kcal mol}^{-1}$.

Se piensa que las diferencias son ocasionadas por las distintas metodologías usadas; en la referencia [48] las simulaciones usan el método Single-Update y truncamiento esférico mientras que en este trabajo se usa el método ANES y las sumas de Ewald.

ii. Estructura agua – hipoxantina.

Para observar el entorno de ciertas zonas de la molécula (Figura 24) por separado, se obtuvieron los perfiles de las funciones de distribución radial locales de estas zonas del tautómero con respecto al oxígeno de las moléculas de agua (Figura 25). Debido a que, químicamente, la molécula de hipoxantina no es homogénea al irse moviendo a través de su estructura se pueden encontrar diferencias en cuanto a la afinidad de las moléculas de agua a la zona.

Las estructuras de los tautómeros sugieren que las zonas marcadas en el anillo pirimidínico (zonas 1 y 2) deberían ser similares entre las especies N1-N7 y N1-N9, mientras que las zonas pertenecientes al anillo imidazólico (zonas 3 y 4) deberían presentar cambios importantes al pasar de un tautómero a otro. Los perfiles obtenidos muestran que esto es cierto (Figura 25): La región del anillo de imidazol (zonas 3 y 4) muestra los cambios más significativos en la estructura del disolvente. La razón es el cambio en la posición del hidrógeno involucrado en la reacción de tautomerización.

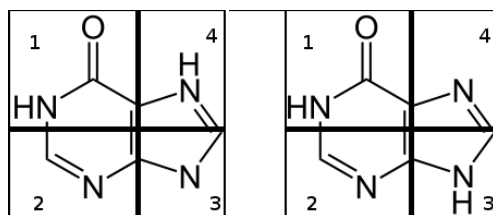


Figura 24. Definición de las zonas para la generación de funciones de distribución radial zona hipoxantina – Oagua.

También puede verse que en uno de los tautómeros el disolvente se encuentra más estructurado:

- En la especie N1-N7, La zona 1 y la zona 4 presentan dos picos bien definidos a distancias menores a 4 Å.
- La especie N1-N9 por otro lado sólo tiene una zona que presenta dos picos bien definidos, la zona 1; mientras, la zona 3 muestra un pico bien definido, aunque con un valor de ρ de 1 y después picos menos acentuados con densidad atómica de 1.

Esto significa que la estructura del disolvente alrededor de la hipoxantina N1-N7 tiene una mayor densidad que alrededor de la hipoxantina N1-N9.

El acomodo de las moléculas de agua alrededor de los tautómeros podría indicar que el tautómero mayormente solvatado fuese el más estable en solución.

De acuerdo a los resultados de la energía configuracional, el tautómero con la menor energía es la especie N1-N9. Sin embargo, el tautómero mayormente solvatado es el N1-N7. Esto puede llevar a pensar que hay otros efectos con mayor peso. En particular, se ha

hablado que debido a la gran diferencia en el momento dipolar entre las especies N1-N9 y N1-N7 es por lo que existe una diferencia entre la población de estas dos especies. De esta manera, es muy probable que este efecto de mayor peso sea el de la polarización de las moléculas de agua alrededor de los tautómeros. Lo anterior justifica la elección de un modelo polarizable.

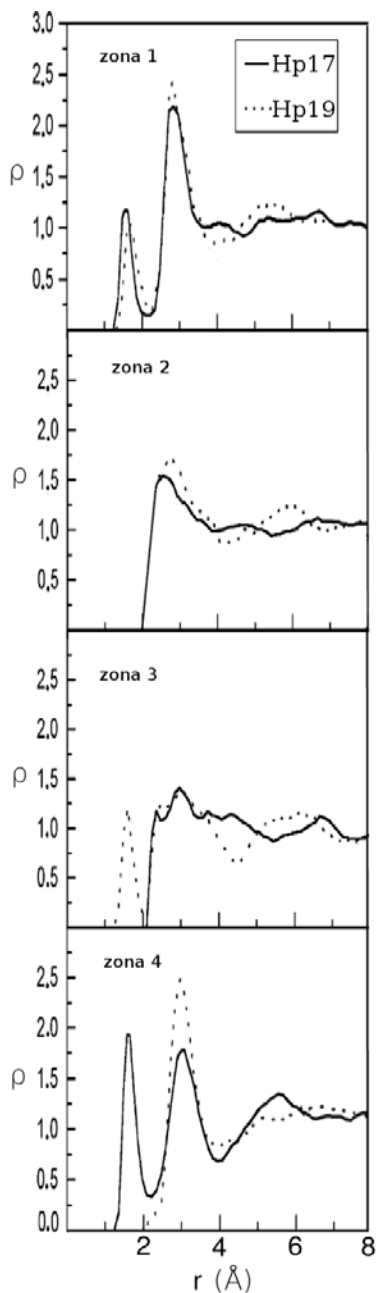


Figura 25. Funciones de distribución radial zona hipoxantina – Oagua.

El arreglo del disolvente tiene otras implicaciones. Se ha visto que éste se estructura en mayor medida alrededor de la Hipoxantina N1-N7. Si la reacción de tautomerización involucra mecanismos donde el agua interviene reaccionando (lo cual es muy probable

como se ha mostrado en infinidad mecanismos de reacciones químicas [1]), no cabe duda que el arreglo del disolvente alrededor de la molécula de hipoxantina debe asumir un papel importante en la reacción y por lo tanto el conocimiento de su estructura toma una mayor relevancia.

21. Energía libre y constante de equilibrio de la reacción de tautomerización.

Dado que el potencial compuesto usado fue:

$$V(\lambda) = (1 - \lambda)V_{H19} + \lambda V_{H17}$$

La energía libre calculada va a estar dada como el paso de $\lambda=0$ a $\lambda=1$ o en otras palabras de la hipoxantina N1-N9 a la N1-N7.

Los resultados obtenidos fueron:

λ_i	$\left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle$
0.00000	4.9995×10^{-2}
0.01592	4.7195×10^{-2}
0.08198	4.4684×10^{-2}
0.19331	3.7468×10^{-2}
0.33787	2.7113×10^{-2}
0.50000	1.5058×10^{-2}
0.66213	1.2864×10^{-4}
0.80669	-1.2422×10^{-2}
0.91802	-2.2150×10^{-2}
0.98402	-3.0083×10^{-2}
1.00000	-3.1382×10^{-2}

La Figura 26 presenta la forma de la curva correspondiente a estos datos. Puede entonces verse que en ese intervalo la curva sólo corta las abscisas una vez, por lo que al menos la curva obtenida puede pertenecer a un polinomio de grado 1. Con 11 puntos evaluados, estos son suficientes para un polinomio de grado 21, por lo que se considera que el número de puntos tomados es adecuado para evaluar la integral $\int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda$.

Con la integración termodinámica se obtiene una energía libre de Helmholtz de reacción (A) de 1.274×10^{-2} Kcal mol⁻¹.

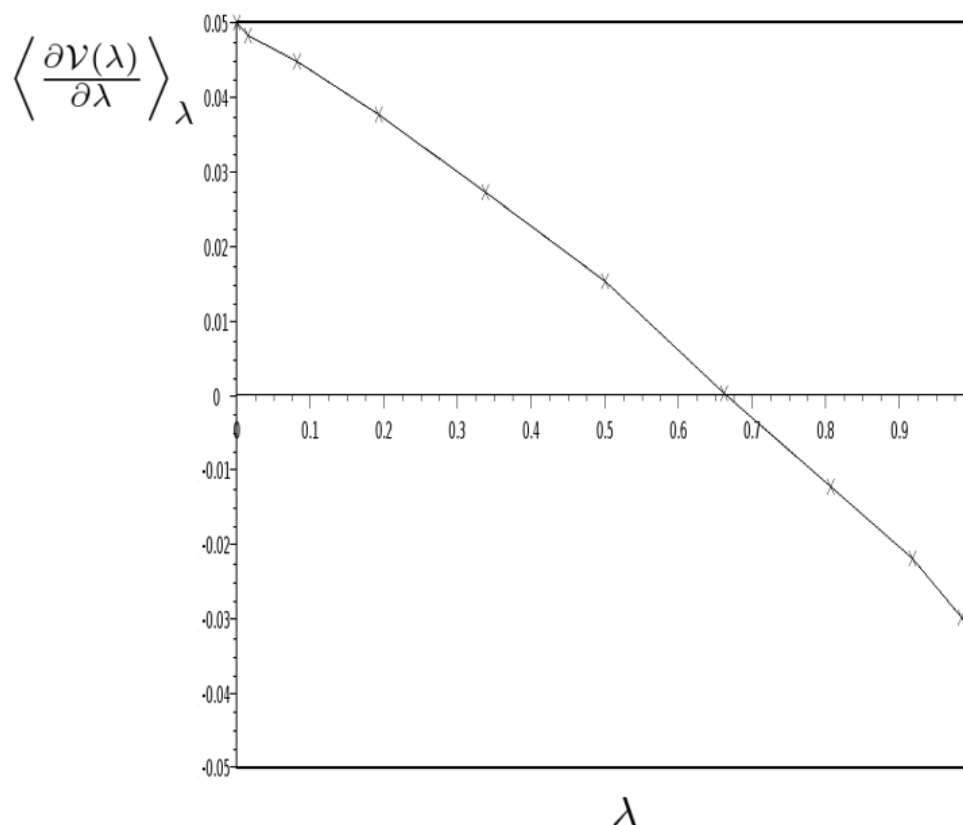


Figura 26. Resultado de la integración termodinámica.

Para calcular la energía libre de Gibbs de reacción (G), que es la de mayor interés, se utilizó la relación termodinámica siguiente:

$$dA = dG + d(PV)$$

Los cambios en volumen son 0 debido a que siempre se trabajó en el colectivo canónico. Por otro lado, se asume que los cambios de presión no son significativos entre las simulaciones. Por lo tanto el término $d(PV)$ puede ser despreciado y que $dA \approx dG$

Debido a la forma en cómo se calculan las energías durante una simulación molecular, el resultado obtenido es una energía libre residual. Así:

$$\Delta G_{H19 \leftrightarrow H17}^{real} = \Delta G_{H19 \leftrightarrow H17}^{ideal} + \Delta G_{H19 \leftrightarrow H17}^{residual}$$

Donde $\Delta G_{H19 \leftrightarrow H17}^{ideal}$ es aquella obtenida en fase gas.

Con $\Delta G_{H19 \leftrightarrow H17}^{gas} = 0.633$ y $\Delta A_{H19 \leftrightarrow H17}^{residual} = 1.274 \times 10^{-2}$ y

$$\Delta G_{H19 \leftrightarrow H17}^{real} \approx \Delta G_{H19 \leftrightarrow H17}^{gas} + \Delta A_{H19 \leftrightarrow H17}^{residual}$$

Se obtiene una energía libre de Gibbs de **0.64574 Kcal mol⁻¹** para la reacción:

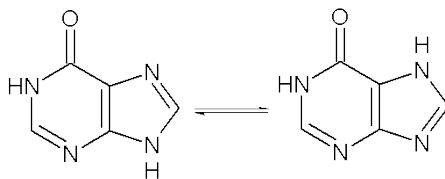


Figura 27. Reacción H19 ↔ H17.

Conociendo la energía libre de reacción, la obtención de la constante de equilibrio y las poblaciones es simple.

Usando la relación

$$\Delta G_{H19 \leftrightarrow H17}^{real} = -RT \ln K_c$$

Donde K_c es la constante de equilibrio de la reacción en solución, **se obtiene una constante de equilibrio $K_c = 0.3361$.**

Entonces, bajo la hipótesis de solución ideal (dilución infinita) $[H17] = 0.3361[H19]$.

En otras palabras, el resultado final es que, de acuerdo a la metodología empleada, en solución acuosa la cantidad de Hipoxantina en la forma de su tautómero N1-N7 es aproximadamente la tercera parte de lo que hay del tautómero N1-N9.

Retomando los datos mencionados en la introducción, en general se reconoce que la tendencia observada experimentalmente sea el predominio del tautómero N1-N9 en solución [11; 31] y aunque la proporción varíe entre diferentes fuentes, el intervalo de población del tautómero N1-N9 estimado se encuentre entre el 60% y 75% en solución acuosa, calculado con diferentes modelos teóricos, inclusive cálculos mecánico cuánticos que introducen los efectos de solvatación mediante disolventes continuos [19; 13; 15; 20; 40]. La concordancia con los resultados de las simulaciones es bastante buena, siendo la proporción obtenida $[Hp17] = 0.3361[Hp19]$ equivalente a decir que el tautómero N1-N9 se encuentra presente en 74.8% y el tautómero N1-N7 en 25.2 %.

Un factor que se piensa puede contribuir también en la magnitud de estos valores es la interacción soluto – soluto, que nunca fue tomada en cuenta explícitamente en la simulación; es importante notar que aunque excluir esta interacción simplifica el manejo de los cálculos es muy probable que esta tenga una contribución importante en la energía. Lo anterior basado en el hecho de que muchos heterociclos orgánicos suelen interactuar entre ellos mismos en solución y formar dímeros [1; 3]. La anterior situación también sugeriría un nuevo efecto a tomar en cuenta para la estimación de las poblaciones relativas.

Conclusiones.

22. Tautomerismo de la hipoxantina MCHO.

De acuerdo al objetivo principal del trabajo:

Se calculo la energía libre de reacción de tautomerización de forma satisfactoria mediante simulaciones moleculares utilizando el método de Metropolis Monte Carlo, el método ANES y la integración termodinámica.

A su vez la constante de equilibrio de la reacción y las poblaciones relativas fueron calculadas también a partir de la energía libre obtenida, cumpliendo con el objetivo principal del trabajo.

Se encontró que la población relativa del tautómero de la hipoxantina N1-N9 está en una proporción **1:0.3361** con respecto al tautómero N1-N7. Dicha relación concuerda bien con las reportadas en la literatura con respecto a otros estudios teóricos que utilizan métodos mecano-cuánticos.

Las hipótesis se comprueban. Uno de los tautómeros se encuentra en mayor proporción con respecto al otro. Además al incluir un modelo que trate la polarización en las moléculas se ha llegado a un resultado concordante con los diversos datos reportados en la literatura.

De acuerdo a las metas planteadas, complementarias a los objetivos principales:

El estudio cuántico en fase gas arrojó los resultados esperados y que se reportan extensivamente en la literatura.

Por otro lado en solución, estructuralmente se observó que el tautómero N1-N7 tiende a tener una mayor cantidad de moléculas de agua a su alrededor que el tautómero N1-N9, lo cual se considera un factor que puede tener relevancia en el mecanismo por el cual ocurre la reacción de tautomerización.

También, pudo llegarse a la conclusión de que aunque, la solvatación estabiliza a los tautómeros tal efecto no es tan potente como los efectos de polarización de las moléculas de soluto en el medio, al menos no en el caso del agua. Así, para modelar sistemas acuosos o moléculas con altos momentos dipolares, incluir efectos de polarización es muy importante.

Debido a lo anterior, el modelo MCHO presenta una propuesta interesante y viable para la realización de simulaciones moleculares que deban incluir polarización. El problema con el modelo es que si no hay datos sobre una interacción en particular que se quiera estudiar se tienen que obtener los parámetros ajustando a una superficie de potencial lo cual es una tarea que requiere de mucho tiempo y recursos computacionales si se quieren obtener parámetros de buena calidad.

Por otro lado, en relación al muestreo de los grados de libertad electrónicos, no se pudo llegar a una conclusión sólida acerca del método Single-Update ya que la sola simulación del agua pura no arrojó datos convincentes para dicho fin. Aunque se planteó la posibilidad, es necesario un estudio formal y más riguroso para concluir contundentemente sobre este aspecto. Aún así, el método ANES se presenta como una alternativa sólida y viable para el muestreo de los grados de libertad electrónicos.

Comentarios finales:

Aunque ambos tautómeros coexisten en solución acuosa hay que recordar que el cálculo de población se hizo tomando en cuenta que no había interacciones explícitas soluto-soluto, por lo que los resultados obtenidos pueden ser refinados aún más al incluir el modelado de estos ya que en la naturaleza este tipo de moléculas tienden a asociarse.

Finalmente, la metodología usada puede extenderse para modelar otros heterociclos ampliamente estudiados como lo son la adenina y la guanina. Resultados satisfactorios en tales moléculas podrían llegar a validar y extender el método para el estudio de sus polímeros: las cadenas de ácidos nucleicos. Tal alcance ayudaría al avance en el campo del modelado de heterociclos y aportaría así nuevos datos al estudio de las reacciones en solución.

Apéndices.

1. Parámetros MCHO para las interacciones agua-agua.

De acuerdo a la ecuación

$$V_{AB} = \sum_{A < B} \left[\sum_{i \in A} \sum_{j \in B} \left(\Lambda_{AB} e^{-n_1 R_{AB}} + \Theta_{AB} e^{-n_2 R_{AB}} + C \frac{q_i q_j}{r_{ij}} + \frac{k}{r_{ij}^{n_3}} \right) \right] + \sum_A \sum_{i \in A} K_A r_{A_i}^2 + \sum_B \sum_{j \in B} K_B r_{B_j}^2$$

Todos los parámetros son tal que las energías quedan dadas en Kcal mol⁻¹, distancias en Bohr y cargas en e⁻ (carga del electrón).

Átomo	q _i	K _A (o K _B)
O	-1.426	59.012
H	0.713	455.059

Par átomo A-B	Λ	n ₁	Θ	n ₂
O – O	3.471402x10 ⁸	3.576	-1.748782x10 ⁹	3.978
O – H	6.919585x10 ²	1.489	-3.447577x10 ¹	0.701
H – H	1.528731x10 ¹	0.634	1.183873x10 ²²	15.203

Par átomo A-B intermolecular	C	k	n ₃
O – O	627.5	0.0	0.0
O – H	627.5	7118.564	6.490
H – H	627.5	0.0	0.0

Par átomo A-A intramolecular	C	k	n ₃
O – O	0.0	0.0	0.0
O – H	293.67	99.54525	2.418
H – H	0.0	0.0	0.0

La geometría de la molécula tiene los siguientes parámetros con sus unidades señaladas:

r _{O-H} (Å)	0.9572
∠ HOH (°)	104.5

2. Parámetros MCHO para las interacciones hipoxantina-agua.

De acuerdo a la ecuación

$$V_{AB} = \sum_{A < B} \left[\sum_{i \in A} \sum_{j \in B} \left(\Lambda_{AB} e^{-n_1 R_{AB}} + \Theta_{AB} e^{-n_2 R_{AB}} + C \frac{q_i q_j}{r_{ij}} + \frac{k}{r_{ij}^{n_3}} \right) \right] + \sum_A \sum_{i \in A} K_A r_{A_i}^2 + \sum_B \sum_{j \in B} K_B r_{B_j}^2$$

Todos los parámetros son tal que las energías quedan dadas en Kcal mol⁻¹, distancias en Bohr y cargas en e⁻ (carga del electrón).

La numeración de los átomos entre paréntesis, igualmente está dada de acuerdo a la mostrada en la Figura 28. El H14 es el hidrógeno que se intercambia en el tautomerismo. Es equivalente en términos de las magnitudes de sus parámetros MCHO en ambos tautómeros. Si no hay numeración entre paréntesis significa que todo átomo de este tipo tiene los mismos valores; para los átomos de agua, se les añade el subíndice “a”.

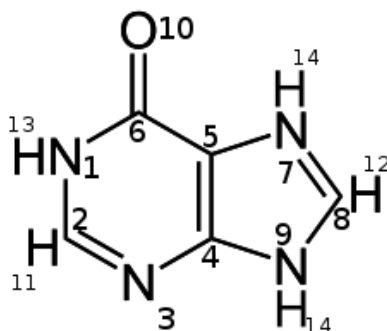


Figura 28. Numeración en los átomos de la hipoxantina.

Átomo	q _i	K
O	-0.668	188.290
N	-0.502	120.590
C(2), C(8)	0.299	79.977
C(4), C(5)	0.352	666.293
C(6)	0.521	398.945
H(13), H(14)	0.099	80.007
H(11), H(12)	0.329	80.020

Par átomo A-B	Λ	n_1	Θ	n_2
$O_a - O$	0.8394529×10^5	1.283	-0.7683871×10^5	1.266
$H_a - O$	0.1204014×10^6	3.310	-0.4856816×10^7	4.698
$O_a - N$	0.2328124×10^7	2.968	-0.1513859×10^7	3.820
$H_a - N$	0.5568232×10^4	2.148	-0.7305201×10^7	4.638
$O_a - C$	0.2150520×10^6	2.089	-0.1092361×10^7	2.539
$H_a - C$	0.7434896×10^6	4.875	-0.9996247×10^3	1.1958
$O_a - H$	0.7007117×10^5	2.728	-0.2823186×10^6	3.367
$H_a - H$	0.9504891×10^6	4.690	-0.1855791×10^4	4.422

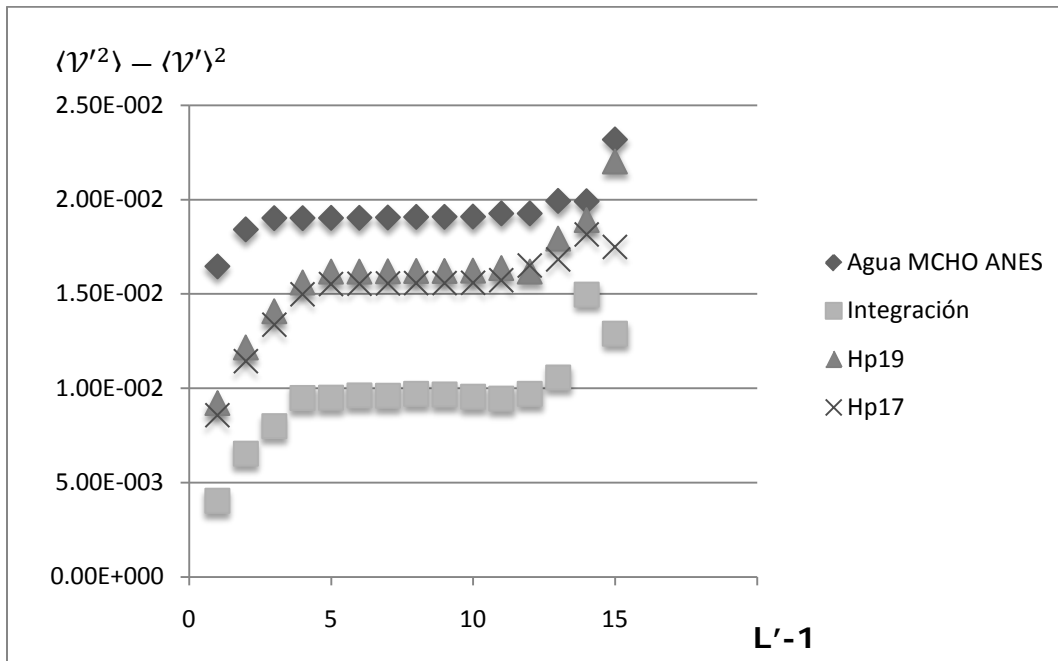
Par átomo A-B	C	K	n_3
$O_a - O$	627.5	0.1542430×10^8	10.572
$H_a - O$	627.5	0.5668371×10^6	9.541
$O_a - N$	627.5	0.7332425×10^9	12.777
$H_a - N$	627.5	0.1494846×10^6	12.134
$O_a - C$	627.5	0.6395584×10^7	11.159
$H_a - C$	627.5	0.8628204×10^6	8.032
$O_a - H$	627.5	0.7522646×10^5	12.310
$H_a - H$	627.5	0.2388443×10^2	6.805

Par átomo A-B	C	K	n_3
$O - N$	62.75	0.7877169×10^4	6.810
$N - N$	131.1475	0.1735182×10^4	1.710
$C - O$	4.7056×10^3	0.2514080×10^5	9.279
$O - H$	87.2225	0.9921085×10^4	9.835
$C - N$	827.045	0.1022239×10^3	10.923
$N - H$	1.7808×10^3	0.4888883×10^3	2.165
$C - C$	85.9675	0.2473304×10^4	3.411
$C - H$	815.75	0.4128202	5.091
$H - H$	2.5307×10^3	0.9006147×10^3	0.756

La geometría de la molécula tiene que ser obtenida de una optimización de geometría en fase gas.

3. Análisis por bloques.

Las gráficas del análisis por bloques se muestran a continuación. Las mesetas indican que las simulaciones son lo suficientemente largas:



Para la integración los valores son muy similares por lo que sólo se muestra uno.

Bibliografía.

1. **Smith M. B., March J.** *Advanced Organic Chemistry*. New York : Wiley Interscience, 2001.
2. **Katritzky A.R., Elguero J., et. al.** *The tautomerism of heterocycles*. New York : Academic Press, 1976.
3. **Saenger, W.** *Principles of nucleic acid structure*. New York : Springer, 1984.
4. **Rosemeyer, H.** : *Chemistry & Biodiversity*, 2004, Vol. 1, p. 361.
5. **Stryer, L.** *Biochemistry*. New York : Freeman, 1988.
6. **Friedberg E. C., Walker G. C., Siede W.** *DNA repair and mutagenesis*. Washington : ASM Press, 1995.
7. **Chemistry, International Union of Pure and Applied.** *Nomenclature of Organic Chemistry*. Oxford : Pergamon, 1979.
8. **Watson, J. D. and Crick, F. H. C.** : *Nature*, 1953, Vol. 171, p. 737.
9. **Kwiatkowski, J. S. and Pullman, B.** : *Adv. Het. Chem.*, 1975, Vol. 18, p. 199.
10. **Topal, M. D. and Fresco, J. R.** : *Nature*, 1976, Vol. 260, p. 285.
11. **Lichtenberg, D., Bergmann, F. and Neiman, Z.** : *Isr. J. Chem.*, 1972, Vol. 10, p. 805.
12. **A. Gerega, L. Lapinski, M. J. Nowak, H. Rostkowska.** : *J. Phys. Chem. A*, 2006, Vol. 110, p. 10236.
13. **Lin J., Yu C., Peng S., Akiyama I., Li K., Li-Kao L., Lebreton P. R.** : *J. Phys. Chem.*, 1980, Vol. 84, p. 1006.
14. **S. Gogia, A. Jain, M. Puranik.** : *J. Phys. Chem. B*, 2009, Vol. 113, p. 15101.
15. **Sheina G. G., Stephanian S. G., Radchenko E. D., Blagoi Y. P.** s.l. : *J. Mol. Struct.*, 1987, Vol. 158, p. 275.
16. **Schmalle H. W., HaEnggi G., Dubler E.** : *Acta Crystallogr.*, 1988, Vol. C44, p. 732.
17. **Psoda A., Shugar D.** : *Biochim. Biophys. Acta*, 1971, Vol. 247, p. 507.
18. **Medeiros G. C., Thomas Jr G. J.** : *Biochim. Biophys. Acta*, 1971, Vol. 238, p. 1.
19. **Ramaekers G., Maes L., Adamowicz A., Dkhissi J.** : *J. of Mol. Struct.*, 2001, Vol. 560, p. 205.
20. **Shukla M. K., Leszczynski J.** : *J. Phys. Chem. A*, 2000, Vol. 104, p. 3021.
21. **Contreras J. G., Alderete J.B.** : *Mol. Engng*, 1992, Vol. 2, p. 29.
22. **Nonella M., HaEnggi G., Dubler E.** : *J. Mol. Struct. (Theochem)*, 1993, Vol. 279, p. 173.
23. **Hernandez B., Luque F. J., Orozco M.** : *J. Org. Chem.*, 1996, Vol. 61, p. 5964.
24. **Costas M. E., Acevedo-Chavez R.** : *J. Phys. Chem. A*, 1997, Vol. 101, p. 8309.
25. **Becke, A. D.** : *J. Chem. Phys.*, 1993, Vol. 98, p. 5648.
26. **Lee C., Yang W., Parr R. G.** : *Phys. Rev. B*, 1988, Vol. 37, p. 785.
27. **Van Bael M. K., Schoone K., Houben L., Smets J., McCarthy W., Adamowicz L., Nowak M. J., Maes G.** : *J. Phys. Chem. A*, 1997, Vol. 101, p. 2397.
28. **Shoone K., Smets J., Houben L., Van Bael M. K., Adamowicz L., Maes G.** : *J. Phys. Chem. A*, 1998, Vol. 102, p. 4863.
29. **Smets J., Schoone K., Ramaekers R., Adamowicz L., Maes G.** : *J. Mol. Struct.*, 1998, Vol. 442, p. 201.
30. **Pullman B., Pullman A.** : *Adv. Heterocycl. Chem.*, 1971, Vol. 13, p. 77.
31. **Chenon M. T., Pugmire R. J., Grant D. M., Panzica R. P., Townsend L. B.** : *J. Am. Chem. Soc.*, 1975, Vol. 97, p. 4636.

32. **Cheng D. M., Kan L.S., Ts'o P. O. P., Pettre C. G., Pullman B.** : J. Am. Chem. Soc., 1980, Vol. 102, p. 525.
33. **Szczepaniak K., Szczepaniack M.** : J. Mol. Struct., 1987, Vol. 156, p. 29.
34. **Person W. B., Szczepaniak K., Szczepaniack M., Kwiatkowski J. S., Hernandez L., Czerminski R.** : J. Mol. Struct., 1989, Vol. 194, p. 239.
35. **D. M. Grant, R. P. Panzica, M. T. Chenon,, R. J. Pugmire, L. B. Townsend.** : J. Am. Chem. Soc., 1975, Vol. 97, p. 4636.
36. **Nowak M. J., Lapinski L., Kwiatkowski J. S.** : Chem. Phys. Lett., 1989, Vol. 157, p. 14.
37. **Santhosh C., Mishra P. C.** : Spectrochim. Acta, 1993, Vol. A49, p. 985.
38. **Dkhissi A., Houben L., Smets J., Adamowicz L., Maes G.** 104, : J. Phys. Chem. A, 2000, p. 9785.
39. **Barone V., Adamo C.** : J. Phys. Chem., 1995, Vol. 99, p. 15062.
40. **B. Hernandez, F. J. Luque, M. Orozco.** : J. Org. Chem., 1996, Vol. 61, p. 5964.
41. **Maginn, E.J.** : J. Phys.: Condens. Matter, 2009, Vol. 21, p. 373101.
42. **Maginn, E.J.** : AIChE Journal, 2009, Vol. 55, p. 1304.
43. **J. L Klepeis, K. Lindorff-Larsen, R. O. Dror, D. E. Shaw.** : Curr. Op. in Stru. Bio., 2009, Vol. 19, p. 120.
44. **Kollman, P.** : Chem. Rev., 1993, Vol. 93, p. 2395.
45. **L. R. Rutledge, C. A. Wheaton, S. D. Wetmore.** : Phys. Chem. Chem. Phys., 2007, Vol. 9, p. 497.
46. **X. J. Sun, J. K. Lee.** : J. Org. Chem., 2007, Vol. 72, p. 6548.
47. **K. Malinakova, L. Novosadova, M. Lahtinen, E. Kolehmainen, J. Brus, R. Marek.** : J. Phys. Chem. A, 2010, Vol. 114, p. 1985.
48. **M. E. Costas, M. L. San Román Zimbrón, R. A. Chávez.** : J. Mol. Struct. (THEOCHEM), 2004, Vol. 711, p. 83.
49. **Allen M. P., Tildesley D. J.** *Computer simulation of liquids.* Oxford : Clarendon Press, 1987.
50. **Frenkel D., Smit B.** *Understanding molecular simulation from algorithms to applications.* : Academic Press, 2002.
51. **G. Bao, S. Suresh.** : Nature Materials, 2003, Vol. 2, p. 715.
52. **M. Karplus, J. A. McCammon.** : Nature Structural Biology, 2002, Vol. 9, p. 646.
53. **McQuarrie, D. A.** *Statistical Mechanics.* New York : Harper Collins Publishers, 1976.
54. **S. H. Friedberg, A. J. Insel, L. E. Spence.** *Algebra Lineal.* México : Publicaciones cultural S.A., 1982.
55. **Luis R. I., Segismundo S. I., José Ignacio S.** : J. of Art. So. and Soc. Sim., 2009, Vol. 12, pág. 6.
56. **N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.N. Teller, E. Teller.** : J. Chem. Phys., 1953, Vol. 21, pág. 1087.
57. **J. G. Kirkwood, E. K. Maun, B. J. Alder.** : J. Chem. Phys., 1950, Vol. 18, p. 1040.
58. **Rotenberg, A. s.l.** : J. Chem. Phys., , Vol. 43, p. 1198.
59. **W. G. Hoover, M. Ross, K. W. Johnson, D. Henderson, J. A. Barker, B. C. Brown.** : J. Chem. Phys., Vol. 52, p. 4931.
60. **Lennard-Jones, J.E.** : Proc. R. Soc. Lond. A, 1924, Vol. 106, pág. 463.
61. **Guillot, B.** : J. Mol. Liq., 2002, Vol. 101, p. 219.
62. **H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans.** *Intermolecular Forces.* Reidel, Dordrecht : B. Pullman, 1981.

63. **W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein.** : J. Chem. Phys., 1983, Vol. 79, p. 926.
64. **J.K. Gregory, D.C. Clary, K. Liu, M.G. Brown, R.J. Saykally.** : Science, 1997, Vol. 275, p. 814.
65. **S.W. Rick, S.J. Stuart, B.J. Berne.** : J. Chem. Phys., 1994, Vol. 101, p. 6141.
66. **W. A. Wakeham, M. Rigby, E. B. Smith, G. C. Maitland.** *The force between molecules.* Oxford : Oxford Science Publications, 1986.
67. **Halgren, T. A.** : J. Am. Chem. Soc., 1992, Vol. 114, p. 7827.
68. **H. J. C. Berendsen, J.R. Grigera, T.P. Straatsma.** : J. Phys. Chem., 1987, Vol. 114, p. 6269.
69. **L. Kuyper, D. Ashton, K.M. Merz Jr., P.A. Kollman.** : J. Phys. Chem., 1991, Vol. 95, p. 6661.
70. **Böttcher, F.C.J.** *Theory of Electric Polarization.* New York : Elsevier, 1973.
71. **P. Cieplak, T.P. Lybrand, P.A. Kollman.** : J. Chem. Phys., 1987, Vol. 86, p. 6393.
72. **S. W. Rick, S. J. Stuart.** : Rev. in Comp. Chem., 2002, Vol. 18, p. 89.
73. **M. Born, K. Huang.** *Dynamical Theory of Crystal Lattices.* Oxford : Oxford University Press, 1954.
74. **B.G. Dick, A.W. Overhouser.** : Phys. Rev., 1958, Vol. 112, p. 90.
75. **C.R.A. Catlow, G.D. Price.** : Nature, 1990, Vol. 347, p. 243.
76. **P.J. Mitchell, D. Fincham.** : J. Phys.: Condens. Matter, 1993, Vol. 5, p. 1031.
77. **Mulliken, R. S.** : J. Chem. Phys., 1934, Vol. 2, p. 782.
78. **R. G. Parr, R. G. Pearson.** : J. Am. Chem. Soc., 1983, Vol. 105, p. 7512.
79. **R.G. Parr, R.A. Donnelly, M. Levy, W.E. Palke.** : J. Chem. Phys., 1978, Vol. 68, p. 3801.
80. **Sanderson, R. T.** : Science, 1954, Vol. 114, p. 670.
81. **W.J. Mortier, K.V. Genechten, J. Gasteiger.** : J. Am. Chem. Soc., 1985, Vol. 107, p. 829.
82. **Field, M.J.** : Mol. Phys., 1997, Vol. 91, p. 835.
83. **Gao, J.** 101, : s.n., 1997, Vol. J. Phys. Chem. B, p. 657.
84. **C. Medina Llanos, H. Saint Martín, I. Ortega Blake.** : J. Chem. Phys., 1990, Vol. 93, p. 6448.
85. **E. Clementi, O. Matsuoka, M. Yoshimine.** : J. Chem. Phys., 1976, Vol. 64, p. 1351.
86. **G. C. Lie, E. Clementi, W. Kolos, G. Ranghino.** : Int. J. Quantum Chem., 1980, Vol. 17, p. 377.
87. **Sprik, M.** : J. Chem. Phys., 1991, Vol. 95, p. 6762.
88. **B. Chen, M.G. Martin, J. I. Siepmann.** : Theor. Chem. Acc., 1999, Vol. 103, p. 87.
89. **B. Chen, J.I. Siepmann.** : J. Chem. Phys., 1998, Vol. 108, p. 3383.
90. **W. F. van Gunsteren, H. J. C. Berendsen.** : Journal of Computer-Aided Molecular Design, 1987, Vol. 1, p. 171.
91. **Kirkwood, J. G.** : J. Chem. Phys., 1935, Vol. 3, p. 300.
92. **A. Michael, Malcolm G., E. Forsythe, C. B. Moler.** *Computer Methods for Mathematical Computations.* Englewood Cliffs : Prentice-Hall, 1977.
93. **Simonson, T.** *Free Energy Calculations.* New York : Marcel Decker, 2001.
94. **G. H. Golub, J. H. Welsch.** : Mathematics of Computation, 1969, Vol. 23, pág. 221.
95. **Zwanzig, R. W. J.** : Chem. Phys., 1954, Vol. 22, pág. 1420.
96. **H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren.** *Molecular dynamics and protein structure.* western springs IL : Hermans J. Ed., 1985.

97. **N. Mizushima, D. Spellmeyer, S. Hirono, D. Pearlman.** : J. Biol. Chem., 1991, Vol. 266, p. 11801.
98. **K. Kuczera, J. Gao, B. Tidor, M. Karplus.** : Proc. Natl. Acad. Sci., 1990, Vol. 87, p. 8481.
99. **Ytreberg F.M., Swendsen R.H., Zuckerman D.M.** : J. Chem. Phys., 2006, Vol. 125, p. 184114.
100. **F.S. Lee, Z. T. Chu, A.J. Warshel.** : J. Comp. Chem., 1993, Vol. 14, p. 161.
101. **D. J. Tobias, C. L. Brooks.** : J. Chem. Phys., 1990, Vol. 92, p. 2582.
102. **P. Cieplak, P. Bash, U.C. Singh, P. A. Kollman.** : J. Am. Chem. Soc., 1987, Vol. 109, p. 6283.
103. **Reynolds, C.A.** : J. Am. Chem. Soc., 1990, Vol. 112, p. 7545.
104. **W.L. Jorgensen, J. M. Briggs, J.J. Gao.** : J. Am. Chem. Soc., 1987, Vol. 109, p. 6857.
105. **P. Cieplak, P. Kollman.** : J. Am. Chem. Soc., 1988, Vol. 110, p. 3734.
106. **Frisch, M. J., et al.** Gaussian 03, Revision C.02. Wallingford CT : Gaussian, Inc., 2004.
107. **H. Flyvbjerg, H.G. Petersen.** : J. Chem. Phys., 1989, Vol. 91, pág. 461.
108. **Salamin, E.** *Application of quaternions to computation with rotations.* Stanford : Stanford Univ., 1974.
109. **Verlet, L.** : Phys. Rev., 1967, Vol. 159, p. 98.
110. **J. W. Perram, S. W. De Leeuw, E. R. Smith.** : Proc. R. Soc. Lond. A., 1980, Vol. 373, p. 27.
111. **J. W. Perram, S. W. De Leeuw, E. R. Smith.** : Proc. R. Soc. Lond. A., 1980, Vol. 373, p. 57.
112. **J. W. Perram, S. W. De Leeuw, E. R. Smith.** : Proc. R. Soc. Lond. A., 1983, Vol. 388, p. 388.
113. **H. Flyvbjerg, H.G. Petersen.** : J. Chem. Phys., 1989, Vol. 91, p. 461.
114. **R.M. Glaeser, G. Hura, J.M. Sorensen, T.H. Gordon.** s.l. : J. Chem. Phys., 2000, Vol. 113, p. 9140.
115. **H. Zhao, D. Beysen, B. Guillot, M.C. Bellisent-Funel, T. Tassaing, Y. Guissani.** : J. Chem. Phys., 1997, Vol. 107, p. 2942.
116. **L. Haar, J.S. Gallagher, G.S. Kell.** *NBS/NRC Steam Tables.* Washington DC : Hemisphere, 1984.
117. **B. Guillot, Y. Guissani.** : J. Chem. Phys., 2001, Vol. 114, p. 6720.
118. **Soper, A.K.** : Chem. Phys., 1996, Vol. 202, p. 295.
119. **Soper, A.K.** : J. Chem. Phys., 1994, Vol. 101, p. 6888.
120. **Hitchings, G. H.** : J. Biol. Chem., 1941, Vol. 142, p. 43.
121. **Johns, C. O.** : J. Biol. Chem., 1911, Vol. 9, p. 67.
122. **Merck.** *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals.* NJ : Merck Co. Inc., 2006.
123. **A.Y. Toukmaji, J.A. Board Jr.** : Computer Physics Communications, 1996, Vol. 95, p. 73.
124. **C.J. Fennell, J.D. Gezelter.** : J. Chem. Phys., 2006, Vol. 124, p. 234104.
125. **S.R. Phillpot, D. Wolf, P. Keblinski, J. Eggebrecht.** : J. Chem. Phys., 1999, Vol. 110, p. 8255.
126. **Dreyfus M., Bensaude O., Dodin G., Dubois J. E.** s.l. : J. Am. Chem Soc., 1976, Vol. 93, pág. 6338.

127. **Kaito A., Hatano M., Ueda T., Shibuya S.** : Bull. Chem. Soc. Jpn., 1980, Vol. 53, pág. 3073.
128. **Lin J., Yu M., Peng S., Akiyama I., Li K., Lee L. K., LeBreton P. R.** : J. Phys. Chem., 1980, Vol. 84, pág. 1006.
129. **Szczepaniak K., Szczesniak M.** s.l. : J. Mol. Struct., 1987, Vol. 29, pág. 156.
130. **Szczepaniak K., Szczesniak M., Person W. B.** : Chem. Phys. Lett., 1988, Vol. 39, p. 153.
131. **Szczesniak M., Szczepaniak K., Kwiatkowski J. S., KuBulat K., Person W. B.** : J. Am. Chem. Soc., 1988, Vol. 110, pág. 8319.
132. **Nowak M. J., Lapinski L., Fulra J.** : Spectrochimica Acta, 1989, Vol. 45A, p. 229.
133. **Gould I. R., Burton N. A., Hall R. J., Hillier I. H.** : J. Mol. Struct. (THEOCHEM), 1995, Vol. 331, p. 147.
134. **Thewalt U., Bugg C. E., Marsh R. E.** : Acta Crystallogr., Sect B., 1971, Vol. 27, pág. 2358.
135. **Estrin D. A., Paglieri L., Coringiu G.** : J. Phys. Chem, 1994, Vol. 98, p. 5653.
136. **Stewart E. L., Foley C. K., Allinger N. L., Bowen J. P.** : J. Am. Chem. Soc., 1994, Vol. 116, pág. 7282.