

8
2 ej.

**UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO**

Unidad Académica de los Cielos
Profesional y de Posgrado
COLEGIO DE CIENCIAS Y HUMANIDADES



**EFFECTOS DE HIDRATACION:
REFINAMIENTO DE UN MODELO
ESTEREO DINAMICO PARA EL AGUA
LIQUIDA**

T E S I S
QUE PARA OBTENER EL GRADO DE
LICENCIADO EN INVESTIGACION
BIOMEDICA BASICA
P R E S E N T A:
CARMEN NINA PASTOR COLON

CUERNAVACA, MOR.

1992

TESIS CON
FALLA DE ORIGEN



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 está 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.

ABREVIATURAS USADAS EN ESTA TESIS

potenciales:

MCHO: Mobile Charges in Harmonic Oscillators (cargas móviles en osciladores armónicos)

MCHOv: Mobile Charges in Harmonic Oscillators; vibrational version (cargas móviles en osciladores armónicos; versión vibracional)

simulaciones numéricas:

MC: Monte Carlo

DM: Dinámica Molecular

cantidades conservadas en los distintos ensembles:

N : número de partículas

V : volumen

T : temperatura

p : presión

E : energía

H : entalpía

μ : potencial químico

TBC: Toroidal Boundary Conditions (condiciones toroidales a la frontera)

PBC: Periodic Boundary Conditions (condiciones periódicas a la frontera)

cálculos ab initio:

STO: Slater Type Orbitals (orbitales tipo Slater)

GTO: Gaussian Type Orbitals (orbitales tipo Gaussiano)

LCAO: Linear Combination of Atomic Orbitals (combinación lineal de orbitales atómicos)

HF: Hartree-Fock

SCF: Self Consistent Field (campo autoconsistente)

HOMO: Highest Occupied Molecular Orbital (orbital molecular ocupado con la energía más alta)

LUMO: Lowest Unoccupied Molecular Orbital (orbital molecular virtual con la energía más baja)

CI: Configuration Interaction (interacción de configuraciones)

CID: Configuration Interaction including Double excitations (interacción de configuraciones incluyendo diexcitaciones)

CISD: Configuration Interaction including Single and Double excitations (interacción de configuraciones incluyendo mono- y diexcitaciones)

RSPT: Rayleigh-Schrödinger Perturbation Theory (teoría de perturbaciones de Rayleigh -Schrödinger)

MBPT: Many Body Perturbation Theory (teoría de perturbaciones de muchos cuerpos)

MPn: Møller Plesset a orden n

MP4SDQ: Møller Plesset 4th order including Single, Double and Quadruple excitations (Møller Plesset a 4^o orden incluyendo mono-, di- y cuadiexcitaciones)

CC: *Coupled Cluster*

CCD: Coupled Cluster including Double excitations (*Coupled Cluster* incluyendo diexcitaciones)

MRSCF: MultiReferential Self Consistent Field (campo autoconsistente multirreferencial)

MRCI: MultiReference Configuration Interaction (interacción de configuraciones multirreferencial)

BSSE: Basis Set Superposition Error (error por superposición de base)

CP: CounterPoise

INDICE

Resumen	1
Capítulo 1: La Hidratación Molecular	2
1.1 Introducción	2
1.2 La bioquímica y la biofísica del agua	4
1.2.1 Los efectos del agua sobre los solutos; el solvente universal	4
1.2.2 Los efectos de los solutos en las propiedades del agua	8
1.3 Las simulaciones numéricas al rescate	10
Capítulo 2: Metodología	12
2.1 Esquema de Simulación Global	12
2.2 Cálculos <i>ab initio</i>	14
2.2.1 Aproximaciones para resolver la ecuación de Schrödinger	15
2.2.1.1 Consideraciones generales	15
2.2.1.2 Correlación electrónica	18
2.2.1.2.1 Interacción de configuraciones completa (<i>full CI</i>)	20
2.2.1.2.2 Teoría de Perturbaciones de Møller-Plesset	21
2.2.1.2.3 <i>Coupled Cluster</i>	22
2.2.1.2.4 Cálculos multirreferenciales	22
2.2.2 Bases moleculares	23
2.2.2.1 Orbitales tipo Slater	23
2.2.2.2 Orbitales tipo Gaussiano	24
2.2.2.3 Geminales	25
2.2.3 Cálculo de energías de interacción	25
2.2.4 Cálculo de propiedades de los complejos de van der Waals	26
2.3 La no aditividad en la interacción molecular	27
2.4 Simulaciones numéricas	29
2.4.1 Consideraciones generales	29
2.4.1.1 La validez de las simulaciones clásicas	30
2.4.1.2 El ensemble de simulación	31
2.4.1.3 El número de partículas simuladas	32
2.4.1.4 Las fronteras del sistema y el corte de los potenciales de interacción	32
2.4.1.5 La longitud de la simulación	35
2.4.2 Monte Carlo	36
2.4.2.1 Descripción	36
2.4.2.2 Dificultades de aplicación	38
2.4.2.3 Información obtenida	38
2.4.3 Dinámica Molecular	38
2.4.3.1 Descripción	38
2.4.3.2 Dificultades de aplicación	40
2.4.3.3 Información que proporciona	41

Capítulo 3: Resultados y Discusión	42
3.1 Obtención de un potencial intramolecular	42
3.2 Ajuste del potencial vibracional	43
3.3 Desarrollo del programa de dinámica molecular	55
3.4 Simulación por dinámica molecular	59
3.5 La no aditividad de MCHOv	62
3.6 Curva de disociación del enlace O-H donador	72
Capítulo 4: Conclusiones y Perspectivas	78
4.1 Conclusiones	78
4.2 Perspectivas	79
Referencias	81
Apéndice A: Bases utilizadas	91
Apéndice B: Coordenadas de los tetrámeros y del dímero optimizado	94
Apéndice C: Listado del programa DNM	06

RESUMEN

Mientras que el papel del agua como reactivo en la bioquímica es claro, su papel como solvente, es decir, la biofísica del agua, todavía presenta muchas interrogantes. El método de Simulación Global^{1,2} es una de las maneras en las que se puede obtener información molecular detallada del comportamiento del agua en la célula. Un prerrequisito para entender la interacción del agua con diversos solutos es entender al agua pura.

Presentamos un ejemplo concreto de la aplicación de esta metodología al estudio del agua líquida. El trabajo consistió en el refinamiento de un potencial intermolecular analítico (MCHO³) mediante la inclusión de grados de libertad intramoleculares. Para este efecto calculamos una superficie vibracional *ab initio* para la molécula de agua, e hicimos un ajuste simultáneo a las superficies *ab initio* de uno, dos, tres y cuatro cuerpos (donde cuerpo se refiere a una molécula de agua). Como nos interesan las propiedades dinámicas y estructurales del agua, desarrollamos un programa de simulación numérica por dinámica molecular. Hicimos dos simulaciones numéricas, una por dinámica molecular y la otra por Monte Carlo. Los resultados de las simulaciones indicaron que el nuevo potencial (MCHOv) exagera la energía potencial en $\approx 30\%$ respecto al valor experimental. Siguiendo el esquema de simulación global, regresamos a analizar el comportamiento de la no aditividad de tres y cuatro cuerpos en dos tetrámeros de moléculas de agua relajados con MCHOv y repetimos el análisis con cálculos *ab initio* a nivel HF-SCF de los mismos tetrámeros. MCHOv, en su versión actual, no es capaz de reproducir los valores *ab initio* ni de las energías de interacción ni de las no aditividades de tres y cuatro cuerpos de agregados de moléculas deformadas. Encontramos que la no aditividad de tres cuerpos es la encargada de estabilizar estos agregados y la no aditividad de cuatro cuerpos es entre el 10 y el 15% de la energía total de interacción. A nivel HF-SCF los dímeros de los tetrámeros relajados no son estables. Hicimos un análisis detallado a nivel CI-MP2 del enlace que sufre el cambio mayor al formarse un dímero de agua, el enlace OH donador. Encontramos que a nivel HF-SCF este enlace prácticamente no se estira, mientras que a nivel CI-MP2 este enlace se estira 0.04 ua respecto a la longitud del enlace del monómero optimizado para la base. Como punto de comparación, en una fase condensada como el hielo Ih, la distancia OH promedio es 0.06 ua⁴ mayor que la longitud del enlace OH del monómero de agua en fase gaseosa. Este valor indica la importancia de los efectos no aditivos en la estabilización de la deformación molecular.

Las conclusiones de este trabajo son que el esquema de simulación global es adecuado para estudiar sistemas moleculares, ya que provee explicaciones para los aciertos y fracasos de los potenciales utilizados y un mecanismo para corregirlos. Este esquema ha mostrado que puede ayudar a entender la fisicoquímica de un sistema tan complejo como el agua líquida. Los resultados obtenidos muestran que los efectos no aditivos parecen ser los responsables de estabilizar la deformación del monómero de agua en sus agregados, y que, como los modos de vibración inter e intramoleculares están acoplados, un potencial de interacción analítico no aditivo y vibracional para el agua debe ajustarse a superficies *ab initio* que incluyan agregados de moléculas deformadas.

CAPÍTULO 1: LA HIDRATACIÓN MOLECULAR

1.1 INTRODUCCIÓN

El agua existe en nuestro planeta desde mucho antes de que apareciera la vida, y ésta se desarrolló en un ambiente predominantemente acuoso. Entre las consecuencias propuestas de este desarrollo acuoso están que tanto la forma como el tamaño de las células están ligados al comportamiento del agua. La ósmosis, que es el desplazamiento de agua a través de una membrana permeable para equilibrar el potencial químico del agua, es un agente morfogenético¹, y el tamaño de las células parece tener un límite superior impuesto por las constantes de difusión de los metabolitos y enzimas en el agua².

Aun en las formas de vida terrestres, el 70% del volumen de sus células es agua, lo cual indica lo fundamental que es esta sustancia para la vida. El estar hechos básicamente de agua hace a los seres vivos especialmente sensibles a las propiedades físicas y químicas de este compuesto. A su vez, estas propiedades son sensibles a los efectos perturbativos causados por cambios en temperatura, presión, y, muy notablemente, por los diversos solutos que se encuentran en una célula.

La imagen que tenemos del agua a nivel celular ha ido cambiando a medida que ha ido mejorando la resolución espacial y temporal de las técnicas utilizadas en biología celular, biología molecular y bioquímica. Los avances en la microscopía han hecho ver que el interior celular, lejos de ser un caldo amorfo, tiene un alto grado de estructuración. Las estructuras intracelulares presentan superficies, y éstas perturban la estructura del agua líquida³. Mientras que esto no está en discusión, lo que permanece como un área activa de investigación es la extensión de la alteración de la estructura del agua por la presencia de superficies, ya sean cargadas, polares o neutras. Por otro lado, la difracción de rayos X y de neutrones, y la resonancia magnética nuclear (RMN) han contribuido a modificar el concepto del agua del de un solvente inerte al de un solvente que participa en las estructuras macromoleculares. Avances recientes en la resolución espacial que se puede alcanzar con estas técnicas ($\approx 1 \text{ \AA}$) han permitido, por ejemplo, asignar las posiciones de las moléculas de agua en la primera esfera de hidratación de proteínas pequeñas, pudiéndose estudiar los diferentes patrones de hidratación de grupos con distinta carga, polares e hidrofóbicos⁴⁻⁶. Por RMN y por estudios cristalográficos se ha confirmado la existencia de moléculas de agua en el interior de algunas proteínas⁴⁻⁶, y también se han hecho estudios sobre la dinámica de las moléculas de agua ligadas a la superficie de macromoléculas^{7,8}. Las moléculas de agua que se localizan en el interior de las proteínas tienen un papel ya sea funcional o estructural, satisfaciendo puentes de H o contactos de van der Waals en general (las cavidades son menos favorables energéticamente que un par de puentes de H no satisfechos).

Recientemente se han empezado a hacer estudios de sistemas biológicos en micelas invertidas y en solventes orgánicos, sistemas que van desde enzimas hasta esporas de bacterias en germinación⁹⁻¹³. En estos estudios el agua es un reactivo limitante. Uno de los resultados de estos estudios es que se requiere agua para que el sistema muestre ac-

tividad, o dicho de otra manera, que un sistema anhidro es un sistema muerto. Finney y Poole¹⁴ llegaron a la misma conclusión hidratando paulatinamente polvos anhidros de lisozima, encontrando que se necesita una cantidad de agua cercana a la necesaria para formar una monocapa de agua alrededor de la enzima para que ésta última tenga actividad; esta monocapa no parece estar distribuida de manera uniforme sobre la superficie de la proteína. El análogo fisiológico de una situación limitante de agua es el estrés hídrico; una de las respuestas a este tipo de estrés es sintetizar carbohidratos que suplen la función del agua, haciendo los puentes de H indispensables para mantener la integridad de las macromoléculas¹⁶. Otro resultado interesante es que el agua es capaz de alterar las constantes de velocidad de diversas reacciones con respecto a lo que sucedería en fase gaseosa, haciendo mayores o menores las barreras de activación de las reacciones. A esto se le llama solvactivación. Además, la energía libre de una reacción puede deberse en parte a la diferencia en la energía libre de solvatación de reactivos y productos. Un ejemplo dramático del efecto de los solventes sobre las constantes de equilibrio es el trabajo del grupo de Klibanov¹⁶⁻¹⁸, quienes, escogiendo con cuidado los solventes en los que solubilizan a varias enzimas, son capaces de cambiar la estereoespecificidad de algunas reacciones respecto a lo que sucedería en fase acuosa; su explicación es que cambia el coeficiente de partición entre el solvente y el sitio activo de la enzima, haciendo que el sustrato se acomode de manera distinta en el sitio activo.

Las simulaciones numéricas por Monte Carlo o dinámica molecular de biomoléculas son una alternativa muy llamativa para explicar resultados experimentales como los descritos arriba. Mientras que los experimentos necesariamente proporcionan información sobre promedios de las propiedades de interés, las simulaciones permiten obtener información con detalle molecular individual, y la disección de las interacciones más relevantes en el sistema; a esto le llamamos explicar. Además del poder explicativo, las simulaciones permiten "hacer experimentos" que no son factibles o son muy costosos en el laboratorio; esta característica hace a las simulaciones muy útiles, por ejemplo, en el diseño de fármacos.

Los resultados de una simulación son tan confiables como el modelo físico y matemático que se resuelve. El modelo físico del sistema se traduce en el potencial de interacción entre los distintos componentes del sistema, y mientras más realista sea este potencial, el sistema simulado se parecerá más al sistema real. En la práctica, los potenciales de interacción son idealizaciones de la interacción real, que en principio está dada por la mecánica cuántica. Las primeras simulaciones de biomoléculas utilizaban modelos muy crudos, tanto para la molécula en sí como para el agua, la cual se modelaba como un medio continuo con una constante dieléctrica de ≈ 80 . Los aumentos en capacidad de memoria y de velocidad de las computadoras han permitido el refinamiento de los modelos moleculares, de tal suerte que ahora es posible simular macromoléculas y agregados moleculares (proteínas, ADN, ARNt, membranas de fosfatidilcolina y complejos enzima - sustrato, proteína - ácido nucleico, etc.) con moléculas de agua explícitas¹⁹. Sin embargo, las simulaciones que incluyen todos los átomos del sistema en estudio son mucho más lentas que aquellas en las que se omiten átomos ligeros (H) y/o se usa un solvente continuo; por esto no se ha extinguido el uso de los modelos menos detallados. Los potenciales para el agua que se usan rutinariamente en las

simulaciones de biomoléculas son los llamados SPC²⁰ y TIP4P²¹ (o su versión simplificada, TIP3P²¹). Estos potenciales son empíricos, efectivos (es decir, representaciones aditivas de un potencial de muchos cuerpos), aditivos por pares, no incluyen polarización y no incluyen grados de libertad intramolecular. Una revisión reciente del desempeño de estos potenciales en el refinamiento de estructuras cristalográficas de proteínas indica que no son suficientemente refinados para proporcionar la resolución estructural y energética que se requiere, y pueden dar información cualitativamente dispar o inclusive incorrecta^{4,23}. Esto establece una necesidad clara de un potencial confiable y económico, en términos computacionales, para el agua, que incluya polarizabilidad, efectos no aditivos y flexibilidad molecular. Los potenciales empíricos están parametrizados con datos de agua de bulto, y esto pudiera no ser adecuado para la simulación de sistemas interfaciales, como son las superficies de las macromoléculas. Los potenciales *ab initio*, a pesar de ser más costoso su desarrollo, tienen la enorme ventaja de poderse utilizar para simular agua en cualquier fase o situación.

Habiendo hecho mención de algunos avances importantes en el estudio del agua en la biología, a continuación presentamos una descripción a nivel molecular del papel que el agua desempeña en la biología celular.

1.2 LA BIOQUÍMICA Y LA BIOFÍSICA DEL AGUA

La bioquímica del agua está básicamente descrita². La fotosíntesis y la respiración son la alternativa biológica para romper y sintetizar la molécula de agua respectivamente, aprovechando la energía de formación de la molécula de agua para sintetizar el material celular. Además, el agua participa como reactivo o como producto en numerosas reacciones del metabolismo celular: condensaciones, hidrólisis, óxido - reducciones, etc. Esto no quiere decir que los mecanismos de reacción estén descritos en detalle. Algunos temas de investigación activa en la actualidad son el mecanismo de lisis de la molécula de agua por el fotosistema II y el mecanismo de oxidación de sustancias tóxicas por catalasas, peroxidasas y citocromos P450, reacciones que producen moléculas de agua, H₂O₂ y superóxido²⁴.

La biofísica del agua no goza del mismo nivel de comprensión. Lo que se quiere entender es, por un lado, el efecto del agua en procesos biomoleculares tales como el plegamiento de las proteínas, el ligado de sustratos y la asociación molecular en general; y por el otro, el efecto de los solutos sobre el agua, o sea, la existencia del agua ligada a superficies moleculares, su extensión espacial y la magnitud de los cambios de las propiedades del agua ligada con respecto al agua de bulto.

1.2.1 El efecto del agua sobre los solutos; el solvente universal

La hidratación es un fenómeno complejo para el que todavía no hay una explicación satisfactoria a nivel molecular para los sistemas más sencillos como las disoluciones diluidas de iones o de gases nobles. La solubilidad mutua está determinada por la interacción

molecular entre solvente y soluto, interacción en la que intervienen factores energéticos y geométricos. Un compuesto *A* se disuelve en un solvente *B* sólo cuando las fuerzas intermoleculares de atracción K_{AA} y K_{BB} para los compuestos puros pueden ser vencidas por las fuerzas K_{AB} en solución.

La existencia de las células se puede explicar en base al comportamiento del agua como buen y mal solvente. Las propiedades de muchas estructuras biológicas como la doble hélice de ADN, las proteínas globulares y ciertos geles depende de un delicado balance entre las características de buen y mal solvente del agua. Este balance confiere a las macromoléculas biológicas una estabilidad marginal, la cual está muy relacionada con la función biológica de estas moléculas.

El agua es un buen solvente de iones y de moléculas polares, como son muchos meta-bolitos, nutrientes y productos de deshecho. Esta capacidad de disolver una gran variedad de sustancias es la base del transporte de energía e información a través de la célula, y de la circulación en organismos más complejos. El coeficiente de partición entre el agua y las membranas y los receptores gobierna la distribución final de los diversos solutos que viajan disueltos en el agua. Además de su función dispersora, por su alta constante dieléctrica, el agua participa en la estabilización de moléculas poli-iónicas, apantallando eficientemente las cargas de proteínas y ácidos nucleicos.

Las interacciones que predominan en este caso son las electrostáticas (por ejemplo, ion - dipolo, dipolo - dipolo y ion - cuadrupolo) y/o la formación de puentes de H. En términos termodinámicos, la ΔG de solvatación es negativa, y la contribución entálpica a la solvatación puede ser positiva o negativa. Aunque la primera esfera de hidratación de los iones en general está formada por moléculas de agua con movilidad muy restringida, la contribución entrópica al proceso de solvatación parece ser positiva. La primera esfera de hidratación puede ser una manera muy selectiva de capturar iones. Por ejemplo, en fase gaseosa, un pentámero de moléculas de agua en configuración pentagonal captura iones de Na, mientras que con un hexágono de moléculas de agua se puede capturar un ion de K^{25} . La diferencia en afinidad se debe a las interacciones carga - dipolo y carga - cuadrupolo y a efectos cuánticos. Hay un modelo²⁶ que propone que, en solución acuosa, alrededor de un ion existen tres zonas: la esfera de hidratación, el seno o bulto del solvente y una zona intermedia en la que el desorden molecular es muy marcado. Se postula que el desorden de esta zona de transición compensa el orden impuesto en la esfera de hidratación; faltan por determinarse la extensión y estructura molecular de estas zonas.

Otro ejemplo del comportamiento del agua como un buen solvente es la formación de geles mucosos, que funcionan como protectores, lubricantes y barreras selectivas de permeabilidad. Estos están formados por soluciones de glicoproteínas de alto peso molecular ($\approx 10^6$ d). El 87% del peso de estas moléculas es carbohidrato, que por tener muchos grupos OH interactúa fuerte y favorablemente con el agua. En una solución diluida coloidal casi no hay interacción entre dos moléculas de glicoproteína, las moléculas están muy extendidas (el 98% del volumen dentro de la esfera de influencia de la molécula es agua), y casi no hay puentes de H intramoleculares. De toda el agua accesible, sólo el 1% está firmemente ligada; el enmascaramiento de los grupos reactivos con poca agua permite a

la molécula flexible expandirse de manera que maximiza su entropía conformacional. Si se baja el contenido de agua aumentan las interacciones intermoleculares, con la consecuente formación de un gel. En estas condiciones las moléculas de glicoproteína tienen contactos intramoleculares, por lo que se dice que el agua atrapada es intramolecular. Si se baja más el contenido de agua, ésta deja de competir de manera efectiva por la formación de puentes de H entre las moléculas y éstas se agregan. Los cambios en la cantidad de agua accesible tienen repercusiones fisiológicas importantes, por ejemplo en el curso de infecciones, modificando las propiedades reológicas de las soluciones mucosas, haciendo más difícil el transporte de patógenos hacia afuera del organismo. Otro ejemplo interesante lo constituyen las transiciones sol - gel que se dan en el movimiento por pseudópodos de diversos protozoarios.

Al comportamiento del agua como un mal solvente se le llama efecto hidrofóbico y se caracteriza por una ΔG de solvatación positiva. Datos calorimétricos de soluciones acuosas de hidrocarburos y de la tensión superficial entre el agua y varias sustancias hidrofóbicas hacen pensar que la contribución entálpica al proceso es favorable y casi cero, y que la responsable de la energía libre positiva es el cambio en la entropía²⁷. Este cambio se interpreta como un ordenamiento del agua alrededor del soluto y/o como un cambio en la dinámica del solvente. El efecto principal de incluir una molécula no polar en agua es crear una cavidad en el líquido por la reorientación de las moléculas de agua que tratan de maximizar el número de puentes de H formados. El resultado es una red (clatrato) que envuelve al soluto no polar; esta red tiene serias restricciones geométricas. En este caso, la interacción soluto - solvente no es suficientemente fuerte para vencer la interacción solvente - solvente; la interacción soluto - soluto en general es muy débil.

Hay dos modelos para explicar por qué es favorable la asociación de solutos no polares en agua. Uno dice que, como el cambio en energía libre para la interacción hidrofóbica es función de la superficie de los grupos expuestos al agua, y cuando se juntan dos grupos la superficie expuesta al agua se reduce, al reducir la superficie de contacto con el agua hay menos moléculas de agua restringidas, con lo que $\Delta S > 0$. El otro dice que al agregar solutos hidrofóbicos se mejora la esteoquímica de las redes de agua de manera que se pueden hacer más puentes de H, haciendo más favorable ΔH . Ambas explicaciones son consistentes con la termodinámica, y no tienen por qué ser excluyentes.

El agua participa como un mal solvente en el ensamblaje y estabilización de componentes estructurales, tales como membranas, transductores de energía, cápsidas virales, etc., todos formados por moléculas anfífilas. Las propiedades del agua que son necesarias para promover el autoensamblaje de estas moléculas son su alta polaridad y su capacidad para formar redes tridimensionales de puentes de H²⁸.

Además de determinar la estructura, la interacción del agua con las macromoléculas biológicas determina su función. Por ejemplo, en cristales de hemoglobina hay varios picos de densidad electrónica a distancia de puente de H de grupos polares, y hay varios

contactos mediados por moléculas de agua que unen a las subunidades. Se piensa que estas moléculas de agua toman parte en la transición alostérica entre la forma activa y la inactiva²⁹. Además de funciones muy locales como la de las moléculas de agua en interfases moleculares, también hay efectos causados por propiedades colectivas del agua. La viscosidad del solvente, a través del acoplamiento a las fluctuaciones conformacionales, puede ser un factor determinante en las actividades moleculares en sistemas biológicos.

Las proteínas globulares parecieran haber evolucionado haciendo el mejor uso posible de las fluctuaciones inducidas por el agua. La hidratación tiene una gran importancia para la dinámica enzimática porque, de acuerdo al teorema de fluctuación - disipación, la viscosidad del solvente es la propiedad cinética que gobierna el movimiento Browniano³⁰. Según este modelo, los dominios de las proteínas deben comportarse como partículas Brownianas, y el agua es el medio que provee las fuerzas azarosas que controlan estos dominios en su función. Por ejemplo, se ha visto que, para algunas enzimas³¹, las fluctuaciones estructurales locales dependientes de la viscosidad controlan la transición sobre la barrera de reacción. El supuesto acoplamiento proteína - solvente no es fácil de explicar dados los pequeños tiempos de residencia de las moléculas de agua en la superficie de las proteínas (10^{-10} s)⁴, así que se ha propuesto un acoplamiento dinámico entre el solvente y el interior de la proteína, ya que la transferencia de momento del solvente al interior de la proteína es muy eficiente³².

El cambio conformacional más dramático en una proteína es el proceso de desnaturalización. Además de la desestabilización del estado nativo por la temperatura y pH's extremos, hay una colección de sustancias que desnaturalizan proteínas. Se han propuesto dos mecanismos de acción de los desnaturalizantes:

- la formación de un complejo desnaturalizante - proteína. Mientras que experimentalmente se encuentra que todas las moléculas que protegen al estado nativo (cosolventes como el glicerol, betaína, etc.) no se asocian con las proteínas, los desnaturalizantes sí muestran asociación³³: la explicación propuesta es que la urea y la guanidina se asocian a los grupos peptídicos y los alcoholes a los grupos no polares, solvatándolos y reduciendo la presión ejercida por el efecto hidrofóbico. Como hay más posibilidad de ligado en el estado desdoblado que en el nativo, el equilibrio se desplaza hacia la forma desdoblada. A pesar de los datos presentados³³, otros autores consideran que no está claro si la asociación de las moléculas a la proteína es un prerrequisito para que haya desnaturalización³⁴. El mecanismo de protección del estado nativo por el glicerol y las betainas no está descrito.
- la alteración de la estructura del agua, que cambia la interacción agua - proteína. Los desnaturalizantes disminuyen la cantidad de agua estructurada (de aquí que se les llame también caotropos) y los estabilizadores ordenan o estructuran al agua. La habilidad protectora de estas moléculas parece estar ligada a la presencia de varios grupos metilo, que tienen la propiedad de estructurar agua a su alrededor³⁴. El corolario de esta propuesta es que el único agente desnaturalizante o estabilizador es el agua misma. Otra explicación para la eficiencia de los cosolventes como el glicerol y la sacarosa en la estabilización de las proteínas es que aumentan la viscosidad del

medio, atenuando las fluctuaciones conformacionales de la proteína y reduciendo la probabilidad de exceso de solvatación de grupos peptídicos que desestabiliza al estado nativo³².

Algunas transiciones estructurales de los ácidos nucleicos también están ligadas al nivel de hidratación. Por ejemplo, en el ADN en forma B, en el surco mayor el agua hace puentes de H con los fosfatos, y en el surco menor los hace con los grupos NH y CO de las bases. Como las moléculas de agua están unidas menos fuertemente a éstas últimas, al deshidratar al ADN se disocian, el surco se contrae, y se obtiene la forma A²⁹.

Los ejemplos que hemos visto hasta ahora se han restringido al efecto del agua sobre macromoléculas aisladas o en grupos pequeños (oligómeros como la hemoglobina). El agua también afecta a agregados de un número grande de biomoléculas, como las membranas biológicas, que cambian de fase al cambiar la cantidad de agua accesible³⁵. Para tener una bicapa parece ser que se requieren al menos cinco capas de agua en su superficie; de haber menos, ocurre una transición hacia la fase hexagonal. Estos cambios inducidos por la hidratación pudieran ser relevantes en los procesos de endocitosis y exocitosis, en los que se fusionan y se disocian secciones de membrana.

1.2.2 Los efectos de los solutos en las propiedades del agua

En la sección anterior vimos el efecto del agua sobre los solutos; en ésta veremos lo que le pasa al agua al ponerse en contacto con los solutos. La presencia de moléculas, iones, microorganismos y partículas dispersas en agua líquida cambia las interacciones entre las moléculas de agua. Este cambio puede a su vez alterar la interacción soluto - soluto.

Los sistemas más estudiados son las soluciones iónicas, debido a su relevancia tecnológica. En éstas cambia el arreglo de las moléculas de agua respecto al agua pura al formarse las esferas de hidratación. Como el agua tiene una constante dieléctrica alta (de 78), se espera que sea capaz de apantallar la carga de los iones rápidamente, de tal forma que el efecto de la carga del ion no sea perceptible a grandes distancias en términos moleculares.

En una solución 0.15M de NaCl (esto corresponde a la concentración en el plasma sanguíneo o de K intracelular), la distancia entre Na y Cl es de 1.9 nm en promedio. Como los radios de Na y Cl hidratados son de ≈ 0.6 nm, el espacio libre entre este par de iones es apenas suficiente para acomodar entre ellos un pentámero de agua, cuyo diámetro es de 0.5 nm. Por lo tanto, se puede concluir que en esta solución la estructura del agua está fuertemente alterada. Como manifestaciones macroscópicas de esta alteración se tiene que la temperatura del punto de fusión baja, y la correspondiente al punto de ebullición aumenta, ambas respecto a los valores para el agua pura.

Un sistema poli-iónico muy importante lo constituyen las proteínas. Para explicar datos de la dinámica del agua adsorbida a la superficie de algunas proteínas se ha propuesto⁷ que hay al menos dos tiempos de relajación. El agua ligada puede girar rápidamente sobre el eje correspondiente al puente de H, pero sobre los otros dos ejes se mueve mucho más lento; la anisotropía rotacional se debe a la diferente energía de ligado a los diversos grupos

funcionales en la superficie. Los tiempos de relajación son lentos, del orden de 10^{-6} s. Otros estudios⁸ de RMN de proteínas hidratadas dicen que no hay más de dos capas de agua estructurada, y que la vida media de un agua ligada es de ≈ 10 ns.

El siguiente nivel en complejidad podría estar dado por sistemas superficiales mesoscópicos. La extensión de los efectos de largo alcance en sistemas acuosos interfaciales es un tema sobre el que las opiniones han oscilado³⁶ entre proponer, o que prácticamente no existen (hasta 5 diámetros moleculares cuando mucho), o que tiene un alcance tan grande (30 a 200 diámetros o más) que toda el agua de una célula se encontraría en un estado semejante al del agua de cristalización.

Uno de los métodos experimentales usados para estudiar sistemas interfaciales es la RMN. En este tipo de estudios se encuentra que el desplazamiento químico (*chemical shift*) del agua celular es similar al del agua de bulto³⁷. Sin embargo, el coeficiente de autodifusión del agua está reducido a la mitad. Según unos autores³⁷, el orden persiste por 1 nm (3 o 4 diámetros moleculares), mientras que otros afirman que las perturbaciones significativas ocurren sólo en las primeras dos capas de hidratación⁸. La única conclusión "universal" es que en los datos experimentales de RMN hay algo que se puede llamar "agua ligada" y que tiene propiedades dinámicas distintas a las del agua de bulto. Experimentalmente, la diferencia en la entalpía de vaporización entre el agua de bulto y la vecinal o ligada es del orden de 0.1 kcal/mol o 0.01 kcal/mol³⁸.

Para dar una idea de la importancia de estos efectos en una célula basta hacer un cálculo aproximado de la cantidad de superficie que se encuentra en una célula eucariótica³⁹. La superficie de membrana de un hepatocito es de $\approx 100,000\mu^2$; si se cubre con una monocapa de agua, captura el 1% del agua intracelular. Como en general se supone que se necesita más de una capa para mantener la integridad de la membrana, se estima que se necesita entre el 5 y el 10% del agua intracelular para este propósito. La superficie del citoesqueleto es de $\approx 500,000\mu^2$, que con 5 capas de hidratación retiraría el 25% del agua intracelular. Todavía faltan por considerarse las esferas de hidratación de los iones, metabolitos, enzimas, y ácidos nucleicos. Pareciera ser que en realidad no queda mucha agua libre, y esto es importante si recordamos que la función de las macromoléculas está ligada a las propiedades estructurales y dinámicas del agua, mismas que son distintas dependiendo del entorno. Para saber qué tan ciertos son estos números se necesita saber la extensión espacial del cambio y cuánto cambia la estructura del agua líquida en la vecindad de una superficie o de moléculas en general, y hemos visto que los resultados experimentales no son sencillos de interpretar dadas las discrepancias en la literatura sobre este tema.

El agua con propiedades distintas a las del bulto es importante al menos en dos aspectos. El más sencillo es que puede llegar un momento en el que no hayan moléculas de agua suficientes para hidratar a todos los componentes de una solución y se precipiten los que tienen interacciones menos fuertes con el agua. Este es un fenómeno muy conocido y poco entendido en bioquímica y de hecho se usa para purificar proteínas². El otro aspecto lo constituye la fuerza de hidratación⁴⁰, que fue propuesta para entender la repulsión entre superficies, además de las interacciones electrostáticas y de van der Waals. Esta fuerza es

repulsiva para casi todas las superficies estudiadas en el intervalo de distancias de 15 a 25 Å y cae exponencialmente. La repulsión por hidratación parece ocurrir en todas las superficies que interactúan a distancias en las que las perturbaciones causadas al solvente por ellas mismas se sobrelapan. Un mecanismo molecular propuesto⁴¹ para esta fuerza es que, dado que las moléculas de agua no son esféricamente simétricas, al formar una superficie adquieren una orientación preferencial, y esto genera un potencial de superficie. Este potencial se puede medir con microscopios de fuerzas atómicas⁴². Los cambios en la estructura del agua se deben a las cargas de la superficie y a los dipolos que inducen una polarización eléctrica local. Como esta fuerza es casi ubicua, su papel en la asociación molecular no debe ser despreciable. De hecho, se supone que es responsable de impedir el contacto y la fusión de las bicapas⁴³ (el calcio favorece la fusión de membranas provocando una deshidratación local).

1.3 LAS SIMULACIONES NUMÉRICAS AL RESCATE

Gran parte de la información del apartado anterior tiene un carácter hipotético, y esto se debe a que los experimentos son complicados y la interpretación de los datos experimentales también es compleja. Parte de la complicación estriba en que los experimentos dan información promediada, y extraer información sobre cada componente del sistema no es sencillo. Algunos puntos donde todavía hay muchas dudas son:

- la estructura del agua alrededor de los iones: el número y disposición de las moléculas de agua que forman la esfera de hidratación y la transición a la estructura del líquido sin perturbar.
- el mecanismo de agregación de moléculas no polares y anfífilas en el agua.
- el acoplamiento dinámico entre el agua y las biomoléculas.
- el efecto de los desnaturalizantes y estabilizadores de proteínas sobre el agua.
- la extensión y la magnitud de los cambios inducidos por superficies sobre el agua.

Una característica común a estos cinco puntos es que estarían resueltos si uno tuviera acceso a información con resolución molecular, y éste es precisamente el tipo de información que proporcionan las simulaciones numéricas. Por supuesto, realizar simulaciones tampoco es trivial. Los sistemas que se pueden tratar de este modo están limitados por el tamaño del sistema (el número de partículas), el alcance de las interacciones y el tiempo característico del proceso en estudio. A pesar de estas limitaciones, el aumento en la capacidad de memoria y de velocidad de las computadoras, los avances en el desarrollo de algoritmos de muestreo y el mejoramiento progresivo de los potenciales de interacción, han permitido el estudio de sistemas moleculares de interés biológico de creciente complejidad⁴⁴⁻⁴⁶.

El detalle más importante de cualquier simulación numérica es el potencial de interacción entre las partículas del sistema. El potencial establece el puente entre el compor-

tamiento microscópico y las propiedades termodinámicas, que son lo que se puede medir, a través de la función de partición. Como discutimos en la introducción, los potenciales para la interacción entre moléculas de agua que se usan comúnmente en simulaciones de sistemas biomoleculares presentan algunos problemas de consistencia y de precisión. En vista de que una parte importante de la estructura y la función de las moléculas relacionadas directamente con la vida depende de su interacción con el agua, un requisito indispensable en las simulaciones de estos sistemas es contar con un potencial para la interacción agua - agua y agua - biomolécula lo más realista posible.

El trabajo de esta tesis está encaminado en este sentido, en particular al desarrollo de un potencial para la interacción agua - agua a partir de primeros principios. Actualmente existe una gran colección de potenciales para describir al agua líquida, pero ninguno es completamente satisfactorio. En particular, las propiedades mecánicas del agua son muy difíciles de predecir si no se ha ajustado el potencial para que las reproduzca (por ejemplo, la presión). El punto de partida para este trabajo es un potencial⁴⁷ *ab initio* para el agua líquida, bajo la aproximación de molécula rígida, que incluye polarizabilidad y efectos no aditivos de muchos cuerpos. El trabajo consistió en el refinamiento de este potencial para incluir grados de libertad intramoleculares, el desarrollo de un programa de simulación numérica por dinámica molecular y la prueba del potencial en simulaciones numéricas de agua líquida. Esta tesis muestra una aplicación del esquema de Simulación Global propuesto por Clementi^{48,49} a un sistema de interés bioquímico .

CAPÍTULO 2: METODOLOGÍA

2.1 ESQUEMA DE SIMULACIÓN GLOBAL

La solución correcta a problemas de escala molecular está dada por la ecuación de Schrödinger. Esto se debe a que el ámbito molecular es territorio de la mecánica cuántica. Sin embargo, esta ecuación (su parte independiente del tiempo) sólo se puede resolver exactamente para sistemas de dos cuerpos, tales como el átomo de H. Existe una serie de aproximaciones para resolver sistemas de muchos electrones, pero a medida que aumenta el número de electrones en el sistema, el tiempo de cómputo crece muy rápido, de manera que los cálculos se vuelven prohibitivamente caros si uno desea una precisión alta. Lo que sucede en la práctica es que, a medida que crece el sistema, se hacen aproximaciones más gruesas y menos fundamentadas.

Además de no ser prácticos los cálculos *ab initio* de sistemas moleculares grandes, los cálculos cuánticos se hacen en el vacío (a 0K) y no se incluye el tiempo. Si uno quiere entender procesos bioquímicos, ésta no es una buena simulación. Lo que se necesita es una conexión operacional (la formal existe desde la década de los 30's) de la mecánica cuántica a la mecánica estadística, de ésta a la termodinámica y de ésta finalmente a la mecánica de fluidos.

Clementi^{1,2} propuso un método general para simulaciones de sistemas químicos complejos. Lo que hace el método es construir un sistema macroscópico a partir de un sistema cuántico en etapas. En cada etapa se aplica la física relevante para el tipo de partículas en estudio, y el resultado de cada paso es el punto de partida para la etapa posterior:

PASO 1: QUÍMICA CUÁNTICA

Se hacen cálculos *ab initio* de una o varias moléculas, con una base de buena calidad e incluyendo las correcciones por correlación electrónica y relativistas si el sistema lo requiere. Si el sistema que se quiere simular está compuesto por muchas moléculas, se calculan los potenciales de interacción, comenzando por el de dos cuerpos. Dependiendo de la velocidad de convergencia de la expansión de muchos cuerpos, se calculan las contribuciones de tres o más cuerpos a la energía de interacción.

PASO 2: CONSTRUCCIÓN DE LOS POTENCIALES DE INTERACCIÓN

A partir de los potenciales numéricos calculados en el paso anterior se obtienen potenciales analíticos. Este es el paso crítico para conectar la química cuántica con la mecánica estadística, y debe tenerse mucho cuidado para reproducir de la manera más precisa posible las superficies *ab initio*. La forma analítica debiera tener las siguientes características: sencillez, rapidez en el tiempo de evaluación en la computadora, ser transferible de una molécula a otra, forma y confiabilidad estandarizadas y facilidad para refinarla y extenderla. Esta última característica es una de las que hacen que este sistema de simulación global sea atractivo, ya que permite un refinamiento progresivo del modelo molecular.

Por diseño no se usan datos experimentales para parametrizar los potenciales. Hay

dos razones para esto: una es que en general no hay suficientes datos consistentes y la otra es que en general es arbitrario el extraer las contribuciones de dos y más cuerpos de un experimento que se hizo en condiciones que corresponden a un potencial de n -cuerpos. Una alternativa para sortear el problema de separar estas contribuciones es usar "potenciales efectivos", que son representaciones aditivas de un potencial de muchos cuerpos. Entre los problemas de este ataque está el que estos potenciales efectivos en general dependen fuertemente de la temperatura y la presión, así que no son transferibles de una fase a otra.

La caracterización de los átomos en una molécula es un punto vital en este paso. Se establecen "clases" de átomos dependiendo de su carga neta y sus primeros vecinos. Esto incluye información sobre la hibridización y la transferencia de carga. Esta clasificación de tipos atómicos hace que el potencial sea transferible de una molécula a otra.

Los potenciales analíticos obtenidos en este paso son una representación clásica de fenómenos cuánticos. La validez de esta representación depende del sistema y/o del fenómeno que se quiera estudiar; la opinión general es que las simulaciones clásicas de fluidos son una buena aproximación (ver sección 2.4.1.1).

PASO 3: PROPIEDADES ESTÁTICAS

El tener un potencial átomo-átomo por pares nos permite pasar con facilidad a la mecánica estadística. Primero se analizan las propiedades estáticas con el método de Monte Carlo (MC). Este tipo de simulación ya toma en cuenta a la temperatura. El incluir la temperatura permite calcular entropías y, por lo tanto, energías libres.

PASO 4: PROPIEDADES DINÁMICAS

Con la dinámica molecular (DM) se introduce el parámetro temporal, haciendo accesible el estudio de procesos de transporte. La configuración inicial en general es una configuración equilibrada del paso tres; las velocidades iniciales en general son cero y se ajustan paulatinamente a la temperatura de simulación deseada.

PASO 5: REPRESENTACIÓN DEL CONTINUO

Los coeficientes necesarios para resolver las ecuaciones de difusión se obtienen de los pasos tres y cuatro. En este paso se hacen simulaciones de hidrodinámica.

Si en los pasos tres, cuatro o cinco se encuentran discrepancias con los datos experimentales, se regresa al paso anterior para analizar posibles fuentes de error; en esta forma se refina el modelo. En general hay que regresar varias veces al paso uno porque las superficies de interacción entre dos o más moléculas son multidimensionales (para un dímero de agua hay doce grados de libertad), y es muy difícil tener un muestreo adecuado al primer intento. Las simulaciones pueden indicar las zonas pobremente representadas de estas superficies.

A continuación se presenta un análisis más detallado de los métodos utilizados en los cálculos *ab initio*, en el estudio de los sistemas de muchos cuerpos y la no aditividad, y en las simulaciones numéricas.

2.2. CÁLCULOS *AB INITIO*^{3,4}

En esta sección presentamos una descripción poco formal de los cálculos *ab initio*. Para el lector interesado en un tratamiento más completo y formal, le recomendamos las referencias 3 y 4.

Desde que fue formulada en 1925 la ecuación de Schrödinger, quedó claro que su solución lleva en principio a una predicción cuantitativa directa de la mayoría de los fenómenos químicos usando sólo los valores de un número pequeño de constantes físicas. Estas constantes son la de Plank, la velocidad de la luz, y las masas y cargas de los electrones y núcleos. Tal procedimiento constituye un ataque *ab initio* a la química y no requiere valores experimentales, salvo los de las constantes ya mencionadas.

Aunque la solución a la química, y por lo tanto a la biología, existe, la solución de la ecuación de Schrödinger es un problema matemático insoluble analíticamente salvo para algunos problemas de dos cuerpos, y las soluciones numéricas con un alto grado de precisión sólo son posibles para sistemas muy pequeños, como el átomo de He o la molécula H₂. En la práctica se tienen que hacer una serie de aproximaciones para resolverla; mientras éstas no incluyan datos experimentales, los métodos siguen siendo *ab initio*, a diferencia de los cálculos semiempíricos, que en el proceso de simplificación se ven forzados a sustituir los valores que debiera predecir la ecuación con datos experimentales, o a suponerlos tan pequeños que se pueden despreciar.

Existen varios niveles a los que se puede resolver esta ecuación, niveles que difieren en su acercamiento a la solución exacta. Hay al menos tres características que deben tener estos niveles de aproximación para asegurar un comportamiento adecuado:

- consistencia de talla: los errores relativos relacionados con el cálculo deben crecer más o menos proporcionalmente al tamaño de la molécula. Esto es particularmente importante si se quieren hacer comparaciones cuantitativas de propiedades de moléculas de distintos tamaños o calcular energías de interacción. En general no se puede satisfacer por completo esta condición, pero lo que sí se puede lograr es que haya consistencia de talla para sistemas infinitamente separados, es decir, que las propiedades de un sistema de n moléculas a separación infinita sea igual a la suma de esas propiedades de cada una de las moléculas.
- ser variacional: el nivel de cálculo debe dar una energía total que es una cota superior al valor que resultaría de la solución exacta de la ecuación de Schrödinger completa. Aunque en general no se sabe qué tan cerca se está de la solución exacta, uno puede analizar la convergencia de las energías predichas al ir refinando los cálculos.
- que se pueda implementar en computadoras. Esta característica hace que el refinamiento

que se pueda alcanzar en el cálculo de estructuras electrónicas sea función de la capacidad computacional con la que se cuente.

La descripción que se presenta en seguida de los métodos que existen para resolver la ecuación de Schrödinger está basada en el modelo de orbitales moleculares. Los funcionales de la densidad⁵ son otra manera de resolver la ecuación de Schrödinger, pero aquí no se hablará de ellos por no ser adecuados para el tratamiento de interacciones débiles. También debe tomarse en cuenta que se hablará de la solución a la parte independiente del tiempo.

2.2.1. Aproximaciones para resolver la ecuación de Schrödinger

2.2.1.1 Consideraciones generales

De acuerdo a la mecánica cuántica, la energía y muchas propiedades de un estado estacionario de una molécula se puede obtener resolviendo la ecuación de Schrödinger, que es una ecuación diferencial parcial

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

donde \hat{H} es el Hamiltoniano, un operador diferencial que representa la energía total, cinética y potencial, del sistema formado por núcleos y electrones. E es el valor numérico de la energía del estado, que corresponde a la energía relativa a un estado en el que los núcleos y los electrones están infinitamente separados y en reposo. Ψ es la función de onda, y depende de las coordenadas cartesianas y de espín de todas las partículas. Su cuadrado, $|\Psi^2|$ se interpreta como una medida de la distribución de probabilidad de las partículas en una molécula.

Los Hamiltonianos más sencillos son no relativistas, aproximación que puede no ser válida para el tratamiento de elementos del segundo período de la tabla periódica en adelante, y tampoco incluyen efectos magnéticos como los acoplamientos espín-órbita y espín-espín. Estos últimos no son muy importantes en el cálculo de energías de interacción.

La ecuación de Schrödinger tiene muchas soluciones para cualquier molécula, correspondientes a diferentes estados estacionarios. El estado con la energía más baja es el estado base, que es el que se obtiene en los cálculos variacionales estándar.

El primer paso en la simplificación del problema molecular en general en la mecánica cuántica es la separación de los movimientos de los núcleos y de los electrones, que es válida por la diferencia tan grande en las masas de éstos. Los electrones ajustan rápidamente su distribución a las posiciones cambiantes de los núcleos, de manera que es razonable proponer que la distribución de electrones depende sólo de las posiciones instantáneas de los núcleos y no de sus velocidades. Esto lleva a una energía electrónica efectiva $E(R)$ que depende de las coordenadas nucleares relativas R . Esta energía efectiva se usa como una energía potencial en estudios subsecuentes del movimiento nuclear. A esto se le llama la aproximación adiabática o de Born-Oppenheimer. El Hamiltoniano total se reduce a un Hamiltoniano electrónico, que contiene términos de energía cinética para los electrones, la energía potencial entre los núcleos y los electrones, la repulsión electrónica y un término del potencial Coulombiano entre los núcleos, que es constante para una configuración nuclear

dada.

$E(R)$ es fundamental en la descripción cuantitativa de estructuras químicas y de procesos de reacción. Los mínimos locales de esta superficie son estructuras de equilibrio, y los puntos silla son estructuras de transición. En la realidad hay movimiento nuclear alrededor de los mínimos por la vibración de punto cero, pero en general estos mínimos son una buena aproximación a la estructura promedio.

Dentro del modelo de orbitales moleculares, el tratamiento aproximado de la distribución electrónica y su movimiento se hace asignando electrones individuales a funciones monoeléctricas llamadas espín-orbitales χ . Estos están formados por un producto de funciones espaciales, llamadas orbitales moleculares ψ , y funciones de espín α o β . Los espín-orbitales tienen completa libertad de extenderse por toda la molécula, y la asignación de electrones a espín-orbitales particulares define la configuración electrónica.

Los orbitales ψ pueden estar doblemente ocupados, semiocupados o no ocupados (también llamados virtuales). Para las moléculas que tienen un número par de electrones, el estado base se puede representar a primera aproximación con una configuración electrónica donde hay orbitales doblemente ocupados y vacíos. A este tipo de configuración se le llama de capa cerrada.

Las únicas soluciones de la ecuación de Schrödinger que son físicamente aceptables son aquellas con la simetría adecuada ante el intercambio de partículas idénticas. Para los bosones, Ψ es simétrica ante el intercambio, pero para fermiones Ψ es antisimétrica. Como los electrones son fermiones, Ψ debe ser antisimétrica. Así se cumple con el Principio de Exclusión de Pauli.

La función de onda más sencilla que se puede proponer es un producto antisimetrizado de espín-orbitales; este producto suele escribirse en forma de determinante porque los determinantes tienen las propiedades matemáticas adecuadas para describir la antisimetría: el intercambio de dos renglones cambia el signo del determinante y el hacer iguales dos columnas hace que el determinante valga cero. Físicamente, el primer movimiento corresponde a una permutación en la asignación de electrones a orbitales, y el segundo caso equivale a asignar más de un electrón al mismo espín-orbital, cosa que está prohibida por el Principio de Exclusión de Pauli.

En la práctica, los orbitales moleculares ψ se restringen a ser combinaciones lineales de un juego de N funciones monoeléctricas ϕ_μ :

$$\psi_i = \sum_{\mu=1}^N c_{\mu i} \phi_\mu$$

Las funciones ϕ_μ se conocen como las funciones de la base, y al conjunto de ellas se le llama simplemente la base. Si estas funciones son los orbitales atómicos de los átomos que constituyen a la molécula, la construcción de ψ se conoce como la aproximación de la combinación lineal de orbitales atómicos (LCAO).

Otras características que deben tener los orbitales son la de ser ortogonales y estar normalizados. Se normalizan porque la probabilidad de encontrar un electrón en cualquier lugar del espacio es 1. La normalización y ortogonalización dan lugar a la función de onda multielectrónica para el estado base de capa cerrada de una molécula con n electrones ocupando doblemente $n/2$ orbitales. Esta representación se llama el determinante de Slater o configuración electrónica.

Al escribir la función de onda como un producto antisimetrizado de orbitales se está suponiendo que el movimiento de los electrones no está correlacionado, salvo para el caso de electrones con el mismo espín que deben cumplir con el Principio de Exclusión de Pauli.

Una vez que se ha construido el determinante de Slater, el siguiente paso es encontrar los valores de los coeficientes de la base. Esto se hace con el método variacional dentro de la teoría de Hartree-Fock (HF). El teorema variacional establece que si ϕ es cualquier función normalizada antisimétrica de las coordenadas electrónicas, el valor de expectación de la energía E' correspondiente a esta función se puede obtener de la integral

$$E' = \int \phi^* \hat{H} \phi d\tau$$

donde se integra sobre las coordenadas espaciales y de espín (x, y, z, ξ) de todos los electrones. Si ϕ es la función de onda exacta Ψ , entonces se satisface la ecuación de Schrödinger, y E' es la energía exacta E :

$$E' = E \int \Psi^* \Psi d\tau = E.$$

Si ϕ es cualquier otra función de onda normalizada y antisimétrica, se puede demostrar que

$$E' = \int \phi^* \hat{H} \phi d\tau > E.$$

Este método se puede usar para determinar los orbitales óptimos de funciones de onda de un determinante. Se escoge una base y los coeficientes $c_{\mu i}$ se ajustan para minimizar E' . E' está tan cerca como es posible de E dentro de las limitaciones impuestas por la función de onda de un determinante y la base empleada.

Físicamente, la aproximación de HF desprecia la repulsión instantánea entre los electrones. Dentro de HF, cada electrón interactúa con la distribución promedio de carga que proviene del resto de los electrones. Esto significa que los electrones se ven unos a otros como distribuciones en lugar de como partículas discretas. Matemáticamente, el resultado es que el término de repulsión de dos cuerpos en el Hamiltoniano molecular se reduce a un término de un cuerpo. El potencial efectivo HF corresponde al campo promedio de la distribución de carga, misma que depende de la función de onda. Por lo tanto, la resolución de la ecuación de Schrödinger HF se hace de manera autoconsistente (teoría de campo autoconsistente o SCF).

Los orbitales moleculares óptimos son los que satisfacen la ecuación de Hartree-Fock (HF):

$$\hat{F}(k)\phi_i(k) = \epsilon_i\phi_i(k)$$

donde ϵ_i es la energía del orbital i y k representa un electrón. El operador de Fock \hat{F} contiene el Hamiltoniano de un solo electrón: su energía cinética y la energía potencial con los núcleos y los demás electrones. La suma de estos operadores no produce el Hamiltoniano total porque se cuentan de más las repulsiones interelectrónicas. Esto es una consecuencia del término de dos cuerpos, que hace que el problema no sea de variables separables. La aproximación SCF cambia el término de repulsión por un operador que representa el campo promedio generado por los otros electrones, y esto sí permite la separación de variables porque todos los operadores son de un cuerpo:

$$\hat{F}(k) = \hat{h}(k) + \hat{g}(k)$$

donde $\hat{h}(k)$ es el Hamiltoniano para el electrón en el campo de los núcleos y $\hat{g}(k)$ representa el campo promedio que el electrón k experimenta debido a la presencia del resto de los electrones. \hat{g} contiene un término Coulombiano y uno de intercambio, que es consecuencia de la antisimetría y no tiene una contraparte clásica. \hat{F} se usa para encontrar los orbitales óptimos reemplazando a \hat{H}_{total} .

Las ecuaciones de HF se pueden resolver numéricamente, pero Roothan mostró que se pueden escribir de forma matricial si se expanden sobre una base de funciones de orbitales atómicos. La determinación de los coeficientes de la expansión se hace variacionalmente.

Como la base es finita, hay errores por el truncamiento de la misma; esto no sucede si se resuelven las ecuaciones de HF numéricamente. El límite de HF es el punto en el que la base de la función de onda SCF hace que esta función de onda esté muy cerca de la que se obtiene de la solución numérica.

Al examinar los orbitales moleculares de la molécula de agua obtenidos a nivel SCF se ve que los enlaces y pares electrónicos están enmascarados por la naturaleza deslocalizada de los orbitales moleculares. Esta descripción no localizada es útil para definir la simetría y para discutir transiciones electrónicas, pero acaba con la intuición química. Los orbitales se pueden localizar transformando los orbitales moleculares de manera que se maximice la energía Coulombiana intraorbital, dando lugar, por ejemplo, a la imagen convencional casi tetraédrica de los orbitales híbridos para el agua. Esta es una imagen cualitativamente correcta.

HF-SCF no da una descripción completa de la distribución electrónica. Para efectos de los modelos que se usan para resolver la ecuación de Schrödinger, los electrones son partículas e interactúan como tales. Dada la repulsión electrónica, una función de onda verdadera debe dar una probabilidad menor de encontrar un electrón en la vecindad de otro en cualquier instante, no sólo en promedio.

2.2.1.2 Correlación electrónica

Las funciones de onda exactas generalmente no se pueden expresar con un sólo determinante. Las representaciones de un determinante no toman en cuenta la correlación entre electrones con espín opuesto, y la correlación de los movimientos de los electrones con el mismo espín (hoyo de Fermi) está parcialmente tratada en virtud de la forma de determinante de la función de onda. Estas limitaciones hacen que las energías calculadas bajo la

teoría de HF, aún en el límite de HF, sean superiores a las exactas. Por convención, la diferencia entre las energías de HF y las exactas no relativistas es la energía de correlación:

$$E(\text{exacta}) = E(\text{HF}) + E(\text{correlación}).$$

El despreciar la correlación entre electrones de espín opuesto (hoyo de Coulomb) lleva a un número de deficiencias cualitativas en la descripción de la estructura electrónica; por ejemplo, no hay una disociación correcta, no se logra la simetría adecuada de la función de onda para describir la estructura electrónica, las energías de disociación de los enlaces se subestiman, etc. Las geometrías de equilibrio calculadas a nivel HF-SCF en general son buenas, pero hay correcciones en las longitudes y ángulos de enlace al incluir correlación electrónica. Las frecuencias de vibración, que son una medida de la forma de la superficie de potencial cerca del mínimo, son muy sensibles al tratamiento de la correlación electrónica.

Hay dos tipos de correlación:

-no dinámica: incluye los determinantes hechos con los orbitales moleculares de valencia, ocupados y virtuales, y es el único tipo de correlación que se puede tratar con una base mínima (ver sección 2.2.2.2). Se presenta cuando un sólo determinante no es capaz de representar toda la Ψ , cuando hay configuraciones degeneradas o casi degeneradas con Ψ_{SCF} o cuando al mezclar un número pequeño de configuraciones se obtiene una Ψ dominante.

-dinámica: incluye los determinantes hechos con orbitales moleculares que no son de valencia; por lo tanto se requiere de una base extendida (ver sección 2.2.2.2) para poder tratar este tipo de correlación. Físicamente corresponde a la energía de dispersión, que proviene de la correlación de las fluctuaciones de la distribución de densidad electrónica de moléculas vecinas. Clásicamente se puede modelar como la interacción de dipolos inducidos. En general la inclusión de la correlación no dinámica mejora más la energía que la de la dinámica.

La libertad matemática requerida para tratar la correlación electrónica viene del mezclado o interacción de configuraciones, que es simplemente permitir que la función de onda sea una combinación lineal de determinantes de Slater,

$$\Psi = a_0 \Psi_0 + \sum_{s>0} a_s \Psi_s,$$

cada uno de los cuales representa una configuración electrónica particular. Al mezclar las configuraciones se hace accesible a los orbitales virtuales (o sea, los orbitales no ocupados) una fracción de la población electrónica que en el esquema de HF estaba confinada a los orbitales ocupados. Por lo tanto, la población de los orbitales ocupados disminuye ligeramente. La dominancia de la Ψ de referencia se puede medir con el coeficiente a_0 de la combinación lineal de determinantes de Slater. Una función de onda será más dominante mientras más cercano a 1.0 sea este coeficiente.

Hay varias maneras de determinar la combinación lineal de configuraciones electrónicas. La manera más directa es hacer que los coeficientes de la combinación lineal sean parámetros variacionales. A este método se le llama interacción de configuraciones (CI).

2.2.1.2.1 Interacción de configuraciones completa (full CI)

En un sistema de n electrones descritos por una base de N funciones ϕ_μ , hay $2N$ espín-orbitales χ . Una vez que se resuelve el problema de HF se obtiene Ψ_0 , que sólo contiene los orbitales ocupados. Los orbitales de la base que no están ocupados se denominan orbitales virtuales. Por convención, los orbitales ocupados se marcan con las letras i, j, k, \dots y los virtuales con las letras a, b, c, \dots . Las funciones de onda determinante distintas a Ψ_0 se pueden construir reemplazando uno o más de los orbitales ocupados χ_i por orbitales virtuales χ_a . Estos determinantes se llaman Ψ_s , con $s > 0$. Ψ_s se pueden clasificar de acuerdo al número de sustituciones, también llamadas excitaciones: monoexcitaciones Ψ_i^a , diexcitaciones Ψ_{ij}^{ab} , triexcitaciones Ψ_{ijk}^{abc} , etc. hasta llegar al determinante en el que todos los orbitales ocupados han sido sustituidos por orbitales virtuales. La función de onda multideterminante es entonces

$$\Psi = a_0 \Psi_0 + \sum_{s>0} a_s \Psi_s$$

y los coeficientes a_s se determinan variacionalmente.

Este es el tratamiento más completo posible dentro de las limitaciones impuestas por la base. Es variacional y tiene consistencia de talla, dos de las características más importantes, pero no es práctico salvo para sistemas muy pequeños porque el número total de determinantes es $(2N!)/\{n!(2N-n)!\}$ (donde N es el número de orbitales y n es el número de electrones). Por ejemplo, para una molécula de agua calculada con la base DZP descrita en el apéndice A, tenemos 10 electrones y 23 orbitales, lo cual da un total de 4 076 350 421 coeficientes para el cálculo variacional.

Hay diferencias importantes entre el tipo de configuraciones en términos de lo que describen y de cómo interactúan. Las sustituciones sencillas o monoexcitaciones no interactúan con la Ψ_{SCF} de acuerdo al teorema de Brillouin. Este teorema implica que el estado base calculado a nivel SCF es estable. Si esto no fuera cierto, el estado base sería fuertemente perturbado por las configuraciones monoexcitadas cercanas. Como sí es cierto, el mezclado directo del estado base sólo se da con niveles energéticos diexcitados, e indirectamente con los estados monoexcitados que interactúan con los multiexcitados⁶. Las configuraciones con sustituciones dobles sí interactúan con el estado base. Las triexcitaciones no, porque el Hamiltoniano sólo tiene operadores de uno y dos cuerpos, pero sí interactúan con las configuraciones mono y diexcitadas. En términos perturbativos, a primer orden sólo entran las diexcitaciones; y las contribuciones de las sustituciones triples y superiores tienden a ser menores y se conocen como efectos de orden superior.

En vista de esto, se han propuesto métodos de Interacción de Configuraciones Limitada:

el CID, que sólo incluye las diexcitaciones:

$$\Psi_{CID} = a_0 \Psi_0 + \sum_{i < j} \sum_{a < b}^{occ \text{ virt}} a_{ij}^{ab} \Psi_{ij}^{ab}$$

el CISD, que incluye mono y diexcitaciones:

$$\Psi_{CID} = a_0 \Psi_0 + \sum_{i < j} \sum_{a < b}^{occ\ virt} a_{ij}^{ab} \Psi_i^{ab} + \sum_{i < j} \sum_{a < b}^{occ\ virt} a_{ij}^{ab} \Psi_{ij}^{ab}$$

Otra opción para hacer más manejable el cálculo es limitar el número de orbitales ocupados que se puede excitar y el número de orbitales virtuales accesibles:

-permitir excitaciones del HOMO (Highest Occupied Molecular Orbital) al LUMO (Lowest Unoccupied Molecular Orbital). Esto tiene el problema de que estos orbitales no están bien definidos en el caso de degeneraciones del estado base.

-permitir excitaciones de los electrones de valencia a todos los orbitales virtuales; esta aproximación recibe el nombre de *frozen core* y es compatible con el desarrollo de pseudopotenciales (ver más adelante).

Al truncar la expansión CI ya no hay consistencia de talla. Esto se arregla en parte con la corrección de Davidson:

$$\Delta E_{D_{av}} = (1 - a_0^2) \Delta E_{CID}.$$

donde a_0 es el coeficiente de la Ψ_0 en la expansión CI.

Otra manera de calcular la energía de correlación es definir la diferencia entre el Hamiltoniano molecular verdadero y el Hamiltoniano SCF como una perturbación y tratar el problema con teoría de perturbaciones.

2.2.1.2.2 Teoría de Perturbaciones de Møller-Plesset

El problema sigue siendo encontrar el eigenvalor más bajo y el eigenvector correspondiente de la matriz completa del Hamiltoniano. Aquí lo que se hace es tratarlo como la suma de dos partes, donde la segunda es una perturbación de la primera. Se introduce un Hamiltoniano electrónico generalizado \hat{H}_λ

$$\hat{H}_\lambda = \hat{H}_0 + \lambda \hat{V}$$

donde \hat{H}_0 es la suma de operadores de Fock monoeléctricos y $\lambda \hat{V}$ es la perturbación, que se define como

$$\lambda \hat{V} = \lambda (\hat{H} - \hat{H}_0)$$

donde \hat{H} es el Hamiltoniano correcto y λ es un parámetro adimensional. El eigenvalor E_λ correspondiente al determinante Ψ_λ es la suma de las energías monoeléctricas ϵ_i de los espín-orbitales ocupados en Ψ_λ .

Ψ_λ y E_λ , la función de onda y energías exactas para el sistema descrito por el Hamiltoniano \hat{H}_λ , se pueden expandir en potencias de acuerdo con la teoría de perturbaciones de Rayleigh-Schrödinger (RSPT):

$$\Psi_\lambda = \Psi^{(0)} + \lambda \Psi^{(1)} + \lambda^2 \Psi^{(2)} + \dots$$

$$E_{\lambda} = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots$$

λ se hace uno y se corta la expansión a varios órdenes. Si se corta después del segundo orden se llama MP2. No existe MP1 porque Møller y Plesset demostraron que la energía HF es correcta dentro de este esquema hasta el primer orden. MP2 contiene una suma sobre todas las doble excitaciones; en este nivel se empieza a tratar la correlación, y probablemente ésta sea la manera más sencilla de tratarla. MP3 es equivalente a CISD. A cuarto orden entran todas las sustituciones sencillas, dobles, triples y cuádruples; como las triexcitaciones son muy difíciles de manejar, existe una aproximación llamada MP4SDQ que sólo incluye las sencillas, dobles y cuádruples.

El tratamiento perturbativo usando la partición de Møller-Plesset a cualquier orden sí tiene consistencia de talla, pero no es variacional.

2.2.1.2.3 Coupled Cluster

Otra manera de tratar la mezcla de configuraciones es la de *coupled cluster* (CC), que genera la función de onda correlacionada aplicando un operador de sustitución exponencial a una configuración de referencia que en general es Ψ_{SCF} . CI se puede expresar de la misma manera, pero el operador de sustitución es lineal en lugar de exponencial. El exponencial es otro operador, que por conveniencia se escribe en términos de una serie de potencias del operador de sustitución. El primer término de la serie es 1, el segundo es el operador de sustituciones dobles y el tercero es ese operador al cuadrado entre 2. Este término hace todas las sustituciones dobles de las sustituciones dobles, generando sustituciones cuádruples. El cuarto término genera las sustituciones héxtuples y así en adelante. De esta manera los efectos de orden superior se incluyen en la función de onda. CCD y CID tienen el mismo número de parámetros, uno por configuración diexcitada. Sin embargo, Ψ_{CCD} tiene muchas más configuraciones, y sus coeficientes de expansión son productos de los de las configuraciones diexcitadas.

2.2.1.2.4 Cálculos multirreferenciales

Hasta ahora, las excitaciones se han construido a partir de una sola función de referencia. Existen cálculos de correlación electrónica a partir de funciones SCF multirreferenciales (MRSCF), en los que se ajustan los orbitales para mejorar la energía de una expansión configuracional. Este es un método muy eficiente para tratar la correlación no dinámica, y si Ψ_{MRSCF} se usa como referencia para un tratamiento CID o CISD, prácticamente se tiene algo semejante a *full CI*³. Este tipo de cálculos se llama MR-CI (interacción de configuraciones multirreferencial).

El programa CIPSI⁷ (Configuration Interaction by Perturbation with multiconfigurational zeroth order wavefunction Selected by Iterative process) utilizado en esta tesis para el cálculo de la energía de correlación hace una expansión perturbativa de Rayleigh-Schrödinger a partir de una función de onda multiconfiguracional. Se define un proceso iterativo que permite el mejoramiento progresivo de la Ψ_0 . La matriz CI se restringe a un subespacio S de configuraciones con interacción fuerte y este espacio se diagonaliza. El

eigenvector resultante se perturba a primer orden por los determinantes que no están en S. Los determinantes que interactúan fuertemente a primer orden (de acuerdo a un criterio seleccionado por el usuario) se incluyen en S y así sucesivamente hasta que se converge, o sea, ya no se seleccionan más determinantes para ese umbral de interacción. La suma de los coeficientes de la expansión CI indica cuánto de la energía del agregado se está obteniendo de manera variacional; la norma de un cálculo se define como la resta de 1 menos esta suma de coeficientes, y mientras más chica sea, mejor será el cálculo. Para el estado base, este procedimiento incluye toda la corrección a segundo orden y los términos más importantes de tercer y cuarto orden. Si se usa la partición de Møller-Plesset, el método tiene consistencia de talla. Aunque tiene una parte variacional, otra parte del tratamiento es perturbativo, así que las energías calculadas con este método ya no están acotadas.

Una regla general es que la precisión aumenta a medida que se incluyen en el cálculo determinantes con sustituciones más numerosas. Por lo tanto, *full CI* es el tratamiento más completo.

2.2.2 Las bases moleculares

Los resultados que se obtienen de cálculos moleculares dependen fuertemente de la base de orbitales atómicos o moleculares que se empleen. En esta sección mostramos tres maneras de construirlos.

Los orbitales moleculares se construyen con una combinación lineal de otras funciones:

$$\psi_i = \sum_{\mu=1}^N c_{\mu i} \phi_{\mu}$$

donde $c_{\mu i}$ son los coeficientes de la expansión.

$\phi_{\mu i}$ puede ser cualquier función, pero en general se usan las que tienen las propiedades de simetría de los orbitales atómicos, así que se pueden clasificar como s, p, d, f, ... de acuerdo a sus propiedades angulares.

Hay dos tipos de $\phi_{\mu i}$ con uso generalizado:

2.2.2.1 Orbitales tipo Slater (STO)

Éstos tienen la forma funcional adecuada para ser soluciones de la ecuación de Schrödinger, una dependencia radial exponencial, por ejemplo:

$$\phi_{1s} = \left(\frac{\zeta_1^3}{\pi}\right)^{1/2} \exp(-\zeta_1 r)$$

donde ζ_1 es una constante que determina el tamaño del orbital. Estas funciones dan una representación adecuada de los orbitales atómicos, pero su evaluación numérica es muy costosa. Por esto se propuso el uso de los orbitales tipo Gaussiano como una aproximación a los STO's.

2.2.2.2 Orbitales tipo Gaussiano (GTO)

Éstos tienen dependencia radial Gaussiana, por ejemplo:

$$g_s(\alpha, r) = \left(\frac{2\alpha}{\pi}\right)^{3/4} \exp(-\alpha r^2)$$

donde α determina el tamaño o extensión radial de la función. Los GTO son menos satisfactorios como representaciones de los orbitales moleculares. Sin embargo, todas las integrales se pueden evaluar explícitamente sin tener que recurrir a la integración numérica. En la práctica se usan combinaciones lineales de GTO:

$$\phi_\mu = \sum_s d_{\mu s} g_s$$

donde g_s son llamadas Gaussianas primitivas y ϕ_μ se llama Gaussiana contraída. En este caso los coeficientes $d_{\mu s}$ son fijos, no entran al teorema variacional salvo cuando se quiere optimizar una base. El optimizar una base para que reproduzca bien la energía no quiere decir que vaya a reproducir otras propiedades moleculares adecuadamente.

Las diferencias matemáticas entre STO's y GTO's son las siguientes. Los GTO tienen la primera derivada igual a cero en el origen, y caen más rápido a cero conforme r crece que los STO. Los STO tienen la primera derivada finita en el origen, donde tienen una cúspide. Por lo anterior, si se quiere reproducir el comportamiento de los STO con GTO, en los GTO primitivos deben tener, por un lado, GTO con α 's grandes, que dan funciones muy localizadas y compactas, para tratar de reproducir la cúspide, y por el otro, GTO con α 's chicas, que dan funciones muy difusas, para reproducir el comportamiento a r grandes. Aun usando un número infinito de Gaussianas primitivas, nunca se llega a reproducir perfectamente un STO con GTO's. Esto se ve reflejado en el hecho de que se necesitan muchos más GTO's que STO's para llegar al límite de HF. A pesar de esta deficiencia de los GTO's, las ventajas computacionales que presentan hacen que sean muy utilizados.

Las ϕ_μ se pueden centrar donde sea, pero comúnmente se centran en los núcleos de los átomos.

La calidad de la base depende de:

- el número total de funciones contraídas.
- el número de funciones primitivas.
- el valor de los exponentes α .
- el valor de los coeficientes de contracción.
- los diferentes tipos de funciones (s, p, d, f, ...) en la base.

El valor de los exponentes y coeficientes de contracción reflejan la optimización de la base al átomo o molécula correspondiente.

Las bases mínimas son aquellas con una función ϕ_μ por orbital ocupado. Como los exponentes de los STO se llaman ζ y hay uno por función, se ha clasificado la extensión

de las bases con el número de ζ 's. Por ejemplo, una base doble (DZ) tiene dos funciones por orbital, una triple ζ tiene tres y así sucesivamente. Mientras más funciones se tengan por orbital se tendrá una mayor flexibilidad para describirlo. Además de las funciones que representan al estado base, a veces se incluyen funciones de polarización, que aumentan aún más la flexibilidad de la descripción de la configuración electrónica, y son indispensables para poder tratar la correlación electrónica, ya que son los orbitales virtuales.

En general, la descripción de los electrones de valencia requiere más flexibilidad que la de los electrones internos; esto ha generado por un lado las bases de valencia escindida³ (*split valence*) y los pseudopotenciales⁸ por el otro. Las bases de valencia escindida tratan a los electrones internos con una base mínima, y a los de valencia con bases de calidad DZ en adelante. Los pseudopotenciales tratan explícitamente a los electrones de valencia nada más, y los electrones internos proveen el campo en el que se mueven. En algunos desarrollos de pseudopotenciales se ajusta un potencial (el pseudopotencial) a los electrones internos y el núcleo. Este potencial tiene una barrera energética que evita que los electrones de valencia se metan al *core*. En el desarrollo de Durand y Barthelat⁹ el pseudopotencial se ajusta a un hamiltoniano compuesto por los electrones internos y el núcleo, y se hace la función de onda de los electrones de valencia igual a cero al entrar al *core*. Cada electrón de valencia se trata como en los casos anteriores: se le asigna un Hamiltoniano y después se puede hacer correlación electrónica, bajo la aproximación de *frozen core* porque no hay electrones internos que se puedan excitar. La aproximación de pseudopotenciales requiere que las bases atómicas se optimicen para éstos³.

2.2.2.3 Geminales

Las moléculas también se pueden resolver dentro del esquema de *valence-bond* propuesto por Pauling, en el que los orbitales moleculares son "geminales" e incluyen dos electrones: para la molécula de agua son cinco, y corresponden a la capa interna, dos enlaces OH y dos pares electrónicos. Este enfoque incluye algo de correlación electrónica.

2.2.3 Cálculo de energías de interacción

Los métodos *ab initio* para el cálculo de energías de interacción se pueden clasificar¹⁰ como supermoleculares, perturbativos o híbridos. El método supermolecular obtiene la energía de interacción como

$$\tilde{E}_{int} = \tilde{E}_{AB} - (\tilde{E}_A + \tilde{E}_B)$$

donde \tilde{E}_X es una aproximación a la energía exacta del estado base E_X . Para interacciones típicas entre sistemas de capa cerrada, la energía de interacción exacta

$$E_{int} = E_{AB} - (E_A + E_B)$$

es de cuatro a siete órdenes de magnitud más chica que las energías totales E_X . Como los errores $E_X - \tilde{E}_X$ son siempre mucho más grandes que la energía de interacción exacta, \tilde{E}_{int} es una buena aproximación a E_{int} sólo si hay una afortunada cancelación de errores. Hay una serie de condiciones necesarias (aunque no suficientes) que debe de cumplir el método supermolecular para dar una descripción razonable de interacciones intermoleculares:

-el cálculo tiene que tener consistencia de talla.

-debe tratar la correlación electrónica suficientemente bien.

-debe eliminarse el error por superposición de bases (BSSE). Este error se debe a que los orbitales de un centro atómico son usados por un átomo vecino porque su base es incompleta. Si la base propia no es suficientemente flexible, la base de átomos cercanos se puebla con la densidad electrónica propia¹¹. Esto lleva a un error porque la energía de una molécula es más baja mientras más deslocalizados estén sus electrones, y los orbitales de los átomos vecinos permiten una deslocalización artificial.

La corrección counterpoise (CP) fue propuesta por Boys y Bernardi¹² para compensar en forma aproximada el problema de BSSE. Se calculan los monómeros en la base del agregado completo para calcular la atracción de los electrones de los monómeros por la base ajena. A esto se le llama *full CP* o CP completo; por limitaciones de cómputo, en ocasiones se calculan los monómeros en la base de subconjuntos de moléculas del agregado en estudio. Al restar estas contribuciones se espera, si no eliminar el error, por lo menos reducirlo. Esta corrección no es variacional y por lo tanto no está acotada. Esto hace que hayan opiniones divididas acerca de si se debe de incluir o no¹³⁻¹⁸.

Los métodos perturbativos calculan E_{int} directamente (no hay que restar números grandes para obtener números pequeños) como una suma de diferentes contribuciones físicas:

$$E_{int} = E_{pol}^{(1)} + E_{exch}^{(1)} + E_{pol}^{(2)} + E_{exch}^{(2)} + \dots$$

donde $E_{pol}^{(1)}$ es la energía de interacción electrostática clásica y contiene la inducción clásica además de la energía de dispersión, que es un efecto cuántico

$$E_{pol}^{(2)} = E_{ind}^{(2)} + E_{disp}^{(2)}$$

$E_{exch}^{(n)}$ son las correcciones por el intercambio.

Para un dímero AB, el Hamiltoniano se divide como $H = H_0 + V$, donde $H_0 = H_A + H_B$ es el Hamiltoniano no perturbado y V es la perturbación, o sea, la interacción molecular.

Los métodos híbridos calculan la energía de interacción a nivel SCF con el método supermolecular y el resto se hace de manera perturbativa:

$$E_{int} = E_{int}^{HF} + E_{intra}^{(1)} + E_{ind-intra}^{(2)} + E_{disp}^{(2)} + E_{exch-disp}^{(2)} + \dots$$

donde los términos E_{intra} corresponden a la contribución de la correlación electrónica intramolecular.

La ventaja de los métodos perturbativos es que cada término de la expansión se puede calcular con una base optimizada para esa propiedad, y además no hay problema de BSSE.

2.2.4 Cálculo de propiedades de complejos de van der Waals

Las interacciones débiles entre átomos y moléculas de capa cerrada juegan un papel importante en una gran variedad de fenómenos físicos y químicos, y además son esenciales

para un entendimiento a nivel molecular de los procesos bioquímicos. Estos últimos en general comienzan con interacciones de van der Waals y/o la formación de puentes de H. El cálculo de estas interacciones no es trivial porque los valores de la energía de interacción son muy pequeños (cuando mucho 10 kcal/mol para una interacción por pares), y muchas veces el error en los cálculos es de esa magnitud. Se requiere de bases extendidas (por lo menos DZ), bien balanceadas, y con funciones de polarización. También se requiere incluir correlación electrónica, tanto la no dinámica como la dinámica, por lo menos a nivel MP2. El tratamiento adecuado de la energía de dispersión es indispensable si se quiere reproducir tanto la energía como la geometría de la interacción. Los agregados de moléculas de agua son complejos de van der Waals; por lo tanto debe tenerse mucho cuidado al realizar cálculos de energías de interacción y de propiedades estructurales.

2.3 LA NO ADITIVIDAD EN LA INTERACCIÓN MOLECULAR

En el agua líquida la interacción entre los monómeros de agua se ve modificada por la presencia de las demás moléculas. Estos son los efectos de muchos cuerpos o la no aditividad. Los efectos no aditivos se pueden explicar en términos clásicos a través de interacciones de inducción electrostática a distancias grandes (varios radios atómicos). Sin embargo, a distancias cortas (al contacto de las esferas de van der Waals) aparecen términos de intercambio, y la explicación ha de ser cuántica.

A medida que una tercera molécula de agua se aproxima a un dímero pueden suceder varias cosas dependiendo de la distancia y la línea de aproximación. La interacción más lejana se da entre las nubes electrónicas del dímero y la carga efectiva de la tercera molécula. Esto altera la polarización de los orbitales del dímero, y por lo tanto su interacción. Al acercarse más la tercera molécula al dímero, la diferencia en electronegatividades de los átomos del agregado puede causar la transferencia de carga entre ellos, relocalizando la carga; esto también cambia la interacción original del dímero. A distancias menores se forma un agregado en el que los electrones tienen acceso a los orbitales de las tres moléculas. Si las moléculas son flexibles, puede haber deformación molecular desde el primer paso, cambiando tanto la polarizabilidad como las longitudes de equilibrio de los enlaces moleculares.

La manera de tratar cuantitativamente este problema es estudiar la convergencia de la expansión de muchos cuerpos. La energía potencial de interacción de un agregado de N moléculas se puede escribir de la siguiente forma:

$$E_N - \sum_i^N E_i = \sum_{i<j}^N V_{ij} + \sum_{i<j<k}^N \epsilon_{ijk} + \sum_{i<j<k<l}^N \epsilon_{ijkl} + \dots$$

donde por definición

$$V_{ij} \equiv E_{ij} - (E_i + E_j)$$

es la energía de interacción de un par de moléculas y las ϵ son las contribuciones de la no

aditividad producida por tres, cuatro o más cuerpos:

$$\epsilon_{ijk} = E_{ijk} - \sum E_i - \sum V_{ij} = E_{ijk} - \sum E_{ij} + \sum E_i$$

$$\epsilon_{ijkl} = E_{ijkl} - \sum E_i - \sum V_{ij} - \sum \epsilon_{ijk} = E_{ijkl} - \sum E_{ijk} + \sum E_{ij} - \sum E_i.$$

E representa la energía potencial total de un agregado y los índices i, j, k y l representan moléculas.

Se ha demostrado que para el caso del agua esta serie converge¹⁹⁻²¹, aunque a diferentes velocidades dependiendo de la fase. Para un pentámero tetrahédrico representativo del hielo²¹, $\sum \epsilon_{ijk}$ es \sim el 6% de la contribución de dos cuerpos, $\sum \epsilon_{ijkl}$ es el 1% de $\sum \epsilon_{ijk}$, y la corrección de cinco cuerpos es cero dentro de la resolución del cálculo. Sin embargo, en estudios de tetrámeros representativos de las fases líquida y gaseosa se encontró¹⁹ que la no aditividad de tres cuerpos puede ser hasta del 30% de la energía de dos cuerpos, y que la no aditividad de cuatro cuerpos puede llegar a ser el 47% de la de tres cuerpos. En resumen, la velocidad de la convergencia de esta serie depende fuertemente de la distancia intermolecular y de la simetría de los agregados. La convergencia es más rápida para agregados simétricos, que además son los más sencillos de calcular y por lo tanto los más estudiados.

Todo lo que se ha discutido se refiere a moléculas rígidas. El tratamiento de las deformaciones moleculares no está descrito, así que proponemos la siguiente alternativa:

a) calcular las contribuciones no aditivas con la expresión canónica, que no trata explícitamente las deformaciones moleculares. El problema de esta alternativa es que no es claro dónde aparecerá la energía de deformación.

b) describir el desarrollo de muchos cuerpos incluyendo un término que trate explícitamente la energía de deformación:

$$E_N - \sum_i^N E_i = \sum_{i<j}^N V_{ij} + \sum_{i<j<k}^N \epsilon_{ijk} + \sum_{i<j<k<l}^N \epsilon_{ijkl} + \sum_i^N \delta_i + \dots$$

donde

$$\delta \equiv E^d - E$$

es la energía de deformación, E es la energía del monómero en la geometría de equilibrio y E^d es la energía del monómero deformado. Para que el tratamiento sea consistente definimos una nueva energía por pares

$$V_{ij} \equiv E_{ij} - (E_i^d + E_j^d)$$

y las no aditividades se calculan de la siguiente manera

$$\epsilon_{ijk} = E_{ijk} - \sum E_{ij} + \sum E_i^d$$

$$\epsilon_{ijkl} = E_{ijkl} - \sum E_{ijk} + \sum E_{ij} - \sum E_i^4$$

Esta forma se reduce a la expresión canónica en el caso de no haber deformación molecular.

El cálculo *ab initio* de no aditividades es caro en términos computacionales, tanto en tiempo de procesador como en memoria. En principio, hay que hacer cálculos con correlación electrónica de todos los agregados del sistema para obtener las E necesarias para calcular las contribuciones de n cuerpos. Sin embargo, se encontró²² que, dentro de la precisión de los cálculos, la energía de correlación es aditiva en su mayor parte, de manera que los cálculos a nivel HF son una buena primera aproximación. Como los valores de las no aditividades son muy pequeños, hay que tener mucho cuidado con la base que se utilice y corregir el error por superposición de bases (BSSE)²⁰ con CP completo (es decir, calcular cada subagregado en la base del agregado total).

El cálculo de la energía de interacción de agregados de tres o más cuerpos incluyendo correlación electrónica es muy costoso. Una manera alternativa para calcular esta energía de interacción es utilizar la expansión de muchos cuerpos, incluyendo correlación electrónica en el cálculo de las energías de deformación de los monómeros y de las energías de interacción de los dímeros, que son la parte aditiva de la expansión.

Una vez establecido el punto donde se va a truncar la expansión de muchos cuerpos, se procede a ajustar un potencial analítico a la colección de datos que se obtuvo con los cálculos *ab initio* de los agregados moleculares del sistema de interés. Este potencial analítico es el puente entre la química cuántica, que sólo puede tratar un número pequeño de átomos, y la mecánica estadística, en particular, las simulaciones numéricas.

2.4 SIMULACIONES NUMÉRICAS

2.4.1 Consideraciones generales

La ambición de cualquier teoría de un sistema de muchos cuerpos es relacionar los detalles microscópicos de las interacciones atómicas y moleculares al comportamiento de las funciones termodinámicas observadas. La mecánica estadística se encarga de esto en principio, pero las dificultades matemáticas que aparecen al tratar Hamiltonianos realistas son formidables. La mecánica estadística matemática sólo puede tratar modelos muy sencillos de fenómenos intrínsecamente complejos, y en el proceso de modelado se pierde toda la complejidad que se quiere estudiar.

Las simulaciones numéricas son una alternativa reciente (las primeras simulaciones se reportaron en 1953 y 1957 para MC y DM respectivamente) que permite calcular en forma aproximada todo lo que la mecánica estadística matemática pretende encontrar de

una manera más elegante, pero inefectiva. Las simulaciones no tienen restricciones en la complejidad del Hamiltoniano, y las restricciones en el número de partículas y el tiempo de simulación van haciéndose menores a medida que la capacidad de cómputo crece.

Las técnicas usadas en la simulación numérica son de dos tipos: Monte Carlo (MC) y dinámica molecular (DM). Su característica común es el hecho de estar basadas en una descripción molecular del sistema de interés. El ingrediente principal en ambos casos es la ley que describe las interacciones entre las partículas constituyentes, ya sean átomos, moléculas o iones. En este esquema de simulación global dichas "leyes" se obtienen a partir de datos *ab initio*.

Las simulaciones pueden ser útiles en las ciencias biológicas: proveen explicaciones teóricas a nivel molecular de fenómenos observados experimentalmente, y predicen valores de cantidades que no se han podido medir pero que son de interés. Las simulaciones dan información que no se puede obtener de otra manera. Para que estas predicciones sean confiables, primero hay que mostrar que el método es capaz de reproducir los datos que sí se pueden medir.

Las propiedades que se pueden obtener de una simulación se dividen en dos grupos dependiendo de la fuente de datos para calcular los promedios:

- propiedades mecánicas: éstas son cantidades que se pueden expresar como función de las coordenadas y momentos de las partículas, y son promedios de ensemble. Entre estas propiedades encontramos la energía potencial y cinética, la presión, las funciones de distribución radial y angular, los factores de estructura y las funciones de correlación directas y cruzadas de posiciones, velocidades y momentos dipolares. De éstas últimas se obtienen propiedades de transporte (coeficientes de difusión, viscosidad y conductividad) y de relajación. De las fluctuaciones de la energía cinética y de la densidad se calculan el calor específico y la compresibilidad.

- propiedades térmicas: éstas dependen del volumen total del espacio fase accesible al sistema. Aquí encontramos la entropía, la energía libre de Gibbs o de Helmholtz, el potencial químico, y todas las cantidades que se derivan de éstas, como son las constantes de asociación y disociación, solubilidades, etc. El cálculo de estos promedios sobre el espacio fase requiere de técnicas especiales que aún están en proceso de validación²³⁻²⁵.

Antes de describir los métodos de MC y DM es conveniente discutir los problemas conceptuales comunes a ambos métodos.

2.4.1.1 La validez de las simulaciones clásicas

El comportamiento a temperatura ambiente de líquidos moleculares como el agua o de macromoléculas en solución se puede describir a primera aproximación con un tratamiento clásico. Sin embargo, se sabe que la mecánica cuántica es esencial en el tratamiento de las vibraciones moleculares y de los movimientos intermoleculares de los líquidos fuertemente asociados como el agua líquida. Estos efectos cuánticos se deben a la masa del H, que es pequeña en comparación a la del O, y esto resulta en frecuencias vibracionales muy

altas, tanto inter como intramoleculares. El efecto total no es muy notable porque hay una cancelación extensa de los efectos intra e intermoleculares²⁹.

El procedimiento estándar para manejar los efectos cuánticos en las simulaciones clásicas es corregir las propiedades que dependen de estos efectos, como son la energía total y la capacidad calorífica. Hay dos maneras de hacer estas correcciones^{26,27}, y cabe señalar que no siempre dan los mismos resultados.

El uso de las correcciones cuánticas no es generalizado, en parte porque las dos maneras de calcularlas no son consistentes. Además, parece ser que sólo se aplican si los resultados clásicos son incapaces de reproducir el experimento.

2.4.1.2 El ensemble de simulación

La elección del ensemble se vuelve irrelevante en el límite termodinámico, pero esto no es necesariamente cierto para los sistemas pequeños que se usan en las simulaciones. Los ensembles típicos para MC y DM son el canónico (NVT) y el microcanónico (NVE) respectivamente, pero hay muchos fenómenos físicos que se pueden discutir de manera más adecuada en otro ensemble. Se facilita la comparación de los resultados con los datos experimentales.

Los ensembles utilizados en MC son NVT, NpT y μ VT²⁸. Para los sistemas de una sola especie molecular, el ensemble NVT es el más adecuado, pero para simular mezclas y cambios estructurales en sólidos, NpT y μ VT son mejores. Lo que se hace al cambiar de ensemble es muestrear, además de las posiciones de las partículas, la variable conjugada al nuevo parámetro fijo. En el caso de NpT, se trata del volumen, que puede cambiar de tamaño y de forma. Los cambios de forma de la celda computacional hacen posible el estudio de transiciones estructurales en sólidos. En el caso de μ VT, la variable conjugada es el número de partículas en la celda. Este ensemble es muy útil para calcular energías libres, pero tiene la desventaja de converger muy lentamente para sistemas densos. Esto se debe a que la mayoría de los intentos de incluir partículas a la caja son rechazados.

Para la DM, la extensión a otros ensembles es más difícil. El ensemble natural para las ecuaciones de Newton es el NVE, de manera que al cambiar de ensemble hay que escribir un equivalente a las ecuaciones de Newton. Lo que se hace es incluir grados de libertad extra, como el volumen, por ejemplo. Un cambio en volumen corresponde a un escalamiento homogéneo de todas las posiciones de centro de masa molecular. Esto implica nuevas masas y momentos también. En este caso ya no se conserva la energía, sino algo muy parecido a la entalpía. Este es el ensemble NPH. Los ensembles NpT y NVT ya no son deterministas porque se acopla al sistema a un baño térmico. Hay varias formas de hacer esto. Una es escalar el tiempo y/o las masas, otra es incluir un término estocástico que modifique la velocidad de una partícula a la vez y la otra es incluir un término de fricción dinámica. Estos nuevos ensembles (NVT, NpT y NPH) permiten simular las fluctuaciones del medio, pero no dan la dinámica verdadera del sistema físico porque hay un factor inercial por cada grado de libertad que se agrega.

2.4.1.3 El número de partículas simuladas

La situación ideal sería poder simular un número de partículas cercano al número de Avogadro ($\approx 10^{23}$), para tener resultados en el límite termodinámico. Obviamente, por mucho que aumente la velocidad de las computadoras y mejoren los algoritmos de programación, esto es irrealizable. Con la capacidad de cómputo actual, cuando mucho se pueden simular $\approx 10^5$ partículas, y estas simulaciones se hacen con potenciales sencillos y de corto alcance, de manera que sólo se calculan interacciones con los primeros o segundos vecinos. A nosotros nos interesa usar un potencial realista, y para el agua, éste es de largo alcance.

Clementi^{29,30} propone el uso de ≈ 500 moléculas de agua para describir estadísticamente el agua líquida a temperatura ambiente. Este número depende del tamaño de la celda computacional a través de la densidad a la que se desea hacer la simulación, y el tamaño de la celda se determina por el alcance de los potenciales que se utilicen. Como el costo computacional sigue siendo elevado para un sistema de este tamaño, se ha buscado el número mínimo de moléculas que permita obtener resultados confiables. Clementi y Corongiu³¹ hicieron una comparación de simulaciones con $N = 64, 125$ y 343 moléculas de agua. La conclusión es que $N < 100$ no es adecuado para simular agua líquida. Por supuesto, mientras más moléculas contenga la celda computacional la estadística será mejor y los resultados también. Resultados recientes obtenidos en el laboratorio indican que, aún en simulaciones de ≈ 1000 moléculas de agua, las propiedades calculadas son dependientes del tamaño del sistema.

2.4.1.4 Las fronteras del sistema y el corte de los potenciales de interacción

Nosotros estamos interesados en la simulación de sistemas polares. Como el alcance del potencial de interacción es mayor que el tamaño del sistema simulado, la parte de largo alcance del potencial tiene que ser incluida de forma aproximada o despreciada de alguna manera sistemática. No hay un tratamiento analítico que indique cuál esquema es el adecuado en el límite termodinámico, así que lo que se ha hecho son ensayos con diferentes métodos.

Dependiendo del método que se vaya a usar, hay dos maneras de manejar el hecho de tener una celda computacional finita y pequeña. La primera consiste en rodear a esta celda con una envoltura de celdas idénticas (26 para el caso de una celda cúbica en tres dimensiones), de manera que si una partícula sale por la pared inferior de la celda, una imagen suya aparezca por la parte superior. Lo mismo se aplica para cada pareja de caras paralelas de la celda. A esto se le llama condiciones toroidales a la frontera (TBC), y genera un sistema periódico, sin bordes, pero finito. La segunda consiste en replicar la celda computacional un número infinito de veces en todas direcciones (x, y y z en el caso de la celda cúbica). El tratamiento para las partículas que se salen de la celda es igual que para TBC. A esto se le llama condiciones periódicas a la frontera (PBC), y genera un sistema periódico infinito. La diferencia entre TBC y PBC es que la primera genera un sistema finito y la segunda, un sistema infinito. Además, para poder aplicar la corrección llamada "sumas de Ewald", se requiere usar PBC. Una tercera solución para simulaciones

de líquidos es simular una gota, pero para esto se necesitan muchas más moléculas que para TBC y PBC porque la cantidad de superficie es muy grande. Las simulaciones de macromoléculas se hacen con TBC o en el vacío.

Tanto TBC como PBC presentan la deficiencia de generar sistemas periódicos, y sabemos que los líquidos no son periódicos. Hasta ahora no se ha propuesto ninguna otra manera de eliminar los bordes. Además de la periodicidad artificial, existe otro problema más sutil que puede verse claramente en sistemas iónicos diluidos. El potencial se corta de manera que una molécula no interactúe con su imagen, pero sí interactúa con imágenes de otras moléculas que la están viendo en su celda. Esta situación no es igual a la que experimenta un ion real en una solución acuosa diluida. El ion real "ve" un solvente puro a su alrededor, mientras que el ion de la simulación "ve" un solvente con la influencia de otro ion cercano. Este es uno de los detalles que complica el concepto de concentración en las simulaciones.

El tamaño de la celda también limita la longitud de las correlaciones espaciales; por esto se requieren técnicas especiales para estudiar transiciones de fase, en las cuales las correlaciones tienen un rango infinito²³.

Cuando el potencial es de corto alcance, como el de van der Waals que cae como r^{-6} , las energías y las fuerzas casi siempre se calculan con la convención del corte esférico. El potencial se hace cero más allá de una distancia de corte que típicamente es de tres o cuatro diámetros de las partículas. Después se puede considerar el efecto de despreciar las interacciones de mayor alcance en las propiedades del sistema.

Si el potencial cae más lentamente que r^{-3} en tres dimensiones, el corte esférico no es apropiado. En sistemas iónicos como las sales fundidas se viola la electroneutralidad y la estructura predicha no es correcta. Para estos sistemas se usa el método de Ewald³² que fue diseñado para sólidos cristalinos. La transformación de Ewald consiste en expresar la energía coulombiana total por celda de un arreglo periódico infinito como la suma de dos series rápidamente convergentes. Una serie es la suma en el espacio real de un potencial de corto alcance. La otra suma se hace sobre vectores de onda en el espacio recíproco, comensurables con las condiciones periódicas a la frontera. Este método exagera la periodicidad de por sí artificial del sistema. Para tratar sistemas iónicos diluidos se puede hacer un corte esférico y luego hacer la corrección de Born³³.

La interacción dipolar cae como r^{-3} , así que la aproximación de corte esférico es menos grave. Un problema adicional es que esta interacción tiene una dependencia angular. Esta combinación hace que la interacción dipolar sea difícil de tratar, y es responsable de las correlaciones entre dipolos distantes. Hay dos problemas en particular: el cálculo de propiedades dieléctricas y que el sistema no conserva la energía.

La constante dieléctrica de un líquido polar está relacionada a las fluctuaciones de longitud de onda larga de la densidad de momento dipolar de la muestra. Estas fluctuaciones dependen fuertemente de las condiciones de frontera, o en el contexto de las simulaciones, de la manera en que se trata el término dipolar. Cortar el potencial tiene el efecto de

eliminar las fluctuaciones de longitud de onda larga, y el resultado de esto es que ciertos componentes angulares de la función de distribución por pares son muy distintos de los que se encontrarían en un sistema infinito.

En cuanto a la no conservación de la energía, el corte del potencial es responsable de pequeñas fluctuaciones en la energía cuando pares de partículas cruzan sus fronteras de corte. Para potenciales isotrópicos, como aquellos entre cargas libres, el corte del potencial no causa una deriva global en la energía. Sin embargo, cuando el potencial depende de los ángulos de interacción, el cruce y recruce de la frontera no se cancela exactamente. Cuando un par de dipolos está cerca de la frontera de corte, hay una cierta correlación entre sus orientaciones. Cuando se alejan más y dejan de interactuar la correlación desaparece con el tiempo. Al momento de volver a interactuar, en general se encontrarán en una configuración con energía menos favorable que la de la configuración con la que dejaron de interactuar. La reorientación para hacer la interacción más favorable inyecta una pequeña cantidad de energía al sistema, y esta energía es totalmente artificial. De hecho se ha violado la primera ley de la termodinámica, ya que, por haber cortado el potencial de interacción, se tiene un ciclo en el que se gana energía interna gratuitamente. La cantidad de energía ganada por ciclo de salida y entrada a la esfera de interacción es casi despreciable, pero las simulaciones son largas, y la suma ya no es despreciable.

Para solucionar estos dos problemas hay dos métodos: uno es el de Ewald³², ya sea en su forma original o en una forma que considera dipolos^{34,35}. El segundo método es el de campo de reacción (*reaction field*)³⁶, que tiene una justificación física más clara. Se usa el corte esférico, pero ahora la región externa a la esfera de interacción se considera como un medio continuo dieléctrico en lugar de considerarlo como el vacío. La presencia del continuo da lugar a un término en la energía potencial de interacción entre el dipolo de una molécula y el campo que actúa sobre la molécula como resultado de la polarización del continuo por el momento dipolar neto de la esfera de corte centrada en esa molécula. La inclusión de este término tiene un efecto dramático en las correlaciones angulares³⁷.

Hay otras dos soluciones propuestas para manejar interacciones de largo alcance. Una es cortar las fuerzas (esto es particular de DM), y es muy mala porque este sistema ya no es conservativo³⁸. La otra es hacer un corte suave del potencial o la fuerza añadiendo un polinomio en la zona de corte que tenga sus primeras dos derivadas continuas. El polinomio hace que el potencial caiga a cero rápidamente pero de forma continua. Este ataque tiene el problema de producir oscilaciones artificiales en la energía, y el sistema puede quedarse atrapado en un mínimo local que ni siquiera es real (P.J. Steinbach, comunicación personal).

No todas las propiedades del sistema son igualmente sensibles al tratamiento del corte del potencial. Lo que se ha encontrado comparando los distintos esquemas es que las propiedades más sensibles son aquellas que dependen de la correlación angular. La energía interna y las otras propiedades que dependen de la distribución atómica son menos sensibles³⁸.

Dado que los potenciales coulombianos tienen simetría esférica, en ocasiones es preferible usar celdas periódicas que son "más esféricas" que la celda cúbica. La celda más esférica que es capaz de llenar el espacio es el octaedro truncado. El reemplazar una celda

cúbica por una no cúbica puede hacer más lenta la simulación, pero la alternativa es simular una celda cúbica más grande, que también es más lento. Las ventajas específicas de las celdas "casi esféricas"²⁸ son que, para el mismo número de partículas, las funciones de distribución radial se pueden calcular para distancias mayores que en la celda cúbica, que hay menos distorsión de los potenciales con dependencia angular y que las simulaciones de macromoléculas requieren menos moléculas de solvente.

2.4.1.5 La longitud de la simulación

Las simulaciones actuales tratan procesos en tiempo real de decenas de ps a ns, dependiendo del tamaño y la complejidad del sistema. Si se quieren estudiar procesos que se dan en intervalos temporales mayores hay que simplificar el modelo molecular y/o discretizar el espacio. El tiempo real accesible a la simulación depende de la frecuencia de vibración más alta del sistema. Si se quieren modelar vibraciones de enlaces, los tiempos propios de estos movimientos son de $\approx 10^{-14}$ s, y por lo menos hay que usar unidades de tiempo dos órdenes de magnitud menores. En MC, como no existe la variable temporal, esto se hace reduciendo los desplazamientos de prueba. Para simular tiempos reales más largos lo que se hace es congelar los grados de libertad con frecuencias más altas. Esto sólo es válido si estos grados de libertad congelados no afectan el movimiento global de las moléculas. En el caso de las macromoléculas se elimina la vibración de los enlaces y se conservan las rotaciones. Existe un método relativamente reciente³⁹ que permite realizar simulaciones mucho más largas. Elimina el problema de rigidez (*stiffness*) del sistema, haciendo que se pueda usar una unidad de tiempo tan grande como uno quiera. Lo que sucede es que una vez más el sistema ha dejado de ser conservativo y no se tiene la dinámica real del sistema.

Antes de poder recolectar datos de la simulación, el sistema tiene que estar equilibrado, y el tiempo para llegar al equilibrio depende de los procesos de relajación del sistema. Por ejemplo, para el agua líquida se requieren al menos 10ps, porque este es el tiempo de relajación rotacional de una molécula de agua en la fase líquida.

Para que la simulación sea válida habría que demostrar que el sistema que se está simulando es ergódico, y esto sólo se puede ver con corridas muy largas (varias decenas de millones de pasos de MC o centenas de miles de pasos de DM). Simulaciones por MC de agua líquida realizadas por H.Saint-Martin indican que si la simulación no es suficientemente larga el sistema tiende a quedarse atrapado en mínimos locales, dando la apariencia de estar equilibrado.

Una vez revisados estos puntos comunes a los dos tipos de simulación, se presentan en detalle los métodos.

2.4.2 Monte Carlo

2.4.2.1 Descripción

El método de MC se propuso⁴⁰ como un método general para calcular las propiedades de cualquier sustancia que pueda ser considerada como compuesta de moléculas individuales en interacción. El método original supone una estadística clásica y se desarrolló para el ensemble canónico (NVT).

El valor de equilibrio en el ensemble NVT de cualquier cantidad mecánica F está dada por

$$\bar{F} = \frac{\int F \exp\left(-\frac{E}{kT}\right) d^{3N}p d^{3N}q}{\int \exp\left(-\frac{E}{kT}\right) d^{3N}p d^{3N}q}$$

en tres dimensiones, donde E es la energía del sistema cuando se encuentra en un punto del espacio fase determinado por el conjunto de p 's y q 's de sus N partículas, y $d^{3N}p d^{3N}q$ es el elemento de volumen en el espacio fase $6N$ dimensional.

Si las fuerzas entre las partículas son independientes de las velocidades de las partículas (se trata de un sistema conservativo), las integrales sobre los momentos se pueden separar e integrarse analíticamente. Una vez hecha esta separación, sólo falta evaluar la integral en el espacio de configuración $3N$ dimensional. Esta integral sólo tiene solución analítica para Hamiltonianos triviales, como el de gas ideal. Para cualquier potencial que dependa de las distancias entre dos o más partículas la única alternativa es hacer una integración numérica; como el espacio sobre el que hay que integrar es multidimensional, se usa el método de MC⁴¹. En lugar de integrar sobre un arreglo regular de puntos, este método integra sobre un muestreo aleatorio de puntos.

Un ejemplo en dos dimensiones de cómo funciona este tipo de integración es el siguiente. Supóngase que se quiere calcular el valor del área bajo una curva localizada entre $[0,x]$ y $[0,y]$ y de la cual no se tiene una expresión analítica. Lo que hace este método es generar al azar una colección muy grande de puntos distribuidos uniformemente en toda el área encerrada entre $[0,x]$ y $[0,y]$; el área total del rectángulo que contiene al área que se quiere calcular se conoce (es simplemente $x \times y$). Se suman los puntos que cayeron bajo la curva y el área bajo la curva se calcula con el cociente de esta suma y el número de puntos totales, todo multiplicado por el área total.

En este ejemplo en dos dimensiones pueden verse dos características muy importantes de este método. La primera es que los puntos tienen que muestrear de manera uniforme todo el espacio accesible, y la segunda es que, para tener resultados precisos se requiere de un número muy grande de puntos. El método es exacto si se genera un número infinito de puntos.

Para integrar en espacios de baja dimensión existen métodos mucho más eficientes y precisos, como los de Runge-Kutta, pero para sistemas multidimensionales, los métodos tradicionales son impracticables y éste es uno de los métodos más utilizados.

Regresando al problema molecular, la manera directa de implementar el algoritmo de

MC es calcular la energía E de una colección de configuraciones y darle a cada configuración un peso $\exp(-\frac{E}{kT})$. Esto es equivalente a la formulación de Gibbs de la mecánica estadística. Sin embargo, esta implementación no es eficiente para configuraciones densamente empacadas (que son típicas en la materia condensada), ya que hay una probabilidad muy alta de escoger una configuración para la cual $\exp(-\frac{E}{kT})$ es muy pequeña y que por lo tanto tiene un peso muy bajo. Lo que proponen Metropolis et al⁴⁰ es escoger las configuraciones con la probabilidad $\exp(-\frac{E}{kT})$ y luego sumarlas con un peso igual. A esto se le llama "muestreo de importancia" y se hace de la siguiente manera: se colocan las N partículas en cualquier configuración (mientras más cercana al equilibrio, mejor) y se mueve cada una de las partículas en forma sucesiva generando una cadena Markoviana. Se escoge al azar la partícula que se va a mover y cuánto se va a mover. Después de mover una partícula se calcula el cambio en energía del sistema ΔE causado por el movimiento. Si $\Delta E \leq 0$ se acepta la movida y se coloca la partícula en su nueva posición. En este caso el movimiento lleva al sistema a un estado de energía más baja. Si $\Delta E > 0$, se acepta la movida con el siguiente criterio: se toma un número al azar ξ entre 0 y 1. Si $\xi < \exp(-\frac{\Delta E}{kT})$ se acepta la movida, y si es mayor o igual se regresa la partícula a la posición de donde partió. Este criterio permite remontar barreras de potencial. La magnitud promedio de las barreras que se pueden cruzar depende de la temperatura de la simulación. A 0K no se aceptan movimientos que tengan $\Delta E > 0$. El número de configuraciones que entran al cálculo del promedio de la propiedad de interés es la suma de configuraciones aceptadas y rechazadas (M):

$$\bar{F} = \frac{1}{M} \sum_{i=1}^M F_i$$

donde F_i es el valor de la propiedad F del sistema después de llevar a cabo la movida i de acuerdo con el algoritmo descrito arriba.

Se puede demostrar^{40,42} que este algoritmo es consistente con un sistema ergódico (que cualquier estado es accesible desde cualquier otro en un número finito de pasos), y que después de un número grande de pasos el sistema simulado tiende a tener una distribución canónica de energías. Sin embargo, no hay nada que indique qué tan rápido se llega a esta distribución canónica; para el agua líquida en una celda computacional con $N = 343$ moléculas esto puede necesitar del orden de 10^7 pasos. La magnitud de los desplazamientos de prueba tiene que escogerse con cuidado. Si es demasiado grande, la mayoría de las movidas es rechazada, y si es demasiado chica, la configuración no cambia lo suficiente. En cualquiera de los dos casos la convergencia al equilibrio es más lenta. En general se busca un cociente de movidas aceptadas sobre rechazadas de 1. El sistema llega al equilibrio cuando los valores de la energía se estabilizan. Algunas veces se llega a observar esta condición sin que el sistema se haya equilibrado realmente; si se generan varios millones de configuraciones más se observa un descenso de la energía, lo cual pareciera indicar que la energía evoluciona a través de mesetas.

Aquí se ha discutido el MC clásico en el ensemble NVT; la implementación de MC cuánticos o en otros ensembles se encuentra en la referencia 28.

2.4.2.2 Dificultades de aplicación

Además de los problemas ya descritos, el MC tiene dos problemas particulares. El primero es que los promedios sobre la distribución de probabilidad que describen el equilibrio térmico se reemplazan por sumas sobre un conjunto finito de puntos. Por lo tanto, la aproximación al resultado "real" es mejor mientras más larga es la simulación. El segundo es que la cadena de números aleatorios que se usa para construir la cadena Markoviana en la simulación en realidad está compuesta de números pseudoaleatorios. Hay que estar seguros de que esta cadena no se repita en el curso de la simulación; de repetirse, se está haciendo un muestreo incompleto de las configuraciones accesibles al sistema. Recordando el ejemplo bidimensional, esto equivale a generar puntos en una zona limitada del área total. La consecuencia es que la estadística está sesgada y ya no es válida.

2.4.2.3 Información obtenida

El MC permite calcular todas las propiedades estáticas de un sistema, o sea, todas las propiedades que dependen de las coordenadas de las partículas y no de sus momentos. No es capaz de proporcionar información dinámica porque la variable temporal no aparece en este tratamiento, y porque la parte cinética es constante en los ensembles que se usan (NVT, NpT y μ VT).

Frente a la DM, presenta dos ventajas en el estudio de los líquidos: converge más rápidamente porque nada más tiene que equilibrar la distribución espacial de las partículas, mientras que la DM tiene que equilibrar la parte de las coordenadas y la de los momentos. Sin embargo, en el caso de simulaciones de proteínas y otras macromoléculas, algunos autores prefieren la DM²³. Esto se debe a que para sistemas complejos con muchos átomos unidos covalentemente sólo se puede obtener un cociente razonable de aceptación de movidas con pasos muy pequeños, y esto hace que la simulación por MC requiera muchos más pasos que una simulación por DM²³.

2.4.3 Dinámica Molecular

2.4.3.1 Descripción

La DM es una técnica más poderosa que el MC porque es capaz de proporcionar información tanto estática como dinámica. Esto se debe a que el MC tiene acceso al espacio de configuración mientras que la DM tiene acceso al espacio fase.

La DM fue propuesta⁴³ como un método para estudiar las relajaciones que acompañan varios fenómenos fuera del equilibrio. El método original resuelve exactamente el sistema de ecuaciones simultáneas clásicas de movimiento para varios cientos de partículas. Se pueden resolver exactamente porque los autores usan el potencial de pozo cuadrado. En este caso la fuerza es cero salvo cuando hay colisiones, ya sean atractivas o repulsivas. La trayectoria en el espacio fase en una simulación con un potencial cuya derivada no es cero entre colisiones se aproxima por saltos discretos en los vectores de posición y velocidad de las N partículas sobre intervalos temporales cortos.

La técnica consiste en obtener una solución en la computadora de la evolución temporal de un número finito de átomos o moléculas totalmente aislado de su entorno. El ensemble correspondiente es el microcanónico (NVE), así que la solución traza la trayectoria de un punto representativo restringida a moverse en una superficie de energía constante en el espacio fase clásico. Los valores de expectación de propiedades termodinámicas de equilibrio se obtienen de la manera propuesta por Boltzmann. Si $F(\mathbf{p}, \mathbf{q})$ es cualquier propiedad definida en el espacio fase, el promedio de ensemble $\langle F \rangle$ es el promedio sobre un tiempo infinito:

$$\langle F \rangle = \lim_{a \rightarrow \infty} \frac{1}{a} \int_0^a F(\mathbf{p}, \mathbf{q}) dt$$

donde a es el número de puntos de la trayectoria temporal. Este promedio es igual al promedio del ensemble microcanónico si el sistema es ergódico⁴⁴:

$$\langle F \rangle = \frac{\int \cdots \int \delta(H_N(\mathbf{p}, \mathbf{q}) - E) F(\mathbf{p}, \mathbf{q}) d\mathbf{p} d\mathbf{q}}{\int \cdots \int \delta(H_N(\mathbf{p}, \mathbf{q}) - E) d\mathbf{p} d\mathbf{q}}$$

$H_N(\mathbf{p}, \mathbf{q})$ es el Hamiltoniano del sistema.

Por supuesto, el límite $a \rightarrow \infty$ no se puede tomar, y en la práctica se reemplaza por valores muy pequeños de a (del orden de 10^5 o 10^6 en el mejor de los casos). Como a no es infinito, es posible que el punto representativo se quede atrapado en un dominio pequeño de la superficie completa y se rompa la ergodicidad. Esto es más probable en una simulación por DM que en una por MC.

Dado un potencial, la evolución del sistema se encuentra integrando numéricamente las ecuaciones de Newton, que en tres dimensiones y para N partículas se convierten en $6N$ ecuaciones simultáneas de primer orden. El control sobre la solución numérica de varios miles de ecuaciones simultáneas se puede lograr monitoreando constantes de movimiento. En general se utiliza el comportamiento de la energía total. Otra manera de comprobar que la integración es estable es cambiar la dirección del paso temporal y ver si el sistema regresa por la misma ruta. Esto no se hace en la práctica de las simulaciones, pero sí en la prueba de los diferentes algoritmos de integración. El algoritmo de Verlet⁴⁵ independiente de velocidades es el más usado en simulaciones de partículas esféricas.

Todo lo que se ha discutido hasta ahora es válido para partículas monoatómicas. Si se quiere simular moléculas diatómicas, los grados de libertad aumentan de $3N$ a $5N$; los nuevos grados de libertad son los ángulos que forma el eje molecular con el marco de referencia Cartesiano localizado en el centro de masa de la molécula. Para moléculas triatómicas como el agua, el número de grados de libertad es de 6 por molécula. La complicación se extiende también a los potenciales. Si se quiere un potencial molécula - molécula se tienen que incluir términos con dependencia angular. El intervalo temporal de integración tiene que reducirse un orden de magnitud respecto al usado en simulaciones de partículas sin estructura si se simulan núcleos ligeros. Esto se debe a las libraciones de estos núcleos ligeros.

Si se simulan moléculas rígidas hay que resolver las ecuaciones de Newton - Euler. Por un lado se tienen las ecuaciones de Newton que describen el movimiento del centro de masa,

y por el otro se tienen las ecuaciones de Euler que describen la rotación de la molécula. El cálculo del vector de momento total en el marco de referencia molecular depende de los puntos sobre los que actúan las fuerzas en la molécula. Una vez que se calculan los componentes de velocidad angular se resuelven seis ecuaciones por molécula, así que en total se tienen que resolver 12 ecuaciones por molécula. Existen métodos numéricos muy poderosos para resolver estos sistemas de ecuaciones⁴⁶.

Una simulación típica de N moléculas de agua comienza, preferentemente, con una configuración equilibrada en una simulación por MC. El primer paso es equilibrar la muestra a partir de esta configuración inicial. Una sugerencia para las velocidades iniciales es ponerlas todas en cero. Como en general el sistema no está en un mínimo de energía potencial, el exceso de energía potencial hace que se aceleren las moléculas. La aceleración puede llegar a calentar al sistema a temperaturas de 10^5 K. El sistema se enfría varias veces hasta que la temperatura converge cerca del valor al que se quiere hacer la simulación. Hay varias formas de monitorear el equilibrio del sistema. La forma propuesta por Boltzmann es observar la función de distribución de velocidades. En el equilibrio, ésta debe ser igual a la distribución de Maxwell-Boltzmann. Otra opción es calcular las temperaturas traslacionales (del movimiento de los centros de masa) y rotacionales y ver qué tan diferentes son. Al principio de la simulación estas dos temperaturas están desacopladas y dan valores que en general son muy dispares; a medida que la simulación avanza, los modos energéticos se acoplan y en el equilibrio estas dos temperaturas prácticamente no difieren. Si se está trabajando con un formalismo átomo - átomo un ataque equivalente es calcular la temperatura de las diferentes especies atómicas; en el equilibrio deben ser muy parecidas.

Una vez que el sistema se ha equilibrado se guardan los datos del espacio fase x y y para cada partícula para cada paso temporal.

2.4.3.2 Dificultades de aplicación

Además de los problemas ya discutidos de no conservación de la energía por el corte de los potenciales de interacción, en la DM hay otra fuente de error que conduce a la no conservación de la energía total. Esta es que se está aproximando una trayectoria continua por una trayectoria discreta, y la suposición más fuerte que se hace a este respecto es que la fuerza es constante entre un punto y el siguiente. Para que esta suposición sea válida el paso de integración temporal Δt tiene que ser pequeño. Qué tan pequeño depende de la forma del potencial. Si el sistema que se quiere simular tiene átomos ligeros como los H, Δt típicamente es de 10^{-16} s, se incluyan vibraciones intramoleculares o no. Este paso tan pequeño se debe a las libraciones (rotaciones impedidas) de las moléculas.

Usar pasos tan pequeños limita considerablemente el tiempo real que se puede simular. Con un millón de pasos se puede llegar a 0.1 ns, y este es el límite actual en las simulaciones de agua líquida. Para compararlo con simulaciones por MC, esto equivale al menos a N millones de pasos de MC para un sistema con N moléculas. Como N es del orden de 10^2 , estamos hablando de 10^8 configuraciones de MC.

En las simulaciones de macromoléculas se hacen muchas aproximaciones para poder simular tiempos más largos. Recordemos que el proceso de doblado de una proteína dura un lapso de ms, tiempo que está muy lejos de poder ser simulado si se incluyen todos los átomos. Una aproximación muy común que está cayendo en desuso pero que es interesante mencionar es la de átomos unidos. Lo que se hace es tomar como unidad al átomo pesado y los H que están covalentemente unidos a él, asignando un radio de van der Waals al conjunto. Esta aproximación refleja los prejuicios de los cristalógrafos de rayos X, porque en esta técnica no se pueden localizar los H. Como ya no se tratan explícitamente los H, Δt puede ser de 10^{-12} s. Otra aproximación casi universal es restringir las longitudes y ángulos de enlace, permitiendo sólo las rotaciones (SHAKE⁴⁷).

2.4.3.3 Información que proporciona

Como ya se mencionó, la DM tiene acceso a una superficie de energía constante en el espacio fase. La consecuencia de esto es que permite el cálculo de cualquier propiedad del sistema que se pueda expresar en términos de las coordenadas y/o momentos de las partículas. En este caso se encuentran todas las propiedades mecánicas del sistema. El cálculo de las propiedades térmicas no es tan sencillo, porque hay que realizar varias DM para sistemas casi iguales de manera que uno abandone la superficie de energía constante y pueda muestrear el espacio fase²⁴.

Habiendo descrito la metodología en general, a continuación se presenta su aplicación al estudio molecular del agua líquida.

CAPÍTULO 3: RESULTADOS Y DISCUSIÓN

3.1 OBTENCIÓN DE UN POTENCIAL INTRAMOLECULAR

En cualquier simulación numérica el potencial de interacción entre las moléculas de interés es el elemento más importante, ya que los resultados que se obtienen en la simulación se derivan de él. En el laboratorio se desarrolló un potencial intermolecular agua-agua (MCHO) que incluye efectos no aditivos de muchos cuerpos y que ha probado reproducir con muy buena aproximación diversas propiedades del agua líquida en simulaciones por Monte Carlo (MC)¹. Este potencial modela los átomos de la siguiente manera: toda la masa del átomo se encuentra concentrada en una esfera que hace las veces de núcleo, y la carga del átomo se localiza en un punto sin masa externo a esta esfera; este punto se encuentra unido a la esfera por un potencial armónico. Las posiciones de las cargas dependen del campo eléctrico de su entorno, y esto es lo que permite reproducir la polarizabilidad y la no aditividad de muchos cuerpos.

Para tratar la parte intramolecular calculamos una superficie vibracional para un monómero de agua. El suponer que con calcular monómeros es suficiente implica que los modos de vibración intra e intermoleculares no están acoplados, o que el acoplamiento es tan débil que es despreciable. La mayoría de los potenciales vibracionales que existen para el agua suponen esto²⁻¹⁵. Hay una familia de potenciales que no hacen distinción entre la parte inter y la intramolecular¹⁶⁻²¹ y un potencial que hace una corrección a nivel HF-SCF a la superficie MCY²⁶ para la interacción de dímeros de moléculas de agua deformadas^{22,23}

La superficie vibracional se calculó con el programa PSHF²⁴ utilizando para el O una base propuesta por van-Duijneveldt²⁵ 10S,8P contraídas a 5S,5P + 2D y optimizada para pseudopotenciales por Leticia Gil. Para el H utilizamos una base de Clementi²⁴ 6S + 2P (base QZ2P). Las bases utilizadas en la tesis están optimizadas para pseudopotenciales y se muestran en el Apéndice A. La superficie calculada además incluye correlación electrónica a nivel CI-MP2 calculada con el programa CIPSI²⁷ (ver sección 2.2.1.2.4). El criterio de selección de determinantes para diagonalización fue de 0.015 para toda la superficie y las normas se encontraron entre 0.0336 y 0.0497, lo cual habla de que el nivel de cálculo es aceptable.

Como el potencial vibracional se añadió al MCHO, su contribución a la energía total debe ser cero si las moléculas no se deforman. Dicho de otra manera, la deformación molecular debe ser tratada como una perturbación. Por esto la superficie se definió en función de $\Delta r = r - r_0$ para cada enlace, donde r es la longitud del enlace en la molécula deformada y r_0 es la longitud correspondiente a la geometría de equilibrio. Se consideraron variaciones en la longitud de los enlaces OH (Δr entre -0.1928 y +0.2702 ua) y en la distancia HH (Δr entre -0.5669 y +0.5669 ua), calculándose un total de 79 puntos, 67 geometrías simétricas y 12 asimétricas. La energía de deformación se calculó como $\Delta E = E - E_0$, donde E es la energía a nivel CI-MP2 del monómero deformado y E_0 es la energía a nivel CI-MP2 del monómero en la geometría experimental. Se calcularon puntos con ΔE tan repulsiva como 33.26 kcal/mol.

3.2 AJUSTE DEL POTENCIAL VIBRACIONAL

Una vez obtenida la superficie vibracional (79 puntos), es necesario proponer una forma funcional y buscar el ajuste de parámetros que de una reproducción adecuada de esta superficie. La forma funcional que propusimos está en función de la longitud de los enlaces OH y de la distancia HH. Decidimos usar la distancia entre los H en lugar del ángulo entre los enlaces OH porque en la simulación es más rápido calcular distancias que ángulos, y además se obtiene un acoplamiento gratuito entre la longitud de los enlaces y el ángulo que forman.

Como en todo proceso de modelado, comenzamos con el modelo físico más sencillo para un enlace no rígido: un oscilador armónico por cada tipo de enlace. El ajuste obtenido con esta forma analítica fue muy pobre (el error cuadrático, que mide la bondad de un ajuste -mientras más cercano a cero es mejor-, era del orden de 10^7), por lo que fue necesario dar mayor flexibilidad al potencial. De esta manera incluimos términos de orden cúbico y cuártico, acoplamos los osciladores a segundo y tercer orden y también probamos osciladores de Morse, que son anarmónicos. Ningún esquema funcionó satisfactoriamente. Eliminamos de la superficie los puntos asimétricos que dieron muchos problemas para ajustarse y aquellos puntos con $\Delta E > 15$ kcal/mol, pero el ajuste tampoco mejoró.

Cabe señalar que el ajuste de un potencial del tipo de MCHO es un problema matemático, ya que se trata de una función implícita: el valor del potencial depende de las coordenadas de las cargas y las coordenadas de las cargas dependen del valor del potencial. Todos los ajustes se realizaron con una rutina de la biblioteca de Harwell²⁸ llamada VA05AD que busca el mínimo de una función de error (suma de los errores cuadráticos entre los valores predichos por el modelo y al valor al que se quiere ajustar) por gradiente y por una búsqueda de tipo simplex. Esto consiste en buscar el mínimo a partir de una proposición inicial de parámetros, evaluando el error cuadrático y modificando los valores en un procedimiento de aproximación sucesiva.

Intentando comprender por qué no funcionó ninguno de los potenciales propuestos analizamos el comportamiento de MCHO frente a deformaciones moleculares. Su respuesta a la deformación molecular se debe a la movilidad de las cargas en el modelo, que responden al campo eléctrico que depende indirectamente de las posiciones de los núcleos. Obviamente, el comportamiento del MCHO no corresponde al de la superficie vibracional ya que la forma analítica nunca fue ajustada para estas condiciones. Para no alterar la parte intermolecular, lo neutralizamos con un polinomio de orden cúbico en función de Δr .

Una vez eliminado el efecto del MCHO obtuvimos un ajuste al parecer muy bueno a la superficie vibracional con un potencial con términos cruzados y 14 parámetros ajustables sólo para la parte intramolecular. Este potencial tiene dos aspectos inconvenientes. El primero es que es una expresión muy complicada y no fácilmente extendible a problemas con un mayor número de enlaces. El segundo es que, como tiene términos cruzados, no se puede tabular (ver la sección 3 de este capítulo) y esto hace más cara su evaluación en la DM. Revisando la literatura encontramos que las formas funcionales que se han propuesto para el potencial vibracional del monómero de agua se pueden dividir en tres clases:

armónicas⁵⁶⁻⁵⁸, de Morse² y expansiones en términos de las coordenadas internas^{29,53-55}. La aproximación armónica es la menos satisfactoria.

El único reporte de un potencial inter e intramolecular *ab initio* es el de Lie y Clementi⁶, quienes hacen una simulación de agua líquida por DM con el potencial MCY²⁰ extendido con un potencial vibracional proveniente de una superficie calculada por Bartlett *et al.*²⁹, el cual se muestra en el cuadro 3.1.

CUADRO 3.1: POTENCIAL DE BARTLETT ET AL.⁶

$$U_{\text{intra}} = \frac{1}{2} f_{RR}(\delta_1^2 + \delta_2^2) + \frac{1}{2} f_{\theta\theta} \delta_3^2 + f_{RR'} \delta_1 \delta_2 + f_{R\theta}(\delta_1 + \delta_2) \delta_3 + \frac{1}{R_e} \left[f_{RRR}(\delta_1^3 + \delta_2^3) + f_{\theta\theta\theta} \delta_3^3 + f_{RRR'}(\delta_1 + \delta_2) \delta_1 \delta_2 + f_{RR\theta}(\delta_1^2 + \delta_2^2) \delta_3 + f_{RR'\theta} \delta_1 \delta_2 \delta_3 + f_{R\theta\theta}(\delta_1 + \delta_2) \delta_3^2 \right] + \frac{1}{R_e^2} \left[f_{RRRR}(\delta_1^4 + \delta_2^4) + f_{\theta\theta\theta\theta} \delta_3^4 + f_{RRRR'}(\delta_1^2 + \delta_2^2) \delta_1 \delta_2 + f_{RRR'R'} \delta_1^2 \delta_2^2 + f_{RRR\theta}(\delta_1^3 + \delta_2^3) \delta_3 + f_{RRR'\theta}(\delta_1 + \delta_2) \delta_1 \delta_2 \delta_3 + f_{RR\theta\theta}(\delta_1^2 + \delta_2^2) \delta_3^2 + f_{RR'\theta\theta} \delta_1 \delta_2 \delta_3^2 + f_{R\theta\theta\theta}(\delta_1 + \delta_2) \delta_3^3 \right]$$

donde δ_1 y $\delta_2 = R_i - R_e$ y $\delta_3 = R_e(\theta - \theta_e)$

$R_e = 1.8096$ ua vs 1.8088 ua (valor experimental)

$\theta_e = 104.59^\circ$ vs 104.52° (valor experimental)

constantes de fuerza en 10^5 dyn/cm:

f_{RR}	8.5120	f_{RRR}	-9.4970	f_{RRRR}	14.0000
$f_{\theta\theta}$	0.7987	$f_{\theta\theta\theta}$	-0.1268	$f_{\theta\theta\theta\theta}$	-0.0320
$f_{RR'}$	-0.0967	$f_{RRR'}$	-0.0310	$f_{RRRR'}$	-0.0500
$f_{R\theta}$	0.2732	$f_{RR\theta}$	-0.0340	$f_{RRR'\theta}$	0.0600
		$f_{RR'\theta}$	-0.5120	$f_{RRR\theta}$	-0.2000
		$f_{R\theta\theta}$	-0.1565	$f_{RRR'\theta}$	0.1000
				$f_{RR\theta\theta}$	-0.0800
				$f_{RR'\theta\theta}$	0.3500
				$f_{R\theta\theta\theta}$	0.1040

La superficie de Bartlett *et al.*²⁹ consta de 36 puntos: 22 deformaciones simétricas de

la geometría del monómero y 14 asimétricas. La deformación máxima corresponde a 2/3 de la amplitud de la vibración de punto cero (0.06 ua). Los cálculos se hicieron con una base 5S,4P,2D/3S,1P (39-STO) *all electron* con correlación electrónica incluyendo excitaciones dobles a orden infinito bajo el esquema MBPT. Estos cálculos son muy refinados y costosos, y permiten reproducir el espectro vibracional de la molécula de agua, el cual es muy sensible a la forma de la superficie vibracional. El potencial ajustado es muy complejo, como puede apreciarse en el cuadro 3.1, incluyendo acoplamientos de orden cuártico entre los osciladores.

Esta forma funcional en términos de coordenadas internas no es conveniente para utilizarse en simulaciones numéricas porque contiene términos cruzados, que corresponden al acoplamiento de los osciladores. Dada la calidad de la superficie sería muy conveniente lograr reproducirla con una forma funcional más sencilla para utilizarla en simulaciones numéricas. Para lograr lo anterior combinamos las superficies, usando la superficie de Bartlett *et al.*²⁹ para deformaciones dentro del intervalo de Δr entre -0.06 y + 0.06 ua, y para garantizar un comportamiento asintótico correcto, añadimos nueve puntos de la superficie calculada en este trabajo, con distorsiones mucho mayores de las que considera Bartlett *et al.*²⁹. Los puntos correspondientes a la superficie de Bartlett *et al.*²⁹ se obtuvieron evaluando el potencial analítico para 61 combinaciones de Δr 's, a intervalos de 0.03 ua, obteniéndose un ΔE máximo de 1.3149 kcal/mol. Un detalle importante es que en esta superficie E_0 no corresponde a la geometría experimental del monómero de agua en fase gaseosa:

$$r_{OH}^0 = 1.511096 \text{ ua vs. } 1.808880 \text{ ua (valor experimental)}$$

$$r_{HH}^0 = 2.863060 \text{ ua vs. } 2.860912 \text{ ua (valor experimental)}$$

Los nueve puntos con los que extendimos la superficie vibracional corresponden a deformaciones simétricas. Tratamos de incluir una asimétrica pero produjo complicaciones con el ajuste así que la descartamos. La justificación para no incluir esta geometría en el ajuste es que los puntos añadidos tienen la función de proporcionar el comportamiento asintótico correcto y deben ser muy poco frecuentes en las simulaciones numéricas de agua líquida. Los Δr y ΔE correspondientes pueden verse en la tabla 3.1.

Una vez determinada la superficie vibracional se procedió a ajustarla con un potencial de Morse y un polinomio de orden cúbico por tipo de enlace. El polinomio cumple la función de cancelar la energía intramolecular espúrea que aparece en el potencial intermolecular al variar la geometría del monómero. En esta fase no modificamos los parámetros de MCHO. Después incluimos el resto de la superficie (la parte intermolecular y momentos dipolar y cuadrupolar) y reajustamos la parte intramolecular de MCHO. Como paso final hicimos un reajuste global de todos los parámetros.

TABLA 3.1: GEOMETRÍAS AGREGADAS A LA SUPERFICIE DE BARTLETT²⁹

geometría	Δr_{OH_1}	Δr_{OH_2}	Δr_{HH}	ΔE
1	0.174	0.174	0.000	6.02024
2	0.000	0.000	-0.380	5.90678
3	0.080	0.080	0.376	4.39771
4	0.174	0.174	0.187	4.83777
5	-0.100	-0.100	-0.192	6.29219
6	-0.100	-0.100	-0.157	6.16268
7	-0.009	-0.009	0.000	5.81749
8	0.000	0.000	0.565	18.42200
9	0.268	0.268	0.000	15.60450

distancias en ua; energías en kcal/mol; $\Delta r = r - r_{eq}$ y $\Delta E = E - E_0$, donde E_0 y r_{eq} son los valores de energía y distancia para la geometría experimental.

En el cuadro 3.2 se muestra la forma funcional del nuevo potencial, llamado MCHOv; en la tabla 3.2 se muestran los parámetros de MCHO con los que se empezó el ajuste y en la tabla 3.3 pueden verse los parámetros finales para MCHOv. Los parámetros tienen las unidades adecuadas para el cálculo de energía potencial en kcal/mol; las distancias deben estar en unidades atómicas (bohrs) y las cargas en unidades atómicas de carga. Los parámetros C deben multiplicarse por 627.5 para obtener kcal/mol. Las diferencias fundamentales entre MCHO y MCHOv son la adición de un potencial nuclear intramolecular, la adición de términos k/r^n para interacciones electrostáticas entre átomos de cargas de signo igual y la adición de la interacción electrostática intramolecular entre H.

CUADRO 3.2: FORMA FUNCIONAL DE MCHOv

parte intermolecular:

a) dependiente de las coordenadas nucleares:

$$A_{ij}e^{-n_{1ij}R_{i,j}} + B_{ij}e^{-n_{2ij}R_{i,j}}$$

b) dependiente de las coordenadas de las cargas:

$$\frac{C_{ij}\delta_i\delta_j}{r_{i,j}} + \frac{k_{ij}}{r_{i,j}^{n3_{ij}}}$$

parte intramolecular:

a) dependiente de las coordenadas nucleares:

$$D_{ij}(e^{-2\alpha_{ij}\Delta r} - 2e^{-\alpha_{ij}\Delta r} + 1) + \text{pol}1_{ij}\Delta r + \text{pol}2_{ij}(\Delta r)^2 + \text{pol}3_{ij}(\Delta r)^3$$

b) dependiente de las coordenadas de las cargas:

$$\frac{C_{ij}\delta_i\delta_j}{r_{i,j}} + \frac{k_{ij}}{r_{i,j}^{n3_{ij}}}$$

c) enlace entre cargas y núcleos de un átomo:

$$a_i r_{j,i}^2 + a_j r_{j,j}^2$$

R son distancias entre los núcleos de moléculas distintas; r son distancias entre cargas o entre núcleos y cargas. Los núcleos se representaron con mayúsculas y las cargas con minúsculas. Δr es la diferencia entre la longitud del enlace y la longitud del enlace con energía mínima.

TABLA 3.2: PARÁMETROS DE MCHO

parámetro	O-O	O-H	H-H
A	3.471402×10^8	691.958520	15.287312
B	-1.748782×10^9	-34.475771	1.183873×10^{23}
n_1	3.575798	1.489193	0.633568
n_2	3.977968	0.700677	15.202894
C_{inter}	1.000000	1.000000	1.000000
k_{inter}		7118.564100	
$n3_{inter}$		6.490107	
C_{intra}		0.467624	
k_{intra}		99.545247	
$n3_{intra}$		2.418201	

O: $a = 59.012287$, $\delta = -1.426165$

H: $a = 455.05927$, $\delta = 0.713087$

TABLA 3.3: PARÁMETROS DE MCHOV

parámetro	O-O	O-H	H-H
A	1.000000×10^8	999.942700	-1422.459000
B	-1.000000×10^8	-749.623300	2.188749×10^{15}
n_1	2.934434	1.040209	2.317225
n_2	2.941385	0.969374	10.610300
C_{inter}	1.000000	1.000000	1.000000
k_{inter}	2.558486	2615.392000	-0.012517
n_{3inter}	5.645374	5.150225	5.763937
C_{intra}		25.374660	1.064440
k_{intra}		13962.380000	0.087112
n_{3intra}		1.011549	1.026968
D		10.362350	8.369590
α		-4.400811	-0.010674
pol1		-37.566050	17.804190
pol2		-31.251470	30.889570
pol3		233.076500	-0.331437
r_{eq}		1.811096	2.863060

O: $a = 60.095280$, $\delta = -1.334915$

H: $a = 172.50120$, $\delta = 0.667457$

En la tabla 3.4 puede verse la bondad del ajuste. Los valores corresponden a una recta ajustada a la comparación de los valores *ab initio* y los valores predichos por el modelo. Mientras mejor es el ajuste hay menos dispersión en los puntos y se puede ajustar mejor una recta a esta gráfica. En la tabla 3.5 se presenta una comparación del ajuste logrado en este trabajo con el ajuste de MCHO¹ a las mismas superficies de 2, 3 y 4 cuerpos. La comparación muestra que la extensión vibracional al potencial intermolecular es muy satisfactoria, produciendo ligeras mejoras en el ajuste a las no aditividades de tres y cuatro cuerpos y una ligera pérdida en el ajuste de la superficie de dímeros. El ajuste global presenta una mejora a pesar de la mayor complejidad del potencial y el mayor número de puntos, un resultado muy alentador.

TABLA 3.4: PARÁMETROS DE LA BONDAD DEL AJUSTE DE MCHO_v A LAS DISTINTAS SUPERFICIES

	<i>m</i>	<i>b</i>	ρ
MCHO_v			
superficie vibracional I + II	0.864	0.182	0.918
superficie vibracional I	0.932	0.040	0.949
superficie vibracional II	0.608	3.034	0.750
superficie de dímeros	1.008	0.045	0.979
n.a. trímeros	0.930	0.018	0.840
n.a. tetrameros	0.367	-0.053	0.399
MCHO			
superficie de dímeros	1.008	0.060	0.982
n.a. trímeros	0.911	0.034	0.763
n.a. tetrameros	0.359	-0.040	0.368

I: superficie vibracional de Bartlett *et al.*²⁹; II: extensión vibracional calculada en este trabajo; n.a.: no aditividad; *m*: pendiente de la recta ajustada; *b*: ordenada al origen de la recta ajustada; ρ : coeficiente de correlación

Los datos corresponden al ajuste a 66 dímeros, 40 trímeros y 15 tetrameros en la parte intermolecular; la superficie vibracional consta de 70 puntos.

Las figuras 3.1, 3.2, 3.3 y 3.4 muestran las gráficas de los ajustes a las superficies de monómeros, dímeros, no aditividad de 3 cuerpos y no aditividad de 4 cuerpos, respectivamente. Del ajuste de los monómeros, puede verse que el ajuste a la parte de Bartlett *et al.*²⁹ es muy bueno, y que la parte correspondiente a geometrías más deformadas no está tan bien ajustada. Esto no nos preocupó porque en principio en la simulación no debiéramos tener deformaciones moleculares tan grandes en promedio (cuando mucho 0.06 ua ²⁰) y sólo nos interesó que el comportamiento del potencial fuera cualitativamente correcto en esta zona de $|\Delta r|$ grandes.

TABLA 3.5: COMPARACIÓN DE LA BONDAD DE LOS AJUSTES DE LOS POTENCIALES MCHO Y MCHOv

	MCHO	MCHOv
ϵ^2	15.942892	18.465800
σ monómeros I + II		0.912262
σ monómeros I		0.074543
σ dímeros	0.151917	0.178054
σ trímeros	0.212838	0.168873
σ tetrámeros	0.178054	0.173491
σ global	0.350197	0.303857
# de puntos	130	200

La bondad del ajuste está definida por los siguientes parámetros:

$$\sigma = \sqrt{\sum [(y(i) - yc(i))^2] / \#puntos}$$

$$\epsilon^2 = \sum [(y(i) - yc(i))^2]$$

$y(i)$ es el valor de la energía (de deformación, de interacción o de la no aditividad) ab initio correspondiente al agregado i , y $yc(i)$ es el valor predicho por el modelo para la misma cantidad

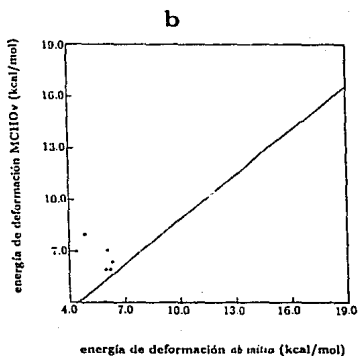
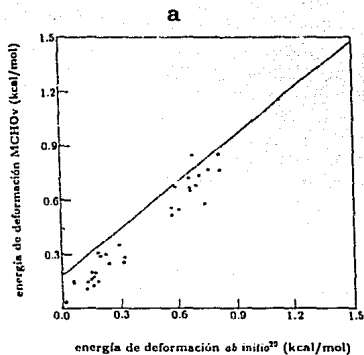


Figura 3.1: energía de deformación de monómeros de agua *ab initio* y calculada con el modelo; a: ajuste a la superficie de Bartlett et al.²⁹; b: ajuste a los puntos de la tabla 3.1; la recta corresponde al ajuste de ambas superficies

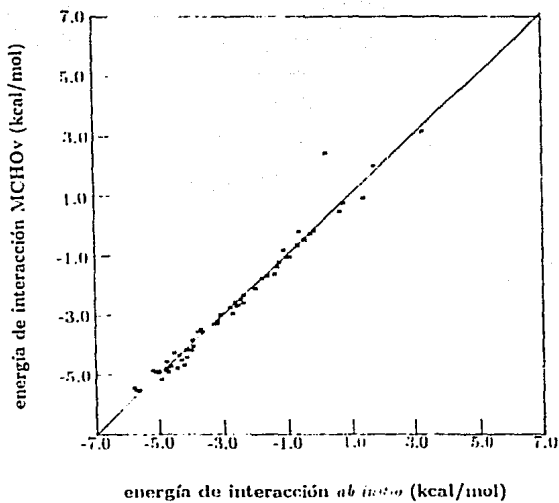


Figura 3.2: energía de interacción de dímeros de agua *ab initio* y calculada con el modelo

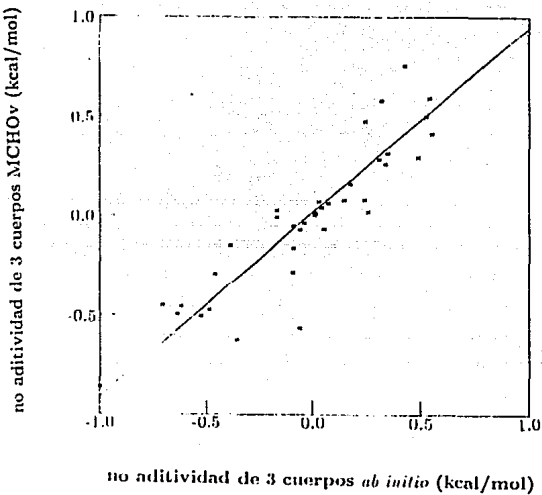


Figura 3.3: no aditividad de tres cuerpos *ab initio* y calculada con el modelo

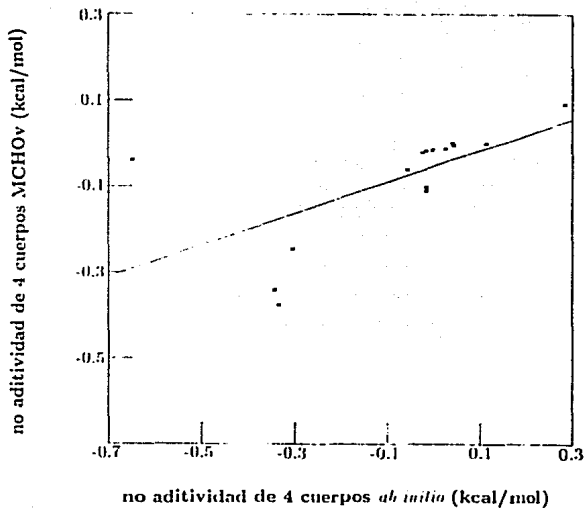


Figura 3.4: no aditividad de cuatro cuerpos *ab initio* y calculada con el modelo

Un punto importante de resaltar es que al ajustar la superficie vibracional con osciladores de Morse el modelo se puede extender hacia una molécula que además de vibrar se disocie, cosa que un potencial parabólico no permite.

El tener la superficie de monómeros con una geometría de energía mínima distinta a la geometría del resto de la superficie no genera problemas, ya que la diferencia (Δr) es muy pequeña y ΔE es del orden de 10^{-6} kcal/mol.

Hay que hacer notar que pese a la apariencia compleja del potencial, la energía entre dos átomos es función exclusiva de la distancia entre ellos; esto tiene muchas ventajas en las simulaciones numéricas. Además, no es una forma funcional *ad hoc* para la molécula de agua, y puede extenderse a otras moléculas. En el laboratorio se está trabajando actualmente con su extensión a la molécula de formamida, con resultados satisfactorios.

3.3 DESARROLLO DEL PROGRAMA DE DINÁMICA MOLECULAR

Como discutimos en la sección 2.4.3, la dinámica molecular permite el cálculo de las propiedades dinámicas y de transporte del sistema simulado, mientras que el Monte Carlo no puede hacer esto. Una de las razones para desarrollar el programa de dinámica molecular es nuestro interés en las propiedades dinámicas del agua, que, como mencionamos en la introducción, están relacionadas con la función de las biomoléculas.

Una pregunta obligada es ¿por qué escribir un programa de este tipo si hay programas comerciales muy buenos? La primera razón es que el modelo molecular que desarrollamos tiene una colección de particularidades, empezando por las cargas móviles sin masa, que los paquetes comerciales no saben manejar. Otra razón es que en el laboratorio no había prácticamente ninguna experiencia con este tipo de simulaciones, y el escribir el programa nos permitió entender su funcionamiento.

Aprovechando que la computadora en la que desarrollamos el programa tiene un compilador que vectoriza, el programa se escribió tratando de que fuera sencilla su vectorización. Además tiene una estructura más o menos modular, de manera que sea sencillo incluir subrutinas con funciones nuevas.

Dado que las computadoras actuales tienen más capacidad de memoria que las versiones anteriores, una de las características que debiera tener el programa es que todos los potenciales y fuerzas estuvieran tabuladas, salvo la parte de la interacción carga-núcleo, que no tiene sentido tabular por ser una simple multiplicación. Tabular un potencial quiere decir evaluarlo para una serie de distancias a intervalos fijos y guardarlo en un arreglo. Cuando se quiere saber el valor del potencial para una distancia en particular basta buscar ese valor en la casilla correspondiente del arreglo.

Los potenciales de interacción y las fuerzas se guardaron en sendas matrices de di-

mensionen $12 \times 30,000$. Doce porque para el agua líquida ese es el número de potenciales distintos por pareja de átomos y 30,000 porque encontramos que era un compromiso razonable entre el espacio ocupado en memoria (cerca de 5 Mbytes) y la resolución tanto en distancias como en energías. Para determinar qué tipo de potencial corresponde a cada interacción se cuenta con unas matrices pequeñas, dependientes de los tipos de átomos que se simulen, donde se indica el número de potencial o fuerza que corresponde al tipo de átomos en interacción.

Los potenciales se evaluaron directamente las expresiones mostradas en el cuadro 3.2. Para tabular las fuerzas se obtuvo la derivada analítica de cada potencial y se multiplicó por $-|\vec{r}|^{-1}$, de manera que al hacer la suma vectorial de fuerzas sólo hubiera que multiplicarla por los segmentos dirigidos en x, y y z correspondientes a la pareja de átomos en cuestión.

Una de las ventajas de tabular las fuerzas y los potenciales es que no importa cuan compleja sea la forma funcional, se evalúa una sola vez. Esto elimina mucho tiempo de cómputo, ya que el cálculo de exponenciales y potencias es muy caro. El único requisito para hacer tablas como la que aquí se usaron es que la función dependa de una sola variable. El precio que se paga es que se convierte una función continua en una discreta, y la resolución tanto en distancias como en energías es ligeramente menor.

El siguiente problema a resolver fueron las unidades de masa, carga, distancia, tiempo, fuerza y energía que se utilizaron en el programa. Algunos autores (por ejemplo Lie y Clementi⁶) utilizan el sistema CGS, pero esto implica usar una colección de números muy pequeños que pueden dar problemas con la precisión de la computadora. Después de ensayar varios juegos de unidades decidimos utilizar las siguientes:

masa : unidades atómicas de masa (uma)
carga : unidades atómicas de carga (e)
distancia : unidades atómicas o bohrs (ua)
tiempo : segundos (s)
fuerza : $\text{uma} \cdot \text{ua}^2/\text{s}^2$
energía : $\text{uma} \cdot \text{ua}^3/\text{s}^2$

El programa reporta energías en kcal/mol para que sean fácilmente comparables con datos experimentales y de otras simulaciones.

Como punto de partida, la DM requiere de las posiciones de las cargas y de los núcleos y de un juego de velocidades iniciales. Las coordenadas se tomaron de una configuración de 343 moléculas equilibradas por MC a 298K con MCHO. Las velocidades deben ser asignadas de acuerdo a una distribución de Boltzmann y debe cumplirse la equipartición de la energía cinética para cada especie atómica o molecular. En el agua, con nuestro modelo, tenemos dos especies atómicas, una con masa 1 y la otra con masa 16. Para lograr la equipartición de energía cinética dividimos la velocidad asignada a los O entre la raíz cuadrada de su masa.

El cálculo de la energía potencial de una configuración no tiene la menor complicación. Simplemente hay que hacer una suma sobre todas las interacciones por pareja de átomos en el sistema. Sin embargo, el calcular la suma de fuerzas sobre un átomo no es tan trivial por lo siguiente: en el modelo MCHO, y su variante MCHOv, cada átomo está representado por un núcleo que no tiene carga pero tiene masa y una carga móvil que no tiene masa, ambos unidos por un potencial armónico. Una solución ingenua es calcular las fuerzas sobre el núcleo y la carga por separado y mover a cada uno según la fuerza resultante. Como las cargas no tienen masa, no se pueden acelerar: $a = F/m$ no tiene solución si $m = 0$. Ni siquiera se les pueden asignar velocidades iniciales en principio porque lo que se asignan son momentos lineales $p = mv$ y si $m = 0$, $p = 0$. Además, falta tomar en cuenta al potencial armónico que une a la carga y el núcleo.

La solución que propusimos es calcular las fuerzas sobre la carga y el núcleo, sumarlas y aplicar la fuerza total al núcleo. El desplazamiento que resulte de esta fuerza, se aplica por igual al núcleo y a la carga, como si estuvieran unidos por una barra rígida. Una vez que se han movido todos los átomos, las nuevas posiciones de las cargas se encuentran con un proceso autoconsistente¹. Básicamente, este proceso consiste en encontrar las posiciones de equilibrio de las cargas de una molécula en el campo eléctrico generado por las cargas de las moléculas vecinas y las propias. El equilibrio se alcanza cuando las posiciones predichas en el ciclo n no cambian en más de un criterio de convergencia externo con las del ciclo $n - 1$. En este caso, el criterio de convergencia es de 1×10^{-3} ua por componente (x, y, z) . La justificación de este modelo es clara: estamos asignando toda la masa del átomo al núcleo y desplazando la nube electrónica a la par, recordando la aproximación de Born-Oppenheimer.

Hemos hablado de mover los átomos. ¿Cómo se hace esto? Tampoco resultó ser trivial. El problema consiste en integrar un sistema de ecuaciones diferenciales de segundo orden:

$$\sum F = m \frac{d^2 x}{dt^2}$$

$$F_{ij} = -\nabla V_{ij}$$

donde m es la masa y x es la posición de una partícula; t es el tiempo; F es la fuerza y V es el potencial que actúan sobre la partícula.

Las características que debe poseer un buen algoritmo de integración son la estabilidad y que las fuerzas se calculen una sola vez por paso (es la parte costosa de la simulación).

La solución más sencilla es la que se presenta en cualquier libro de cinemática:

$$x = x_0 + v_0 t + \frac{at^2}{2}$$

$$v = v_0 + at$$

$$a = \frac{\sum F}{m}$$

donde m es la masa, r es la posición, v es la velocidad, y a es la aceleración de una partícula; t es el intervalo temporal y F es la fuerza que actúa sobre la partícula.

Resulta que esta solución tiene un error del orden de t^3 y es muy inestable para t grandes (para nuestro sistema esto resultó ser del orden de 10^{-16} s). Los algoritmos que tienen un orden menor que tercer orden (éste es de segundo) no son eficientes en DM de sistemas moleculares³¹.

Otro algoritmo que probamos es el de *leap-frog* o de diferencia central³¹:

$$v(t + \Delta t/2) = v(t - \Delta t/2) + \frac{\sum F(t)}{m} \Delta t$$

$$r(t + \Delta t) = r(t) + v(t + \Delta t/2)\Delta t$$

donde m es la masa, r es la posición y v es la velocidad de una partícula; Δt es el intervalo temporal y F es la fuerza que actúa sobre la partícula. Tiene la propiedad de ser invariante ante la inversión temporal y esto debiera evitar una deriva de la energía total del sistema. Este algoritmo es de tercer orden en el paso temporal de integración y es menos preciso que algoritmos de mayor orden. Sin embargo, debiera ser más estable y eficiente para valores grandes de Δt (10^{-13} s). Este algoritmo tampoco fue satisfactorio ya que había que estar enfriando el sistema con mucha frecuencia, lo cual habla de una muy pobre conservación de la energía.

El siguiente algoritmo que probamos fue el de Verlet³², que se obtiene de escribir dos expansiones de Taylor para las coordenadas de las partículas, una hacia adelante y otra hacia atrás en el tiempo:

$$q(t + \delta t) = q(t) + (dq/dt)\delta t + (d^2q/dt^2)\delta t^2/2! + \dots$$

$$q(t - \delta t) = q(t) - (dq/dt)\delta t + (d^2q/dt^2)\delta t^2/2! + \dots$$

y sumarlas:

$$q(t + \delta t) = 2q(t) - q(t - \delta t) + (d^2q/dt^2)\delta t^2 + O(\delta t^4)$$

(q son las coordenadas de la partícula y δt es el intervalo de tiempo).

Este es el algoritmo independiente de velocidades. Estas se calculan como:

$$v(t) = \frac{q(t + \delta t) - q(t - \delta t)}{2\delta t}$$

Este método es escasamente inferior y a veces hasta superior a los esquemas de predictor-corrector, y además requiere guardar menos información³³. Este método sí fue estable con el Δt que se recomienda⁶ para estudiar sistemas con vibraciones: 10^{-16} s.

El siguiente paso fue decidir qué información guardar de la simulación. Cada paso se calculan las energías potencial, cinética y total, la temperatura, y los promedios de estas cantidades. Además, guarda las coordenadas y velocidades de la configuración con energía potencial más baja que se haya encontrado durante la simulación en archivos que se reescriben. Cada n pasos (con n decidido por el usuario) se escriben las energías, la temperatura, sus promedios y se calculan y escriben en archivos separados:

- coordenadas de todas las partículas
- velocidades de los núcleos
- autocorrelación de momentos dipolares
- autocorrelación de velocidades por especie atómica
- desplazamientos cuadráticos medios por especie atómica
- longitudes de enlace por tipo de enlace
- histograma de velocidades por especie atómica
- número de vecinos a distintos radios para la función de correlación radial.

Por supuesto, mientras más frecuente sea el monitoreo de estas propiedades más lenta será la simulación.

El programa además puede congelar especies atómicas, de manera que uno puede por ejemplo relajar los H de un grupo de moléculas de agua sin mover los O.

El listado (en Fortran 77) del programa se presenta en el Apéndice C.

3.4 SIMULACIÓN POR DINÁMICA MOLECULAR

condiciones de la simulación

Se simularon 343 moléculas de agua a una densidad $\rho = 0.997$ g/cm³; la simulación se hizo a ≈ 300 K con condiciones toroidales a la frontera y corte de energía de interacción y fuerzas esférico a $L/2$ (20.55 ua), donde L es la longitud de la celda computacional. El paso de integración fue de 2×10^{-16} s.

Aunque el programa de DM permitió obtener resultados iniciales, no fue probado por completo, y aún subsisten algunos problemas que deben ser corregidos. Por razones aún desconocidas no pudimos hacer simulaciones largas (de más de 350 pasos o 7×10^{-14} s). En total, se simuló cerca de 1 ps de tiempo real.

La DM tardó mucho en equilibrarse porque aunque partimos de una configuración espacial que debiera estar cercana al equilibrio ya que los potenciales MCHO y MCHOv son muy semejantes, la parte cinética ciertamente estaba muy alejada del equilibrio. De hecho, las distribuciones de velocidades escogidas al azar nos dieron problemas de convergencia en las posiciones de las cargas. Al principio de la simulación tuvimos que usar un paso de integración de 2×10^{-17} s y subirlo posteriormente al paso recomendado por Lie y Clementi⁶ de 2×10^{-16} s.

Ya que en el proceso de equilibrar el sistema en la DM la energía potencial no se estabilizó, y además daba valores más negativos que el valor experimental, decidimos caracterizar el potencial con un método más barato en términos de tiempo de cómputo. Para esto modificamos un programa de MC, desarrollado en el laboratorio, para que maneja vibraciones. Partiendo de la misma configuración equilibrada por MC con MCHO, hicimos una simulación por MC de 1×10^6 pasos con MCHOv. La energía potencial del sistema convergió en ≈ -13.6 kcal/mol, valor que debe compararse con -9.92 kcal/mol, correspondiente al potencial medido experimentalmente³⁴. Una simulación por MC de 5×10^7 pasos bajo condiciones idénticas con MCHOv pero con moléculas rígidas da una energía potencial de -11.14 kcal/mol (H.Saint Martin, comunicación personal). La diferencia entre los resultados obtenidos con el modelo de molécula rígida y la flexible es muy grande, sobre todo si se considera la adición de grados de libertad vibracionales como una modificación fina al potencial.

El cambio tan drástico en la energía potencial al incluir vibraciones no es un problema exclusivo de nuestro potencial. En la tabla 3.6 se encuentra una lista de los potenciales intermoleculares extendidos para incluir vibraciones y de los potenciales que no distinguen entre interacciones intra e intermoleculares. En la tabla 3.7 se muestran los resultados de algunas simulaciones de agua líquida hechas con potenciales flexibles y con su versión de molécula rígida. El único potencial *ab initio* es el MCY-L⁶, y puede verse que hay un descenso en la energía potencial de casi 1 kcal/mol al incluir vibraciones. El valor de la energía potencial ya incluye correcciones cuánticas, tanto para la parte rígida como para la parte flexible, y esto pudiera enmascarar un cambio en energía potencial mayor que el reportado. El que los demás potenciales den valores de energía potencial cercanos al experimental se debe a que fueron ajustados para que dieran ese valor. Fuera del potencial MCY-L, que es un potencial de dos cuerpos, todos los demás son potenciales efectivos. Hay un reporte de una simulación con el potencial NCC³⁵ con una extensión que considera moléculas flexibles. Este es un potencial *ab initio* que incluye efectos no aditivos. La energía potencial pasa de -10.65 kcal/mol para molécula rígida a -10.68 kcal/mol para molécula flexible³⁶, correcciones cuánticas incluidas.

TABLA 3.6: POTENCIALES EXTENDIDOS Y POTENCIALES CENTRALES**POTENCIALES INTERMOLECULARES:**

potencial	ajuste	ref
RWK2	empírico	37
ST2	empírico	38
SPC	empírico	39
PSPC	empírico	40
TIPS	empírico	41
CF	empírico	16-18
MCY	ab initio	26,42,43

POTENCIALES CENTRALES:

potencial	ajuste	ref
CF	empírico	16-18
PM	empírico	19-21

TABLA 3.7: SIMULACIONES CON POTENCIALES FLEXIBLES

ref	potencial	# moléculas	DM/MC	ΔE rígido	ΔE flexible	$\overline{r_{OH}}$	$\overline{\theta}$
44	SPC-M	125	DM	-7.84	-10.18	1.9559	106.63
11	SPC-H	216	DM	-9.08	-9.56 a -10.51	1.9200	104.90
12	SPC-V	216	DM	-10.08	-10.13		
13	PSPC-ar	125	DM	-9.08	-9.94	1.9219	104.00
13	PSPC-anar	125	DM	-9.08	-10.21	1.9219	105.00
13	PSPC-anar,q *	125	DM	-9.08	-7.98	1.9030	104.00
13	PSPC-anar,q, α *	125	DM	-9.08	-8.89	1.9124	105.00
4	RWK2-M	64	MC	-10.02	-10.06	1.8500	104.63
12	TIPS-V	216	DM	-8.9	-9.10	1.8368	
6	MCY-L	343	DM	-8.53	-9.42	1.8425	103.50
34,45	experimental				-9.92	1.8255	102.81

distancias en ua; ángulos en grados; energías en kcal/mol; $\overline{r_{OH}}$ = distancia intramolecular OH promedio; $\overline{\theta}$ = ángulo intramolecular promedio; * q = cargas variables; α = polarización variable; ar = potencial armónico; anar = potencial anarmónico

El descenso en energía potencial tan marcado que observamos con MCHO_v se debe muy probablemente a que las polarizabilidades predichas por el modelo son más grandes que las experimentales. Esto tiene la consecuencia de permitir una relocalización de carga que hace que la interacción electrostática sea más favorable. Una manera de corregir este problema es modificar el modelo agregando un término que trate explícitamente a la polarizabilidad (como hacen algunos modelos que incluyen efectos no aditivos^{21,41,58}) y ajustar este nuevo modelo a superficies de energía de interacción *ab initio* de agregados de moléculas deformadas; otra opción es ver si el modelo como está es capaz de ser ajustado a estas superficies. Ante un caso similar, los paquetes comerciales de simulación numérica de sistemas moleculares ofrecen la opción de reescalar algunos parámetros de la interacción molecular, como los valores de las cargas, hasta que el usuario esté satisfecho con el resultado obtenido. Creemos que esta manera de solucionar el problema no es adecuada porque el método pierde generalidad y poder explicativo.

De todas las propiedades que se pueden estudiar con MC y DM, la más "resistente" a errores es la energía del sistema. Ejemplos muy claros de propiedades sensibles son la presión y detalles finos de la función de correlación radial. En vista de que encontramos que la energía estaba mal en un 30%, no tiene caso analizar otros parámetros del sistema sin corregir el potencial. La corrección del potencial se hizo dentro del esquema de simulación global, regresando al paso en el que se construye el potencial analítico y se ajusta a los datos *ab initio*. Como el potencial está ajustado a superficies de hasta cuatro cuerpos, estudiamos el comportamiento del potencial en agregados de cuatro moléculas de agua para ver cómo maneja la energía de deformación, la energía de interacción entre dímeros y la no aditividad, y compararlos con los valores *ab initio*.

3.5 NO ADITIVIDAD DE MCHO_v

Estudiamos tres tetrámeros formados con monómeros en la geometría experimental (tabla 3.8). Los valores de energías de interacción y de no aditividad son muy razonables y muy parecidos al comportamiento del MCHO, cosa que no es de extrañarse dada la similitud del ajuste.

El tetrámero 1 lo minimizamos por MC con MCHO_v. Al tetrámero 2 le fijamos los O e hicimos una DM a 20K con un Δt de 1×10^{-18} s por 10^4 pasos. Las coordenadas de los tetrámeros relajados están reportadas en el Apéndice B.

Hicimos el análisis energético detallado del tetrámero 1 con las geometrías de los monómeros predichas por el potencial. Los resultados se muestran en la tabla 3.9. Comparando las tablas 3.8 y 3.9 puede verse que, para el tetrámero 1, la deformación de las moléculas se ve favorecida en los trímeros y el tetrámero (≈ 6 kcal/mol de estabilización respecto al mismo tetrámero pero sin deformar), mientras que para los dímeros hay una desestabilización de entre 1.5 y 2.5 kcal/mol. Obviamente, en los monómeros no se favorece

la deformación.

Las no aditividades fueron calculadas con los dos esquemas propuestos en la sección 2.3, ya sea con la expresión canónica que no incluye de forma explícita la energía de deformación molecular o con la extensión que sí la incluye de manera explícita.

TABLA 3.8: ENERGÍAS DE INTERACCIÓN Y NO ADITIVIDADES TETRÁMEROS DE MONÓMEROS NO DEFORMADOS

	tetrámero 1	tetrámero 2	tetrámero 3
E_{12}	-1.6899	-3.5947	-4.2260
E_{13}	-5.4510	-1.9037	-1.9037
E_{14}	-5.3666	-3.5464	-4.2012
E_{23}	-5.3666	-3.5947	-4.2260
E_{24}	-5.4510	-1.9037	-1.9037
E_{34}	-1.6899	-3.5947	-4.2260
$\sum 2c$	-25.0150	-18.1379	-20.6866
E_{123}	-13.3176	-9.7886	-11.0301
E_{124}	-13.3419	-9.6541	-10.9114
E_{134}	-13.2125	-9.7649	-11.0032
E_{234}	-13.2394	-9.7886	-11.0301
E_{1234}	-28.4692	-21.2210	-23.8518
na_{123}^1	-0.8101	-0.6955	-0.6744
na_{124}^1	-0.8344	-0.6093	-0.5805
na_{134}^1	-0.7050	-0.7201	-0.6723
na_{234}^1	-0.7319	-0.6955	-0.6744
$\sum na^1$	-3.0814	-2.7204	-2.6016
na^4	-0.3728	-0.3627	-0.5636

energías y no aditividades en kcal/mol; E = energía de interacción del agregado; na_m^n = no aditividad de n cuerpos del agregado m

**TABLA 3.9: ENERGÍAS DE INTERACCIÓN Y NO ADITIVIDADES
TETRÁMERO 1 CONSTRUIDO CON MONÓMEROS DEFORMADOS
DESARROLLO CANÓNICO DE MUCHOS CUERPOS**

agregado	E interacción	no aditividad
mon 1	1.9921	
mon 2	1.9515	
mon 3	1.9257	
mon 4	1.9592	
Σ mon	7.8285	
dim 12	1.0402	
dim 13	-3.6577	
dim 14	-3.6533	
dim 23	-3.7425	
dim 24	-3.6318	
dim 34	0.9584	
Σ dim	-12.6867	
tri 123	-13.7489	-7.3889
tri 124	-13.7147	-7.4698
tri 134	-13.7943	-7.4417
tri 234	-13.7917	-7.3758
Σ na ³		-29.6762
tet	-34.8995	7.4634

DESARROLLO DE MUCHOS CUERPOS CON δ EXPLÍCITA

agregado	E interacción	no aditividad
mon 1	1.9921	
mon 2	1.9515	
mon 3	1.9257	
mon 4	1.9592	
Σ mon	7.8285	
dim 12	-2.9031	
dim 13	-7.5755	
dim 14	-7.6066	
dim 23	-7.6194	

dim 24	-7.5422	
dim 34	-2.9265	
Σ dim	-36.1733	
tri 123	-13.7489	-1.5199
tri 124	-13.7147	-1.5653
tri 134	-13.7943	-1.5627
tri 234	-13.7917	-1.5397
Σ na ³		-6.1876
tet	-34.8995	-0.3668

energías y no aditvidades en kcal/mol; mon = monómero; dim = dímero; tri = trímero; tet = tetrámero

La tabla 3.9 muestra los comportamientos dispares de las dos expansiones de muchos cuerpos que definimos en la sección 2.3. La expansión que no considera explícitamente la energía de deformación presenta un comportamiento errático, ciertamente no convergente. Además, la no aditividad de tres cuerpos es el 85% de la energía de interacción del tetrámero, valor patológico según la discusión de la sección 2.3. Este valor tan alto podría deberse al potencial en sí. Por otro lado, la expansión que sí toma en cuenta explícitamente la energía de deformación muestra un comportamiento convergente, con valores razonables.

Dentro del esquema de simulación global a partir de primeros principios podemos verificar las predicciones hechas por el modelo y rectificar el potencial en caso de ser necesario. Para este efecto, calculamos *ab initio* los tetrámeros 1 y 2 de moléculas relajadas. Con la base DZP calculamos los tetrámeros completos y los trímeros, dímeros y monómeros correspondientes en la base del tetrámero (*full CP*). La no aditividad calculada con los dos esquemas mencionados en la sección 2.3 se muestra en la tabla 3.10.

Como las no aditvidades pueden ser fuertemente dependientes de la base con la que se calculan, buscamos una base más extendida para repetir los cálculos. Tomamos como punto de partida una base de Clementi^{26,49} y probamos diversas contracciones de las funciones Gaussianas primitivas con el dímero de agua como calibrador⁵⁰. La base que dio mejores resultados es la QZP (ver Apéndice A).

Con esta base recalculamos los tetrámeros 1 y 2 relajados, también con *full CP*. Los resultados se presentan en la tabla 3.10, y en la tabla 3.11 se muestra la convergencia de los dos desarrollos de muchos cuerpos que pueden escribirse para tratar moléculas deformadas.

TABLA 3.10: NO ADITIVIDADES PARA LOS TETRÁMEROS 1 Y 2
CONSTRUIDOS CON MONÓMEROS DEFORMADOS
CÁLCULOS *AB INITIO*

		canónica	con δ
BASE DZP			
TETRÁMERO 1	na_{123}^3	-14.8441	-2.6870
	na_{124}^3	-14.9759	-2.7235
	na_{134}^3	-14.9025	-2.7133
	na_{234}^3	-14.8008	-2.7121
	$\sum na^3$	-59.5233	-10.8359
	na^4	15.2991	-0.9281
TETRÁMERO 2	na_{123}^3	-25.8028	-2.2515
	na_{124}^3	-25.4897	-2.2540
	na_{134}^3	-26.8294	-2.3399
	na_{234}^3	-25.8768	-2.3054
	$\sum na^3$	-103.9987	-9.1508
	na^4	30.8435	-0.7725
BASE QZP			
TETRÁMERO 1	na_{123}^3	-14.7061	-2.5238
	na_{124}^3	-14.8291	-2.5671
	na_{134}^3	-14.7582	-2.5495
	na_{234}^3	-14.6634	-2.5483
	$\sum na^3$	-58.9568	-10.1887
	na^4	15.3461	-0.9099
TETRÁMERO 2	na_{123}^3	-25.3642	-2.1103
	na_{124}^3	-25.0366	-2.1122
	na_{134}^3	-26.3625	-2.1937
	na_{234}^3	-25.3686	-2.1537
	$\sum na^3$	-102.1319	-8.5754
	na^4	30.4413	-0.7442

no aditvidades en kcal/mol; na_m^n = no aditvidad de n cuerpos del agregado m

TABLA 3.11: COMPARACIÓN DE LOS DESARROLLOS DE MUCHOS CUERPOS

DESARROLLO CANÓNICO

agregado	E_n^{int}	$\sum 2c$	$\sum na^3$	na^4
TETRÁMERO 1 BASE DZP				
1234	-10.3851	33.8392	-59.5233	15.2991
123	1.9741	16.8183	-14.8441	
124	2.1806	17.1565	-14.9759	
134	2.0846	16.9871	-14.9025	
234	1.9158	16.7166	-14.8008	
TETRÁMERO 1 BASE QZP				
1234	-6.0422	37.5684	-58.9568	15.3461
123	3.9702	18.6763	-14.7061	
124	4.1892	19.0183	-14.8291	
134	4.0963	18.8545	-14.7582	
234	3.9244	18.5878	-14.6634	
TETRÁMERO 2 BASE DZP				
1234	7.0518	80.2071	-103.9987	30.8435
123	14.0554	39.8582	-25.8028	
124	13.6701	39.1598	-25.4897	
134	14.8404	41.6698	-26.8294	
234	13.8496	39.7264	-25.8768	
TETRÁMERO 2 BASE QZP				
1234	10.2678	81.9584	-102.1319	30.4413
123	15.4534	40.8176	-25.3642	
124	14.9816	40.0182	-25.0366	
134	16.1644	42.5269	-26.3625	
234	15.1855	40.5541	-25.3686	

DESARROLLO CON δ EXPLÍCITA

agregado	E_m^{int}	$\sum 2c$	$\sum na^2$	na^4	$\sum \delta$
TETRÁMERO 1 BASE DZP					
1234	-10.3851	-14.8423	-10.8358	-0.9281	16.2272
123	1.9741	-7.4961	-2.6870		12.1572
124	2.1806	-7.3361	-2.7236		12.2463
134	2.0846	-7.3913	-2.7133		12.1892
234	1.9158	-7.4610	-2.7121		12.0888
TETRÁMERO 1 BASE QZP					
1234	-6.0422	-11.1996	-10.1887	-0.9099	16.2560
123	3.9702	-5.6883	-2.5238		12.1823
124	4.1892	-5.5057	-2.5671		12.2620
134	4.0963	-5.5628	-2.5495		12.2086
234	3.9702	-5.6425	-2.5483		12.1151
TETRÁMERO 2 BASE DZP					
1234	7.0518	-14.6408	-9.1508	-0.7725	31.6160
123	14.0554	-7.2445	-2.2515		23.5513
124	13.6701	-7.3116	-2.2540		23.2357
134	14.8404	-7.3091	-2.3399		24.4894
234	13.8496	-7.4164	-2.3054		23.5714
TETRÁMERO 2 BASE QZP					
1234	10.2678	-11.5981	-8.5754	-0.7442	31.1855
123	15.4534	-5.6902	-2.1103		23.2539
124	14.9816	-5.8307	-2.1122		22.9245
134	16.1644	-5.8106	-2.1937		24.1688
234	15.1855	-5.8646	-2.1592		23.2093

energías y no aditividades en kcal/mol; E_m^{int} = energía de interacción del agregado m; $\sum 2c$ = suma de energías de dos cuerpos; na^n = no aditividad de n cuerpos

En las tablas 3.10 y 3.11 pueden verse varios puntos importantes:

- 1- aunque las energías de interacción de los dímeros, trímeros, y tetrameros son distintas con las dos bases, las no aditividad de tres y cuatro cuerpos son muy semejantes (≈ 0.2 kcal/mol de diferencia). Esto no es una coincidencia para estos dos tetrameros, ya que este mismo comportamiento se observó anteriormente en el análisis de varios tetrameros de moléculas de agua en la geometría experimental⁵¹. Por lo tanto uno puede esperar una buena aproximación en el cálculo de las no aditividad con la base DZP aunque los valores sean pequeños.
- 2- usando cualquiera de los dos esquemas para calcular la no aditividad, lo que estabiliza al tetramero es la no aditividad de tres cuerpos.
- 3- el esquema que no considera explícitamente a las deformaciones moleculares converge con menos rapidez que el esquema que sí las considera; en el primer caso, la deformación molecular aparece en la no aditividad de cuatro cuerpos. Este es el mismo comportamiento que encontramos en la tabla 3.9 para el potencial MCHOv, dato que indica que el problema de no convergencia no es un defecto del potencial, sino de la expansión de muchos cuerpos.
- 4- los dímeros de estos tetrameros NO son estables. Las energías de interacción para dímeros, trímeros y el tetramero 2 son positivas. El MCHOv sólo dio energías positivas para dos de los dímeros del tetramero 1, y las energías de interacción predichas para los trímeros son muy negativas. Un detalle que no debe perderse de vista es que MCHOv está ajustado a superficies de dímeros no deformados calculadas con CI, y todos los cálculos que hicimos en esta sección están a nivel HF-SCF. Lo que puede asegurarse es que el potencial MCHOv, en su versión actual, no es capaz de reproducir la no aditividad de tres y cuatro cuerpos *ab initio* del tetramero 1 relajado.

También minimizamos un dímero por MC con MCHOv. Este dímero fue calculado después con la base DZP. Mientras que para MCHOv este dímero es ≈ 0.3 kcal/mol más estable que el más estable de MCHO, a nivel HF-SCF con esta base es ≈ 0.4 kcal/mol menos estable que el dímero más estable según Szalewicz *et al.*⁵⁰ para molécula rígida calculado con la misma base.

El análisis hecho hasta ahora indica que la no aditividad de tres cuerpos es la responsable de la estabilización de agregados construidos con monómeros deformados, ya que los dímeros que se calcularon no son estables. El hecho de que los monómeros deformados no sean estables es de esperarse, pero el comportamiento de los dímeros merece un estudio más profundo.

La mayoría de los modelos vibracionales para el agua líquida predicen que el dímero de agua de moléculas deformadas es más estable que el de moléculas no deformadas (tablas 3.12 y 3.13). Una excepción notable es el potencial YMD-V^{22,23} que es una corrección a nivel HF-SCF a la superficie MCY²⁶.

El cambio más notable en longitudes de enlace que ocurre al juntarse dos moléculas de agua se da en el enlace OH que dona el protón para la formación del puente de H. En el hielo la elongación llega a ser de 0.06 ua³⁰. El otro enlace de la molécula donadora se encoge ligeramente y los enlaces de la molécula aceptora se estiran un poco. Los valores predichos por los distintos modelos se muestran en la tabla 3.13 y la longitud de los enlaces

en el monómero se da en la tabla 3.12.

TABLA 3.12: POTENCIALES VIBRACIONALES

SUPERFICIES VIBRACIONALES USADAS EN POTENCIALES AGUA-AGUA:

	nombre	ref
<i>ab initio</i>	BSP	29
	KRS	52
empírica	KM	53
	CCL	54
	HMS	55
	RW	2
	EK	56
	SO	57
	S	58

POTENCIALES VIBRACIONALES

potencial	# sitios	intra	ajuste	r_{OH}	θ	ref
RWK2-M	4	RW,KM	empírico	1.8088	104.52	2-4,9
RWK2-K	4	KM	empírico	1.8088	104.52	14
TIPS-V	4	EK	empírico	1.8141	104.50	12
MST-FP	5	EK	empírico	1.8088	104.52	59
SPC-TK	3	KM	empírico	1.8897	109.47	5
SPC-H	3	S	empírico	1.8897	109.47	10
SPC-V	3	EK	empírico	1.8897	109.47	12
PSPC-H	3	KM	empírico	1.8897	109.47	13
PSPC-M	3	KM	empírico	1.8897	109.47	13
CF1	3	EK	empírico	1.8111	104.45	16,60
CF2	3	EK	empírico	1.8111	104.45	17,18,60
MCF	3	KRS	empírico	1.8111	104.45	6
BJH	3	CCL	empírico	1.8088	104.52	8,15,60
BJH-G	3	CCL	empírico	1.8088	104.52	15,60,61
PM	3	SO	empírico	1.8111	104.45	19,20
PM6	3	SO	empírico	1.8111	104.45	21
MCY-L	4	BSP	<i>ab initio</i>	1.8096	104.59	6

MCY-I	4	BSP	ab initio	1.8096	104.59	60
MCY-B	4	BSP	ab initio	1.8096	104.59	60
MCY-C	4	BSP	ab initio	1.8096	104.59	60
YMD-V	4	CCL	ab initio	1.8088	104.52	22,23
MCHOv	6	BSP	ab initio	1.8111	104.45	*

r_{OH} y θ = longitud del enlace OH (en ua) y ángulo intramolecular (en grados) para el monómero aislado; la columna de ajuste se refiere a la parte intermolecular del potencial; para la parte intramolecular, véase la tabla inmediata anterior; # sitios: número de centros sobre los que se evalúan fuerzas y energía potencial por molécula; * = este trabajo

TABLA 3.13: LOS DÍMEROS MÁS ESTABLES DE POTENCIALES FLEXIBLES

potencial	ΔE_r	ΔE_v	r_{OH}^b	r_{OH}^{nb}	θ_D	r_{OH}	θ_A
RWK2-M	-5.90	-6.15	1.8493	1.8070	104.10	1.8155	104.70
RWK2-K	-5.90	-5.70	1.8340	1.8085	103.00	1.8130	103.80
BJH	-5.52	-5.63	1.8302	1.8087	103.05	1.8124	103.90
PM	-5.40	-6.95	1.8897	1.7971	101.86	1.8444	104.14
MCY-L	-5.87	-5.94	1.8285	1.8073	104.00	1.8186	104.50
MCY-L*	-5.87	-5.98	1.8317	1.8039	104.02	1.8168	104.53
YMD-V	-5.87	-5.58	1.8187	1.8081	104.33	1.8104	104.82
MCHOv	-5.87	-6.10	1.8631	1.7889	102.91	1.8200	103.91

energías en kcal/mol; distancias en ua y ángulos en grados; * corresponde al mínimo analítico de la superficie de interacción de dímeros¹¹; r_{OH}^b = distancia OH del enlace de la molécula donadora que forma el puente de H; r_{OH}^{nb} = distancia OH del enlace de la molécula donadora que no forma el puente de H; θ_D = ángulo de la molécula donadora; r_{OH} = distancia OH de los enlaces de la molécula aceptora; θ_A = ángulo de la molécula aceptora; ΔE_r = energía de interacción en el dímero de monómeros rígidos; ΔE_v = energía de interacción en el dímero de monómeros flexibles.

Un problema muy serio es que no hay datos experimentales para la geometría de los monómeros en un dímero de agua. Los datos de espectroscopía y de conductividad térmica se interpretaron suponiendo que la geometría de la molécula de agua no cambia al asociarse⁶²⁻⁶⁴. Lo más cercano a unos datos experimentales es el análisis hecho por el grupo de Reimers y Watts^{2-4,9,65} del espectro vibracional del dímero de agua con su potencial RWK2-M.

Una incertidumbre de 1.4 kcal/mol en la energía de interacción experimental⁶⁴ (-5.4 ± 0.7 kcal/mol) es molesta, porque esto hace que no sea fácil distinguir entre modelos buenos y malos para la interacción en el dímero de agua. En la tabla 3.14 se muestran las predicciones de los modelos *ab initio* para la energía de interacción del dímero de agua, suponiendo que no hay deformación molecular. Si hay una estabilización de ~ 0.5 kcal/mol por la deformación (valor que parece estar cercano a la realidad^{3,9,65}), prácticamente todos los modelos siguen estando dentro de las barras de error experimental, habiendo diferencias de casi 1 kcal/mol entre las energías predichas por los modelos. Si el análisis se extiende a los modelos empíricos, uno encuentra valores de energía desde -4.31 hasta -6.88 kcal/mol^{16,22,23,41,42,46,47,60,66-69} (o hasta -7 kcal/mol para MCY-L⁷⁰).

TABLA 3.14: ENERGÍA DE INTERACCIÓN DEL DÍMERO DE AGUA CALCULADA CON POTENCIALES RÍGIDOS *AB INITIO*

potencial	E_{int}	ref
MCY-I	-5.72	60,66
MCY-II	-5.87	60,66
MBY-B	-5.64	60,67
MCY-C	-5.70	10,42
YMD	-5.78	22,23
NCC	-5.19	46
NEMO	-4.96	47
MCHO	-5.87	1

energías en kcal/mol; E_{int} = energía de interacción

3.6 CURVA DE DISOCIACIÓN DEL ENLACE O-H DONADOR

En la sección anterior quedó claro que el potencial MCHOv debe reajustarse a datos de superficies de interacción y no aditividad de agregados de moléculas deformadas, desde dímeros hasta tetrámeros al menos, si queremos que el modelo MCHOv sea capaz de reproducir la energía de interacción y los cambios en la geometría de la molécula de agua que se dan al asociarse dos o más moléculas. Comenzamos calculando el agregado más simple, dímeros de agua en los que sólo variamos la longitud de un enlace, manteniendo el resto de las coordenadas fijas. Calculamos una curva de disociación del enlace O-H donador en un dímero lineal de agua por ser el enlace que sufre el cambio más notable al formarse el dímero. Además de formar parte de la nueva base de datos para reajustar el potencial, este es el modelo con relevancia biológica más sencillo para estudiar los cambios en geometría que suceden al formar un puente de H. Los cálculos se hicieron a nivel HF-SCF y CI-MP2

con la base QZP. Para calcular estas curvas hicimos lo siguiente.

Buscamos la geometría de energía mínima para el monómero en la base QZP a nivel HF-SCF. Esto lo hicimos con la intención ver cuánto se estira o encoge el enlace OH al formar el dímero, partiendo de su longitud de equilibrio, y la geometría de equilibrio depende de la base. Sólo se calcularon estructuras simétricas, haciendo una variación simultánea de las distancias de enlace OH y entre los H a intervalos de 0.03 ua alrededor de las distancias experimentales. Una vez encontrada una posible localización del mínimo en la superficie vibracional se hizo un muestreo más fino de esta zona, a intervalos de 0.005 ua. La geometría de equilibrio que encontramos tiene una longitud de enlace OH de 1.7894 ua (vs 1.808876 ua) y una distancia HH de 2.8559 ua (vs 2.860912 ua), que corresponde a un ángulo de 105.88°. En total se calcularon 75 geometrías.

Con este monómero construimos un dímero en la misma geometría que la de Szalewicz *et al.*⁵⁰; misma distancia entre O (5.67 ua), lineal y con un ángulo de 150° entre la bisectriz de la molécula aceptora y la línea que une a los dos O. Esta configuración corresponde a una optimización de geometría con moléculas rígidas hecha con una base extendida y con correlación electrónica⁵⁰. Las coordenadas del dímero construido se encuentran en el Apéndice B.

Calculamos la energía de interacción de este dímero a nivel HF-SCF, CI-MP2 y a niveles superiores de cálculo de la correlación electrónica (incluyendo tri y cuatriexcitaciones en la diagonalización). Con los programas que tenemos, cualquier esquema superior a CI-MP2 no tiene consistencia de talla, así que no puede utilizarse para calcular energías de interacción (para una discusión más detallada ver la sección 2.2.3). Por esto nos quedamos a nivel CI-MP2.

Aún dentro de CI-MP2 hubo que encontrar una manera de balancear correctamente la energía de correlación que se obtiene por diagonalización para el dímero y los monómeros. Como puede verse en las tablas 3.15 y 3.16, si está desbalanceado el tratamiento pueden obtenerse resultados muy malos para la energía de interacción. Aunque no está del todo claro si debe o no incluirse la corrección de CP, nosotros decidimos sí corregir siguiendo la recomendación de Szalewicz *et al.*⁵⁰. La energía de interacción tanto a nivel HF-SCF como CI-MP2 está muy cercana a los valores límites obtenidos con cálculos muy precisos⁵⁰: -3.87 kcal/mol a nivel HF-SCF y -4.87 kcal/mol a nivel CI-MP2 (cálculo de la tabla 3.15 con 227, 125 y 136 determinantes para el dímero, el monómero donador y el monómero aceptor, respectivamente). El valor experimental es de -5.4 ± 0.7 kcal/mol⁵⁴.

En estas tablas se pueden ver una serie de cosas interesantes. Existe la idea de que mientras mayor sea el número de determinantes que se diagonalice, mejor será el cálculo de la energía de correlación de la molécula. En la tabla 3.15 tenemos un ejemplo de cómo esto no se extiende a la mejor predicción de la energía de interacción: el caso en el que el dímero tiene 557 determinantes no da la mejor energía de interacción, porque para las condiciones del cálculo, los monómeros obtuvieron mucha más energía por diagonalización que el dímero y el cálculo está por consecuencia desbalanceado.

Aún para el dímero cuyo espacio de diagonalización tiene 557 determinantes, se puede ver que la mayor parte de la energía de correlación calculada está en la parte perturbativa ($\approx 72\%$), lo cual muestra la dificultad de tratar sistemas de van der Waals.

TABLA 3.15: ENERGÍA DE INTERACCIÓN DEL DÍMERO DE AGUA CON CORRECCIÓN CP

dim NCF	mon1 NCF	mon2 NCF	dim % diag	mon1 % diag	mon2 % diag	E_{int}^{MP2}
20	17	38	3.66	6.47	11.28	-5.35
96	66	85	9.05	13.76	18.07	-5.60
192	113	154	14.47	18.52	24.71	-5.16
200	175	175	18.04	26.44	30.03	-4.57
227	125	136	18.69	22.48	27.06	-4.87
227	93	127	18.69	6.90	7.37	-4.20 *
557	325	328	28.22	35.52	38.94	-4.07

energías en kcal/mol; dim = dímero; mon1 = monómero donador; mon2 = monómero aceptor; NCF = número de determinantes en el espacio de diagonalización; % diag = porcentaje de la energía de correlación calculada (diagonalización + perturbación) obtenida por la diagonalización; * los determinantes de los monómeros provienen del espacio de diagonalización del dímero; E_{int}^{MP2} = energía de interacción calculada a nivel CI-MP2

TABLA 3.16: ENERGÍA DE INTERACCIÓN DEL DÍMERO DE AGUA SIN CORRECCIÓN CP

dim NCF	dim80 NCF	dim % diag	dim80 % diag	E_{int}^{MP2}
227	345	18.69	42.51	-2.64
110	173	12.84	33.67	-5.15
119	173	13.45	33.67	-4.72
121	173	13.64	33.67	-4.71
200	199	18.04	36.23	-4.55
200	173	18.04	33.67	-4.80
200	163	18.04	32.68	-4.96
200	153	18.04	31.55	-4.99

energías en kcal/mol; dim80 = dímero separado 80 ua

Si se quiere comparar, hay cálculos *ab initio* de dímeros lineales de moléculas rígidas que incluyen correlación electrónica^{26,35,42,50,71-73}, y los valores de energía de interacción van desde -4.50 hasta -5.70 kcal/mol.

Habiendo establecido el nivel de cálculo de la energía de correlación, calculamos el dímero sin distorsionar y el dímero con $\Delta r_{OH}^b = 0.1$ ua, unimos los espacios de interacción de configuraciones sobre los que se diagonalizaría y con esta suma de espacios calculamos toda la curva a nivel CI-MP2 con el programa CIPSI²⁷, 23 puntos a intervalos de 0.005 ua en Δr_{OH}^b , muestreando el intervalo comprendido entre -0.01 ua y +0.1 ua. Unimos los espacios de diagonalización para tratar de una manera uniforme todas las geometrías y no tener oscilaciones artefactuales en la energía de interacción a nivel CI-MP2. El cálculo de la curva a nivel HF-SCF se hizo con el programa PSHF²⁴.

En la tabla 3.17 y la figura 3.5 se muestran las curvas a nivel HF-SCF y CI-MP2. En las abscisas se dan los desplazamientos respecto a la longitud de enlace del monómero en equilibrio (en ua) y en las ordenadas se graficaron las diferencias en energía del complejo (en kcal/mol) respecto a la geometría sin deformar.

TABLA 3.17: CURVAS DE DISOCIACIÓN DEL ENLACE OH DONADOR

Δr	ΔE_{HF-SCF}	ΔE_{CI-MP2}
-0.010	0.0446	0.1679
-0.005	0.0176	0.0791
0.000	0.0000	0.0000
0.005	-0.0075	-0.0692 mín HF-SCF
0.010	-0.0056	-0.1289
0.015	0.0050	-0.1792
0.020	0.0257	-0.2205
0.025	0.0546	-0.2529
0.030	0.0922	-0.2765
0.035	0.1387	-0.2913
0.040	0.1939	-0.2978 mín CI-MP2
0.045	0.2573	-0.2964
0.050	0.3288	-0.2869
0.055	0.4085	-0.2696
0.060	0.4957	-0.2448
0.065	0.5905	-0.2126
0.070	0.6934	-0.1733
0.075	0.8038	-0.1272
0.080	0.9212	-0.0742
0.085	1.0454	-0.0146
0.090	1.1772	0.0515

0.095
0.100

1.3159
1.4608

0.1240
0.2024

energías en kcal/mol; distancias en ua

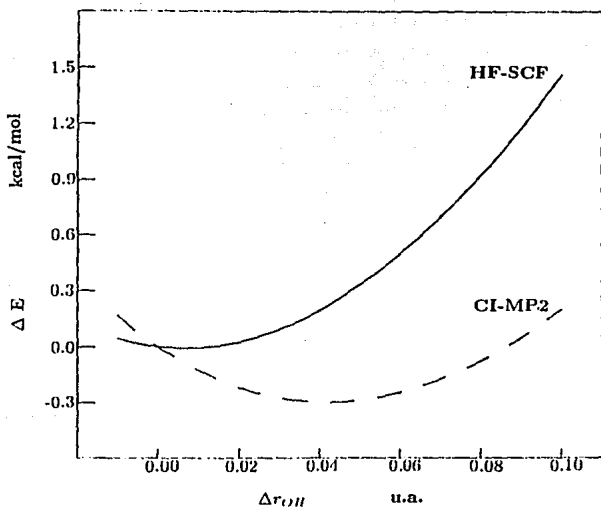


Figura 3.5: curvas de disociación del enlace OH donador en un dímero lineal de agua; la línea continua corresponde al cálculo a nivel HF-SCF y la línea a trazos al cálculo CI-MP2. En las ordenadas se muestra la diferencia en energía (en kcal/mol) de los dímeros deformados respecto al dímero sin deformar; en las abscisas se muestra la deformación del enlace OH donador (en ua) respecto a la longitud de ese enlace optimizado a nivel HF-SCF con la base QZP.

De estas curvas puede concluirse lo siguiente:

- 1- a nivel HF-SCF este enlace no se estira más de 0.005 ua, con una energía -0.007 kcal/mol más baja que la geometría de referencia.
- 2- a nivel CI-MP2 la deformación correspondiente a la energía mínima es de 0.04 ua, con una energía de estabilización respecto a la geometría de referencia de ~ -0.3 kcal/mol.

Cabe señalar que no hicimos una optimización completa de todos los grados de libertad del dímero, de manera que no tenemos un valor ni para la energía total de deformación ni para la geometría de equilibrio. En la tabla 3.18 se muestran cuatro cálculos a nivel HF-SCF reportados con distintas calidades de base para optimizaciones globales de geometría del dímero de agua. Como las geometrías de energía mínima de los monómeros no son idénticas a la experimental, las longitudes de enlace en el complejo nada más dan una idea cualitativa de la dirección del cambio al formar el complejo. De los cuatro cálculos, sólo el último tiene corrección CP, y además falta considerar la contribución de la correlación electrónica.

**TABLA 3.18: CÁLCULOS AB INITIO DE DÍMEROS
OPTIMIZACIÓN COMPLETA DE GEOMETRÍA**

base	ΔE_v	r_{OH}^b	r_{OH}^{b*}	θ_D	r_{OH}	θ_A	ref
4-31G	-8.23	1.8104	1.7933	111.32	1.7971	111.96	74
6-31G*	-5.62	1.7990	1.7896	105.44	1.7915	105.91	74
6-31G**	-5.54	1.7915	1.7801	106.20	1.7839	106.22	74
QZ2P(3+,3+)	-3.85	1.7839	1.7744	106.40	1.7782	106.60	75

energías en kcal/mol; distancias en ua; ángulos en grados; nomenclatura de la tabla 3.13.

Los resultados de la curva de disociación indican que no es de esperarse que los dímeros deformados sean más estables al incluir la energía de correlación en el análisis de la no aditividad. Esto se debe a que las deformaciones de los dímeros de los tetrámeros analizados son tan grandes que se encuentran en la parte de la curva CI-MP2 donde los dímeros de moléculas deformadas son menos estables que el dímero de referencia. Por lo tanto, todo parece indicar que la no aditividad de tres cuerpos es en efecto responsable de la deformación de los enlaces de la molécula de agua que se encuentran en fases condensadas como el agua líquida y los diversos tipos de hielo. Este resultado concuerda con datos preliminares de Yoon *et al.*²³ para el hielo Ih.

CAPITULO 4: CONCLUSIONES Y PERSPECTIVAS

4.1 CONCLUSIONES

- Una manera de hacer estudios moleculares es el esquema de simulación global propuesto por Clementi^{1,2}, cuya principal virtud es que se conocen las aproximaciones que se hacen en cada paso. Esto permite un refinamiento progresivo del modelo que representa al sistema en estudio. En general el refinamiento va a la par del aumento en capacidad computacional. Este esquema ha mostrado ser muy conveniente para el estudio del agua líquida.
- El sistema al cual le aplicamos esta metodología es el agua líquida. El agua es un sistema complejo muy estudiado, tanto teórica como experimentalmente, de manera que hay muchos datos contra los cuales validar el modelo y la metodología. A pesar de ser un sistema tan estudiado, las propiedades del agua no han sido entendidas completamente. La obtención de un potencial vibracional analítico confiable para el agua pura permitirá, además de contribuir al entendimiento de las propiedades del agua, la ampliación de una serie de estudios sobre la hidratación de diversos solutos de relevancia bioquímica.
- Los resultados de esta tesis son:
 - a . La forma analítica de MCHO_v permite tratar vibraciones moleculares sin recurrir a términos cruzados, dando un muy buen ajuste a la superficie vibracional de Bartlett *et al.*³.
 - b . El potencial MCHO_v, construido a partir del potencial MCHO⁴ y la superficie vibracional de Bartlett *et al.*³, exagera la energía potencial de interacción en un 30% respecto al valor experimental en simulaciones por MC y DM. La exageración de la corrección a la energía de interacción por incluir grados de libertad intramoleculares parece ser un problema común en las extensiones vibratoriales de otros potenciales para el agua líquida.
 - c . En lugar de reparametrizar el potencial para que reproduzca los resultados experimentales, regresamos a analizar el comportamiento del potencial y a comparar este comportamiento con datos *ab initio*.
 - d . El potencial MCHO_v subestima la no aditividad en la interacción de tetrámeros de moléculas de agua deformadas.
 - e . Para corregir este potencial parece necesario incluir en las superficies de ajuste puntos correspondientes a dímeros, trímeros y tetrámeros construidos con moléculas de agua deformadas.
 - f . Del análisis *ab initio* de la no aditividad de dos tetrámeros relajados con MCHO_v se encuentra que la no aditividad de tres cuerpos es la responsable de la estabilidad del agregado.

- g . La expansión de muchos cuerpos que no considera explícitamente la deformación molecular no converge; la expansión que sí la considera explícitamente sí parece converger, aunque hay que calcular la contribución de cinco cuerpos en pentámeros de moléculas deformadas para corroborar esta convergencia.
- h . Los cálculos a nivel HF-SCF y MRCI-MP2 de la elongación del enlace OH donador en un dímero lineal de agua muestran que este enlace se deforma menos en el dímero (0.04 ua) que en el hielo (0.06 ua)⁵, indicando la importancia de la no aditividad para estabilizar estructuras con monómeros deformados.
- La metodología utilizada en este trabajo todavía no está lista para ser totalmente automatizada; es un trabajo artesanal en el que hay que ser muy cuidadoso en cada paso y estar regresando constantemente para garantizar dentro de lo posible que no hay inconsistencias en el proceso de modelado y simulación. El hecho de no ser un proceso automático tiene la ventaja de hacer más evidentes las aproximaciones hechas y las limitaciones del modelo. Por lo tanto, los resultados que se obtienen de paquetes comerciales de simulación, donde todo está automatizado, deben analizarse con cuidado, especialmente para no extrapolar los resultados obtenidos a situaciones en los que el modelo utilizado ya no es válido.

4.2 PERSPECTIVAS

Este trabajo ha dejado abiertas varias líneas sobre las que se puede trabajar. En cuanto a la no aditividad de agregados de moléculas deformadas, lo siguiente es, por un lado, calcular más tetrámeros y algunos pentámeros para estudiar la convergencia de la expansión de muchos cuerpos y corroborar el papel estabilizador de la no aditividad de tres cuerpos. Por otro lado, todo el análisis se ha hecho a nivel HF-SCF y, aunque la energía de correlación es básicamente aditiva, debe incluirse en las energías de los monómeros y de interacción de dímeros, que son la parte aditiva de la expansión de muchos cuerpos.

En cuanto al refinamiento del potencial, con los puntos calculados en el inciso anterior se puede intentar un reajuste de MCHOv y probarlo en simulaciones por MC. Un proyecto a más largo plazo es el cálculo de nuevas superficies de dos, tres y cuatro cuerpos (o más si el análisis de muchos cuerpos así lo indica), incluyendo geometrías deformadas, todas con la misma base y el mismo nivel de aproximación, el ajuste a estas superficies y su prueba en simulaciones por MC y DM.

La perspectiva más interesante en cuanto al modelo es la posibilidad de extenderlo para que permita la disociación molecular. El potencial vibracional de MCHOv tiene la forma correcta para producir disociación, si bien fue ajustado nada más en una zona muy cercana al mínimo del potencial. Este modelo disociable haría accesible el estudio de la química del agua bajo un esquema de primeros principios.

En cuanto al campo de simulaciones numéricas de sistemas biomoleculares, las pers-

pectivas son muy optimistas. A pesar de las limitaciones técnicas y conceptuales, el tamaño y la complejidad de los sistemas estudiados ha crecido en los últimos años, contribuyendo de manera importante al entendimiento de estos sistemas. Por ejemplo, ahora ya se puede simular el efecto de algunos antivirales sobre la cápside de un rinovirus^{6,7}. Las simulaciones numéricas están demostrando que sí pueden proporcionar respuestas inteligentes a problemas moleculares de interés biológico y médico.

REFERENCIAS

Resumen

- 1 . E.Clementi *Lecture Notes in Chemistry # 19: Computational Aspects for Large Chemical Systems* (1980) Springer-Verlag, Heidelberg.
- 2 . E.Clementi *Supercomputers and Supercomputing in Chemical Research and Development in Theoretical Biochemistry & Molecular Biophysics Volume 1: DNA* edited by D.L.Beveridge and R.Lavery (1991) Adenine Press, New York, U.S.A.
- 3 . H.Saint-Martin, C.Medina-LLanos and I.Ortega-Blake. *J.Chem.Phys.* (1990) **93**, 6448-6452.
- 4 . E.Whalley *J.Glaciology* (1978) **21**, 13.

Capítulo 1

- 1 . H.G.Hempling *Osmosis: the push and pull of life in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 2 . A.L.Lehninger *Biochemistry* (1977) Worth Publishers, Inc. New York.
- 3 . H.Saint-Martin and I.Ortega-Blake. *The Microscopic Structure and Dynamics of Water at a Surface in Biomolecules in Organic Solvents.* edited by A.Gómez-Puyou (1992) CRC Press, Boca Ratón.
- 4 . M.M.Teeter *Annu.Rev.Biophys.Biophys.Chem.* (1991) **20**, 577-600.
- 5 . G.M.Clore and A.M.Gronenborn. *Annu.Rev.Biophys.Biophys.Chem.* (1991) **20**, 29-63.
- 6 . H.F.J.Savage and J.L.Finney. *New Approaches in Studying Biomolecule - Water Interactions in Biomolecules in Organic Solvents.* edited by A.Gómez-Puyou (1992) CRC Press, Boca Ratón.
- 7 . F.G.Bryant *NMR Relaxation and Water Dynamics at the Protein Surface and Water at the Protein Surface in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 8 . B.Halle, L.Piculell, G.Carlström, T.Anderson, H.Wennerström and B.Lindman. *NMR of Water Nuclei in Heterogeneous Systems. Relaxation Theory and Oxygen-17 Data from Aqueous Solutions of Proteins, Polyelectrolytes and Micelles.* in *Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 9 . A.J.Russell, S.Chatterjee, I.Rapanovich and J.G.Goodwin, Jr. *Mechanistic Enzymology in Anhydrous Organic Solvents in Biomolecules in Organic Solvents.* edited by

- A.Gómez-Puyou (1992) CRC Press, Boca Ratón.
- 10 . G.Garza-Ramos, D.A.Fernández-Velasco, M.Tuena de Gómez-Puyou and A.Gómez-Puyou. *Water and Enzymes in Organic Solvents in Biomolecules in Organic Solvents*. edited by A.Gómez-Puyou (1992) CRC Press, Boca Ratón.
 - 11 . S.T.Ferreira and E.Gratton. *Enzyme Activity in Reverse Micelles: Roles of Diffusion and Extent of Hydration in Biomolecules in Organic Solvents*. edited by A.Gómez-Puyou (1992) CRC Press, Boca Ratón.
 - 12 . E.Escamilla, M.Contreras, L.Escobar and G.Ayala. *Biological Electron Transport in Reverse Micellar Systems. From Enzymes to Cells*. in *Biomolecules in Organic Solvents*. edited by A.Gómez-Puyou (1992) CRC Press, Boca Ratón.
 - 13 . N.Pfammatter, M.Famiglietti, A.Hochköppler and P.L.Luisi. *Solubilization of Microorganisms in Organic Solvents by Reverse Micelles in Biomolecules in Organic Solvents*. edited by A.Gómez-Puyou (1992) CRC Press, Boca Ratón.
 - 14 . J.L.Finney and P.L.Poole. *Comments Mol.Cell.Biophys.* (1984) 2, 129-157.
 - 15 . J.H.Crowe and L.M.Crowe. *Cryobiology* (1981) 18, 613.
 - 16 . A.M.Klibanov *Trends Biochem. Sci.* (1989) 14, 141.
 - 17 . A.Žaks and A.M.Klibanov. *J.Am.Chem.Soc.* (1986) 108, 2767.
 - 18 . T.Sakurai, A.Margolin, A.J.Russell and A.M.Klibanov. *J.Am.Chem.Soc.* (1988) 110, 7236.
 - 19 . *Theoretical Biochemistry & Molecular Biophysics Volumes I and II*, edited by D.L.Beveridge and R.Lavery (1991) Adenine Press, New York.
 - 20 . H.J.C.Berendsen, J.P.M.Postma, W.F.van Gunsteren and J.Hermans. *Interaction Models for Water in Relation to Protein Hydration in Intermolecular Forces*, edited by B.Pullman (1981), Reidel, Dordrecht.
 - 21 . W.L.Jorgensen, J.Chandrasekhar, J.D.Madura, R.W.Impey and M.L.Klein. *J. Chem.Phys.* (1983) 79, 926-935.
 - 22 . M.M.Teeter *Annu.Rev.Biophys.Biophys.Chem.* (1991) 20, 577-600.
 - 23 . J.W.Brady *The Development of Molecular Mechanics Parameters for Carbohydrates* presentado en *Expanding Frontiers in Polypeptide and Protein Structural Research* Whistler, B.C., Canada (1990).
 - 24 . The FASEB Journal (1992) 6,.
 - 25 . D.T.Edmonds *Some Electrical Properties of Small Isolated Water Molecule Arrays in Biophysics of Water*. edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.

- 26 . H.S.Frank and W.Y.Wen. *Disc.Faraday Soc.* (1957) **24**, 133.
- 27 . C.Tanford *The Hydrophobic Effect: Formation of Micelles and Biological Membranes.* (1980) John Wiley & Sons, New York.
- 28 . D.F.Evans and D.D.Miller. *A Reappraisal of the Role of Water in Promoting Amphiphilic Assembly and Structure in Water Science Reviews Vol 4.* edited by F.Franks (1989) Cambridge University Press, Great Britain.
- 29 . M.F.Perutz *Foreword in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 30 . G.Careri *Molecular Hydration and Biological Function in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 31 . B.Gavish and M.M.Werber. *Biochemistry* (1979) **18**, 1269.
- 32 . R.B.Gregory and A.Rosenberg. *The Influence of the Dynamic Properties of Water on Protein Fluctuations in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 33 . S.N.Timasheff, T.Arakawa, H.Inoue, K.Gekko, M.J.Gorbunoff, J.C.Lee, G.C.Na, E.P.Pittz and V.Prakash. *The Role of Solvation in Protein Structure in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 34 . B.Schobert *Evidence for a Protein Stabilizing Mechanism in Plant Cells under Water Stress in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 35 . J.H.Crowe, L.M.Crowe and D.W.Deamer. *Hydration Dependent Phase Changes in a Biological Membrane in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 36 . W.Drost-Hansen *The Occurrence and Extent of Vicinal Water in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 37 . W.Derbyshire *Dynamics of Water in Cellular Systems in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 38 . W.Drost-Hansen *Phys.Chem.Liquids* (1977) **7**, 295.
- 39 . J.S.Clegg *Alternative Views on the Role of Water in Cell Function in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 40 . R.P.Rand and V.A.Parsegian. *Annu.Rev.Physiol.* (1986) **48**, 201.
- 41 . S.Marčelja *Molecular Mechanism of the Hydration Force in Biophysics of Water.* edited by F.Franks and S.Mathias (1982) John Wiley & Sons, New York.
- 42 . H.-J.Butt *Biophys.J.* (1991) **60**, 777.

- 43 . J.H.Crowe and L.M.Crowe. *Lyotropic Effects of Water on Phospholipids in Water Science Reviews Vol 5*. edited by F.Franks (1990) Cambridge University Press, Great Britain.
- 44 . H.J.C.Berendsen *Curr.Op.Struc.Biol.* (1991) **1**, 191-195.
- 45 . J.A.McCammon *Curr.Op.Struc.Biol.* (1991) **1**, 196-200.
- 46 . W.C.Cornell, A.E.Howard and P.Kollman. *Curr.Op.Struc.Biol.* **1**, 201-212.
- 47 . H.Saint-Martin, C.Medina-LLanos and I.Ortega-Blake. *J.Chem.Phys.* (1990) **93**, 6448-6452.
- 48 . E.Clementi *Lecture Notes in Chemistry # 19: Computational Aspects for Large Chemical Systems* (1980) Springer-Verlag, Heidelberg.
- 49 . E.Clementi *Supercomputers and Supercomputing in Chemical Research and Development in Theoretical Biochemistry & Molecular Biophysics Volume 1: DNA* edited by D.L.Beveridge and R.Lavery (1991) Adenine Press, New York, U.S.A.

Capítulo 2

- 1 . E.Clementi *Lecture Notes in Chemistry # 19: Computational Aspects for Large Chemical Systems* (1980) Springer-Verlag, Heidelberg.
- 2 . E.Clementi *Supercomputers and Supercomputing in Chemical Research and Development in Theoretical Biochemistry & Molecular Biophysics Volume 1: DNA* edited by D.L.Beveridge and R.Lavery (1991) Adenine Press, New York, U.S.A.
- 3 . W.J.Hehre, L.Radom, P.v.R.Schleyer and J.A.Pople. *Ab Initio Molecular Orbital Theory*. (1986) John Wiley & Sons, U.S.A.
- 4 . C.E.Dykstra *Studies in Physical and Theoretical Chemistry # 58: Ab Initio Calculation of the Structures and Properties of Molecules* (1988) Elsevier, The Netherlands.
- 5 . J.Sauer *Chem.Rev.* (1989) **89**, 199-225.
- 6 . P.W.Atkins *Quanta: a Handbook of Concepts*. Oxford Chemistry Series (1974) Clarendon Press, Oxford.
- 7 . B.Huron, J.P.Malrieu and P.Rancurel. *J.Chem.Phys.* (1973) **58**, 5745-5759.
- 8 . L.Szasz *Pseudopotential Theory of Atoms and Molecules* (1985) John Wiley & Sons, U.S.A.
- 9 . Ph.Durand and J.C.Barthelat. *Theoret.Chim.Acta* (1975) **38**, 283. J.C.Barthelat, Ph.Durand and A.Serafini. *Mol.Phys.* (1977) **33**, 179.
- 10 . S.Rybak, B.Jeziorski and K.Szalewicz. *J.Chem.Phys.* (1991) **95**, 6576-6601.

- 11 . A.J.Grant and B.T.Pickup. *Int.J.Quantum Chem.* (1991) **40**, 97-126.
- 12 . S.F.Boys and F.Bernardi. *Mol.Phys.* (1970) **19**, 553-566.
- 13 . M.M.Szczęśniak and S.Scheiner. *J.Chem.Phys.* (1986) **84**, 6328-6335.
- 14 . A.E.Reed, F.Weinhold, L.A.Curtiss and D.J.Pochatko. *J.Chem.Phys.* (1986) **84**, 5687-5705.
- 15 . S.K.Loushin, S.-Y.Liu and C.E.Dykstra. *J.Chem.Phys.* (1986) **84**, 2720-2725.
- 16 . J.E. del Bene and I.Shavitt. *Int.J.Quantum Chem.:Quantum Chem.Symp.* (1989) **23**, 445-452.
- 17 . G.Chalasiński, M.M.Szczęśniak, P.Cieplak and S.Scheiner. *J.Chem.Phys.* (1991) **94**, 2873-2883.
- 18 . K.Szalewicz, S.J.Cole, W.Kolos and R.J.Bartlett. *J.Chem.Phys.* (1988) **89**, 3662-3673.
- 19 . L.Gil-Adalid and I.Ortega-Blake. *J.Chem.Phys.* (1991) **94**, 3748-3752.
- 20 . E.Clementi, W.Kolos, G.C.Lie and G.Ranghino. *Int.J.Quantum Chem.* (1983) **17**, 337.
- 21 . K.S.Hermansson *J.Chem.Phys.* (1988) **89**, 2149-2159.
- 22 . P.Habitz, P.Bagus, P.Siegbahn and E.Clementi. *Int.J.Quantum Chem* (1983) **23**, 1803.
- 23 . W.F.van Gunsteren *Classical Molecular Dynamics Simulations: Algorithms and Applications, Stochastic Dynamics and Free Energies* pp 136-174.
- 24 . D.L.Beveridge and F.M.DiCapua. *Annu.Rev.Biophys.Biophys.Chem.* (1989) **18**, 431-492.
- 25 . B.Jayaram and D.L.Beveridge. *J.Phys.Chem.* (1990) **94**, 7288.
- 26 . J.C.Owicky and H.A.Scheraga. *J.Am.Chem.Soc.* (1977) **99**, 7403-7412.
- 27 . P.H.Berens, D.H.J.Mackay, G.M.White and K.R.Wilson. *J.Chem. Phys.* (1983) **79**, 2375-2389.
- 28 . G.Ciccotti, D.Frenkel and I.R.McDonald. *Simulations of Liquids and Solids: Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics* (1987) North Holland, Netherlands.
- 29 . E.Clementi *J.Phys.Chem.* (1985) **89**, 4426.
- 30 . E.Clementi, G.C.Lie, L.Hannon, D.C.Rapaport and M.Wojcik in *Structure and Dynamics of Nucleic Acids, Proteins and Membranes* edited by E.Clementi and S.Chin

(1986) Plenum Publishing Corporation.

- 31 . E.Clementi and G.Corongiu. *Int.J.Quantum Chem.: Quantum Biol. Symp.* (1983) **10**, 31.
- 32 . P.P.Ewald *Annln.Phys.* (1921) **64**, 253.
- 33 . T.P.Straatsma and H.J.C.Berendsen. *J.Chem.Phys.* (1988) **89**, 5876-5886.
- 34 . D.J.Adams and I.R.McDonald. *Mol.Phys.* (1976) **32**, 931.
- 35 . S.Nosé and M.L.Klein. *Mol.Phys.* (1983) **50**, 1055-1076.
- 36 . J.A.Barker and R.O.Watts. *Mol.Phys.* (1973) **26**, 789.
- 37 . M.Neumann *Mol.Phys.* (1983) **50**, 841-858.
- 38 . T.A.Andrea, W.C.Swope and H.C.Andersen. *J.Chem.Phys.* (1983) **79**, 4576-4584.
- 39 . T.Schlick, S.Figueroa and M.Mezei. *J.Chem.Phys.* (1991) **94**, 2118-2129.
- 40 . N.Metropolis, A.W.Rosenbluth, M.N.Rosenbluth, A.H.Teller and E.Teller. *J.Chem. Phys.* (1953) **21**, 1087-1092.
- 41 . I.M.Sóbol *Método de Monte Carlo* (1976) Mir, Moscú.
- 42 . H.Saint-Martin Posada *No Aditividad en un Potencial Analítico Intermolecular. La Interacción Agua-Agua*. Tesis de Maestría en Investigación Biomédica Básica, (1988) UNAM
- 43 . B.J.Alder and T.E.Wainwright. *J.Chem.Phys.* (1957) **27**, 1208-1209.
- 44 . J.Kestin and J.R.Dorfman. *A Course in Statistical Thermodynamics* (1971) Academic Press, New York.
- 45 . L.Verlet *Phys.Rev.* (1967) **159**, 98-103.
- 46 . G.W.Gear *Numerical Initial Value Problems in Ordinary Differential Equations* (1971) Prentice Hall, New Jersey.
- 47 . J.-P.Ryckaert, G.Ciccotti and H.J.C.Berendsen. *J.Compu.Phys.* **23**, 327-341.

Capítulo 3

- 1 . H.Saint-Martin, C.Medina-LLanos and I.Ortega-Blake. *J.Chem.Phys.* (1990) **93**, 6448-6452.
- 2 . J.R.Reimers and R.O.Watts. *Mol.Phys.* (1984) **52**, 357-381.
- 3 . J.R.Reimers and R.O.Watts. *Chem.Phys.* (1984) **85**, 83-112.

- 4 . J.R.Reimers and R.O.Watts. *Chem.Phys.* (1984) **91**, 201-223.
- 5 . K.Toukan and A.Rahman. *Phys.Rev.B* (1985) **31**, 2643-2648.
- 6 . G.C.Lie and E.Clementi. *Phys.Rev.A* (1986) **33**, 2679-2693.
- 7 . P.Demontis, G.B.Suffritti, E.S.Fois and A.Gamba. *Chem.Phys.Lett.* (1986) **127**, 456-461.
- 8 . Z.Slanina *Chem.Phys.Lett.* **172**, 367-371.
- 9 . D.F.Coker, R.E.Miller and R.O.Watts. *J.Chem.Phys.* (1985) **82**, 3554-3562.
- 10 . T.Schlick, S.Figueroa and M.Mezei. *J.Chem.Phys.* (1991) **94**, 2118-2129.
- 11 . O.Teleman, B.Jönsson and S.Engström. *Mol.Phys.* (1987) **60**, 193-203.
- 12 . L.X.Dang and B.M.Pettitt. *J.Phys.Chem.* (1987) **91**, 3349-3354.
- 13 . A.Wallqvist *Chem.Phys.* (1990) **148**, 439-449.
- 14 . R.O.Watts *Chem.Phys.* (1977) **26**, 367-377.
- 15 . P.Bopp, G.Jancsó and K.Heinzinger. *Chem.Phys.Lett.* (1983) **98**, 129-133.
- 16 . H.L.Lemberg and F.H.Stillinger. *J.Chem.Phys.* (1975) **62**, 1677-1690.
- 17 . A.Rahman, F.H.Stillinger and H.L.Lemberg. *J.Chem.Phys.* **63**, 5223-5230.
- 18 . F.H.Stillinger and A.Rahman. *J.Chem.Phys.* (1978) **68**, 666-670.
- 19 . F.H.Stillinger and C.W.David. *J.Chem.Phys.* (1978) **69**, 1473-1484.
- 20 . F.H.Stillinger and C.W.David. *J.Chem.Phys.* (1980) **73**, 3384-3389.
- 21 . T.A.Weber and F.H.Stillinger. *J.Phys.Chem.* (1982) **86**, 1314.
- 22 . B.J.Yoon and M.S.Jhon. *Chem.Phys.Lett.* (1991) **178**, 253-258.
- 23 . B.J.Yoon, K.Morokuma and E.R.Davidson. *J.Chem.Phys.* (1985) **83**, 1223-1231.
- 24 . M.Dupuis, J.Rys and H.F.King. *Quantum Chem.Prog.Exchange* (1977) **11**, 335; adaptación para pseudopotenciales por J.P.Daudey, Université Paul Sabatier, Lab. de Physique Quantique (1989).
- 25 . J.G.C.M.van Duijneveldt-van de Rijdt, and F.B.van Duijneveldt. revisión sobre bases atómicas, Depto. de Química, Universidad de Varsovia, Polonia.
- 26 . O.Matsuoka, E.Clementi and M.Yoshimine. *J.Chem.Phys.* (1976) **64**, 1351-1361.
- 27 . B.Huron, J.P.Malrieu and P.Rancurel. *J.Chem.Phys.* (1973) **58**, 5745-5759.

- 28 . VA05AD. Harwell Subroutine Library (1974) Harwell Laboratories, Harwell, U.K.
- 29 . R.J.Bartlett, I.Shavitt and G.D.Purvis III. *J.Chem.Phys.* (1979) **71**, 281-291.
- 30 . E.Whalley *J.Glaciology* (1978) **21**, 13.
- 31 . W.F.van Gunsteren *Classical Molecular Dynamics Simulations: Algorithms and Applications, Stochastic Dynamics and Free Energies* pp 136-174.
- 32 . L.Verlet *Phys.Rev.* (1967) **159**, 98-103.
- 33 . G.Ciccotti, D.Frenkel and I.R.McDonald. *Simulations of Liquids and Solids: Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics* (1987) North Holland, Netherlands.
- 34 . V.G.Dashevsky and G.N.Sarkisov. *Mol.Phys.* (1974) **27**, 1249-1269. E.S.Rudakov *Thermodynamics of Intermolecular Interaction* (1968) Nauka Publishing House, Moscú.
- 35 . U.Niesar, G.Corongiu, E.Clementi, G.R.Kneller and D.K.Bhattacharya. *J.Phys. Chem.* (1990) **94**, 7949-7956.
- 36 . E.Clementi, G.Corongiu, G.C.Lie, U.Niesar and P.Procacci. *Modern Techniques in Computational Chemistry* edited by E.Clementi (1989), ESCOM, Leiden.
- 37 . J.R.Reimers, R.O.Watts and M.L.Klein. *Chem.Phys.* (1982) **64**, 95.
- 38 . F.H.Stillinger and A.Rahman. *J.Chem.Phys.* (1974) **60**, 1545.
- 39 . H.J.C.Berendsen, J.P.M.Postma, W.F.van Gunsteren and J.Hermans. *Interaction Models for Water in Relation to Protein Hydration in Intermolecular Forces*, edited by B.Pullman (1981), Reidel, Dordrecht.
- 40 . P.Ahlström, A.Wallqvist, S.Engström and B.Jónsson. *Mol.Phys.* (1989) **68**, 563.
- 41 . W.L.Jorgensen, J.Chandrasekhar, J.D.Madura, R.W.Impey and M.L.Klein. *J. Chem.Phys.* (1983) **79**, 926-935.
- 42 . V.Carraveta and E.Clementi. *J.Chem.Phys.* (1984) **81**, 2646-2651.
- 43 . D.G.Bounds *Chem.Phys.Lett.* (1983) **96**, 604.
- 44 . J.Anderson, J.J.Ullo and S.Yip. *J.Chem.Phys.* (1987) **87**, 1726-1732.
- 45 . W.E.Thiessen and A.H.Narten. *J.Chem.Phys.* (1982) **77**, 2656-2662.
- 46 . Z.Slanina *Chem.Phys.Lett.* (1991) **179**, 355-360.
- 47 . A.Wallqvist and G.Karlström. *Chem.Scripts* (1989) **29A**, 131-137.
- 48 . M.Sprík and M.L.Klein. *J.Chem.Phys.* (1988) **89**, 7556-7560.

- 49 . E.Clementi and H.Popkie. *J.Chem.Phys.* (1972) **57**, 1077.
- 50 . K.Szalewicz, S.J.Cole, W.Kolos and R.J.Bartlett. *J.Chem.Phys.* (1988) **89**, 3662-3673.
- 51 . L.Gil-Adalid and I.Ortega-Blake. *J.Chem.Phys.* (1991) **94**, 3748-3752.
- 52 . W.P.Kraemer, B.O.Ross and P.E.M.Siegbahn. *Chem.Phys.* (1982) **69**, 305.
- 53 . K.Kuchitsu and Y.Morino. *Spectrochim.Acta* (1966) **22**, 33-46.
- 54 . G.D.Carney, L.A.Curtiss and S.R.Langhoff. *J.Mol.Spectry.* (1976) **61**, 371-381.
- 55 . A.R.Hoy, I.M.Mills and G.Strey. *Mol.Phys.* (1972) **24**, 1265-1290.
- 56 . D.Eisenberg and W.Kauzmann. *The Structure and Properties of Water.* (1969) Oxford University Press, London.
- 57 . D.F.Smith Jr. and J.Overend. *Spectrochim.Acta A* (1972) **28**, 471.
- 58 . T.Schlick *J.Comp.Chem.* (1989) **10**, 951.
- 59 . S.-B.Zhu, S.Singh and G.W.Robinson. *J.Chem.Phys.* (1991) **95**, 2791-2799.
- 60 . Z.Slanina *Chem.Phys.* (1991) **150**, 321-329.
- 61 . G.Jancsó, P.Bopp and K.Heinzinger. *Chem.Phys.* (1984) **85**, 377-387.
- 62 . T.R.Dyke, K.M.Mack and J.S.Muenter. *J.Chem.Phys.* (1977) **66**, 498-510.
- 63 . R.H.Page, J.G.Frey, Y.-R.Shen and Y.T.Lee. *Chem.Phys.Lett.* (1984) **106**, 373-376.
- 64 . L.A.Curtiss, D.J.Frurip and M.Blander. *J.Chem.Phys.* (1979) **71**, 2703-2711.
- 65 . D.F.Coker and R.O.Watts. *J.Phys.Chem.* (1987) **91**, 2513-2518.
- 66 . Z.Slanina *J.Chem.Phys.* (1980) **73**, 2519-2521.
- 67 . Z.Slanina *Chem.Phys.Lett.* (1986) **127**, 67-72.
- 68 . P.Cieplak, P.Kollman and T.Lybrand. *J.Chem.Phys.* (1990) **92**, 6755-6760.
- 69 . P.Barnes, J.L.Finney, J.D.Nicholas and J.E.Quinn. *Nature* (1979) **282**, 459-464.
- 70 . K.Refson, G.C.Lie and E.Clementi. *J.Chem.Phys.* (1987) **87**, 3634-3638.
- 71 . R.J.Vos, R.Hendricks and F.B.van Duijneveldt. *J.Comp.Chem.* (1990) **11**, 1-18.
- 72 . M.D.Newton and N.R.Kestner. *Chem.Phys.Lett.* (1983) **94**, 198-201.
- 73 . S.Rybak, B.Jeziorski and K.Szalewicz. *J.Chem.Phys.* (1991) **95**, 6576-6601.
- 74 . E.Honegger and S.Leutwyler. *J.Chem.Phys.* (1988) **88**, 2582-2595.

75 . C.J.Marsden, B.J.Smith, J.A.Pople. J.E.Schaefer III and L.Radom. *J.Chem.Phys.* (1991) **95**, 1825-1828.

Capítulo 4

- 1 . E.Clementi *Lecture Notes in Chemistry # 19: Computational Aspects for Large Chemical Systems* (1980) Springer-Verlag, Heidelberg.
- 2 . E.Clementi *Supercomputers and Supercomputing in Chemical Research and Development in Theoretical Biochemistry & Molecular Biophysics Volume 1: DNA* edited by D.L.Beveridge and R.Lavery (1991) Adenine Press, New York, U.S.A.
- 3 . R.J.Bartlett, I.Shavitt and G.D.Purvis III. *J.Chem.Phys.* (1979) **71**, 281-291.
- 4 . H.Saint-Martin, C.Medina-LLanos and I.Ortega-Blake. *J.Chem.Phys.* (1990) **93**, 6448-6452.
- 5 . E.Whalley *J.Glaciology* (1978) **21**, 13.
- 6 . W.F.Lau and B.M.Pettitt. *J.Med.Chem.* (1989) **32**, 2542-2547.
- 7 . T.Ç ađın, M.Holder and B.M.Pettitt. *J.Compu.Chem.* (1991) **12**, 627-634.

APÉNDICE A

BASES UTILIZADAS PARA CÁLCULOS *AB INITIO*

base QZ2P:

función	exponente	coeficiente
1	0.191016	0.235598
1	0.508922	0.576945
2	1.291628	0.327075
2	4.113478	-0.050941
3	7.340125	-0.118923
3	15.816495	0.023571
4	51.237577	0.002478
4	112.492305	0.001170
5	349.948212	0.000553
5	2164.792662	0.000091
6,7,8	0.128892	0.130525
6,7,8	0.337597	0.344688
9,10,11	0.850223	0.373684
9,10,11	2.102525	0.255876
12,13,14	5.133064	0.113754
12,13,14	14.621809	0.037002
15,16,17	46.533367	0.006708
18,19,20	200.000000	0.000590
21,22,23,24,25	0.622222	1.000000
26,27,28,29,30	0.060000	1.000000

función	exponente	coeficiente
1	0.089859	1.000000
2	0.258165	1.000000
3	0.798266	1.000000
4	2.825680	1.000000
5	12.418500	1.000000
6	82.702800	1.000000
7,8,9	0.200000	1.000000
10,11,12	0.020000	1.000000

base QZP:

oxígeno: 8S,7P → 4S,4P + 2D		
función	exponente	coeficiente
1	0.219032	-0.325770
2	0.623318	-0.604971
2	1.553162	-0.208183
3	5.935557	0.155509
3	34.096860	-0.120210
4	35.537520	0.111077
4	86.272893	-0.003970
4	631.351031	-0.000420
5,6,7	0.154302	0.191650
8,9,10	0.428755	0.393162
8,9,10	1.128480	0.369012
11,12,13	2.972370	0.211348
11,12,13	8.320770	0.076620
14,15,16	26.876700	0.017721
14,15,16	114.863000	0.001613
17,18,19,20,21	4.277000	0.168660
17,18,19,20,21	1.341000	0.584800
22,23,24,25,26	0.522000	1.000000

hidrógeno: 6S → 3S + P		
función	exponente	coeficiente
1	0.089859	0.048364
1	0.258165	0.134580
2	0.798266	0.179965
2	2.825680	0.042371
3	12.418500	0.010282
3	82.702800	0.001256
4,5,6	1.798300	0.177050
7,8,9	0.466300	0.885600

base DZP:

oxígeno: 4S,4P → 2S,2P + D		
función	exponente	coeficiente
1	5.779268	-0.139478
1	1.272318	0.283079
1	0.559623	0.534163
2	0.247498	0.302855
3,4,5	12.662193	0.063529
3,4,5	2.971050	0.256392
3,4,5	0.829612	0.505112
6,7,8	0.214450	0.416022
9,10,11,12,13	0.715847	1.000000

hidrógeno: 4S → 2S + P		
función	exponente	coeficiente
1	8.020733	0.084153
1	1.480731	0.335516
1	0.671579	0.546603
2	0.179534	0.291626
3,4,5	0.856698	1.000000

APÉNDICE B:

coordenadas en ua de los tetrámeros relajados y del dímero con geometría de los monómeros optimizada para la base QZP.

coordenadas de los tetrámeros relajados

TETRÁMERO 1

O1	0.112653	3.669542	0.094406
H11	1.755970	2.634186	-0.056088
H12	0.621957	4.900011	1.298568
O2	0.074441	-3.411988	-0.010050
H21	-1.566125	-2.366523	-0.073410
H22	-0.422991	-4.722323	1.113295
O3	-3.449840	0.150178	-0.157109
H31	-2.410856	1.790112	-0.036542
H32	-4.736725	0.670096	-1.297414
O4	3.618717	0.119405	-0.222040
H41	2.583928	-1.525416	-0.109692
H42	4.923155	-0.401912	-1.341733

TETRÁMERO 2

O1	0.000000	3.670000	0.000000
H11	1.743700	2.721980	-0.071090
H12	0.558050	4.797510	1.322580
O2	-3.670000	0.000000	0.000000
H21	-2.664500	1.691690	0.107310
H22	-4.854380	0.583410	1.235450
O3	0.000000	-3.670000	0.000000
H31	-1.741060	-2.719010	-0.148710
H32	-0.617760	-4.757910	1.288520

O4	3.670000	0.000000	0.000000
H41	2.623990	-1.685850	0.040220
H42	4.769480	-0.554190	1.320310

coordenadas del dímero con la geometría de los monómeros optimizada
con la base QZP

O1	0.000000	0.000000	0.000000
H11	1.789444	0.000000	0.000000
H12	-0.489542	1.721179	0.000000
O2	5.670000	0.000000	0.000000
H21	6.603975	-0.539204	1.427956
H22	6.603975	-0.539204	-1.427956

APÉNDICE C

PROGRAMA DE DINÁMICA MOLECULAR

```

program dnm
C version de marzo de 1991
C un solo cutoff para todas las interacciones intermoleculares,
C con el mismo numero de particiones en la tabulacion y los
C mismos limites inferior y superior; un juego diferente para la
C parte intramolecular, que no tiene cutoff
C los vecinos se deciden por molecula
C dinamica molecular, por ahora, para 1030 atomos de hasta 3 tipos
C y hasta dos tipos de moleculas diferentes
C supone que TODOS los potenciales y fuerzas estan tabuladas
C salvo el potencial intratomico
C algoritmo de Verlet
C UNIDADES: distancias y cargas en uas; energias en Kcal/mol;
C tiempo en segundos; masa en umas

```

```

implicit none
include 'dmpars'
integer numpas, numatom, numtip, numpot, numpart, nesc, itmax
integer i, j, k, n, h, m, l, ntipmol, nmol, conti, contj, numint
integer tipo(maxatom), num(maxatom), ini(maxmol), fin(maxmol)
integer internuc(maxtip,maxtip), intercamb(maxtip,maxtip)
integer intranuc(maxtip,maxtip), intracar(maxtip,maxtip)
integer cont(maxtip), nummol(maxtmol), natmol(maxtmol)
integer tipom(maxtmol,maxat), nlim, splim(maxlim)
integer fdr(maxin,maxpin)
real*8 rmin, rmax, rminc, rmaxc, rminn, rmaxm, cutoff, cf2
real*8 deltat, sig, vo, err, to, mat, vt, ec, t, ep, et
real*8 qn(maxatom,3), qc(maxatom,3), qvn(maxatom,3), qvc(maxatom,3)
real*8 qan(maxatom,3), qac(maxatom,3)
real*8 qnt(maxatom,3), qct(maxatom,3), qi(maxatom,3)
real*8 qnm(maxtmol,maxat,3), qem(maxtmol,maxat,3)
real*8 v(maxatom,3), vi(maxatom,3)
real*8 fu(maxpot,maxpart), u(maxpot,maxpart)
real*8 eta, epa, eca, emin, rmean(maxdis)
real*8 mas(maxtip), car(maxtip), kin(maxtip), dosk(maxtip)
real*8 caja(3), caja2(3), mu0(maxmol,3),eref(maxtmol)
real*8 vcm(3), f(maxatom,3), acsom, req(maxtip,maxtip)
real*8 deltar2(maxtip), acvel(maxtip), mu(maxmol,3), ri(maxdis)
real*8 rminf, rmaxf
character*12 enesal, velsal, r2sal, dipsal, dissal, corsal, vsal
character*12 hissal, monentra, cormin, velmin, fdreal
character*12 fue, pot
logical reinicio, escribe, lim(maxatom), rescale

```

```

common/min/intercar, intracar, internuc, intranuc
common/sis/ nmol, ini, fin
common/esc/ k, nesc
common/tab/ tipo, numpart
common/num/ numatom
common/osm/ itmax, sig
common/ntab/ dosk
common/coi/ qn, qc
common/cov/ qvn, qvc
common/box/ qnt, qct
common/boi/ caja, caja2
common/cut/ cf2, rmin, rmax, rminc, rmaxc, rminn, rmaxm
common/epo/ ep, u, kin
common/fue/ f, fu
common/eci/ ec, err, to, rescale
common/vel/ v, mas, vcm, mat

```

```

common/sca/ vo
common/ver/ qan, qac
common/tis/ deltat
common/his/ numint, numtip
common/qan/ qnm, qem, eref, ntipmol, natmol, tipom
common/raf/ req
common/lim/ lim
common/ldr/ fdr, rminf, rmaxf

```

C LECTURA DE DATOS:

```

C numero de pasos de simulacion, numero de atomos, numero de tipos de
C atomos, numero de potenciales diferentes, intervalo de tiempo, cada cuantos
C pasos debe escribir resultados, numero de iteraciones (para osmol),
C el criterio de convergencia (para osmol) y el criterio de error para
C reescalar las velocidades a la temperatura inicial, el numero de
C intervalos para hacer el histograma
  read (5,*) numpas, numatom, numtip, numpot, deltat, nesc,
  &
  itmax, sig, err, numint

C si debe asignar velocidades nuevas o leer un archivo, si debe calcular
C autocorrelaciones, distancias y desplazamientos, y si debe reescalar las
C velocidades en el paso cero
  read (5,*) reinicio, escribe, rescale

C numero de especies atomicas que no se mueven y cuales son
  read (5,*) nlim, (splim(i), i = 1, nlim)

C archivo con la informacion de los monomeros
  read (5,*(a12')) monentra
  open (1, file = monentra, status = 'old')

C distancias de equilibrio para los enlaces
  do i = 1, numtip
    read (1,*) (req(i,j), j = 1, numtip)
  enddo

C numero de especies moleculares distintas
  read (1,*) ntipmol

C numero de molculas de cada especie
  read (1,*) (nummol(i), i = 1, ntipmol)

C numero de atomos de cada especie molecular
  read (1,*) (natmol(i), i = 1, ntipmol)

C coordenadas de los monomeros de referencia
  do i = 1, ntipmol
    do j = 1, natmol(i)
      read (1,*) (qnm(i,j,k), k = 1, 3), (qem(i,j,k), k = 1, 3),
      &
      tipom(i,j)
    enddo
  enddo

close (1)

C los semilados de la caja de la simulacion
  read (5,*) (caja(i), i = 1, 3)
  do i = 1, 3
    caja2(i) = 2*caja(i)
  enddo

```



```

        enddo

C coordenadas del nucleo, coordenadas de la carga, tipo de atomo,
C numero de molecula
C tipo de atomo: H = 1, O = 2, I = 3
    do i = 1, numatom
        read (5,*) (qn(i,j), j= 1,3), (qc(i,j), j= 1,3),
            tipo(i), num(i)
    enddo

C cargas para cada tipo de atomo
    read (5,*) (car(i), i = 1, numtip)

C masas para cada tipo de atomo
    read (5,*) (mas(i), i = 1, numtip)

C velocidad inicial del centro de masa del sistema para distribuir velocidades
C y la temperatura deseada de simulacion
    read (5,*) vo, to

C matrices con el tipo de interaccion que corresponde a una pareja de atomos
C de los tipos ya descritos
    do i = 1, numtip
        read (5,*) (internuc(i,j), j = 1, numtip)
    enddo
    do i = 1, numtip
        read (5,*) (intercar(i,j), j = 1, numtip)
    enddo
    do i = 1, numtip
        read (5,*) (intranuc(i,j), j = 1, numtip)
    enddo
    do i = 1, numtip
        read (5,*) (intracar(i,j), j = 1, numtip)
    enddo

C numero de particiones tabuladas
    read (5,*) numpart

C nombres de los archivos donde estan tabulados fuerzas y potenciales
    read (5, '(2s12)') fue, pot
    open (1, file = fue, status = 'old', form = 'unformatted')
    open (2, file = pot, status = 'old', form = 'unformatted')
    read (1) fu
    read (2) u
    close (1)
    close (2)

C distancias extremas de evaluacion de potenciales y fuerzas
C para las interacciones intermoleculares, intramoleculares coulombianas,
C intramoleculares de los Morses y cutoff
    read (5,*) rmin, rmax, rminc, rmaxc, rminn, rmaxn, cutoff
    cf2 = cutoff * cutoff

C distancias minima y maxima para la fdr
    read (5,*) rminf, rmaxf

C constantes de resorte para el osmol (lo unico que no esta tabulado)
C uno por tipo de atomo

```

```

read (5,*) (kin(i), i = 1, numtip)
do i = 1, numtip
  dosk(i) = kin(i)*2
enddo

```

C nombres de los archivos de salida para energias, autocorrelacion de velocidades, desplazamientos cuadraticos medios, autocorrelacion de momentos dipolares, distancias intramoleculares, coordenadas para pasos intermedios, velocidades para pasos intermedios, histograma de velocidades, configuracion de energia potencial minima, velocidades correspondientes y sumas parciales (o finales) para la fdr

```

C coordenadas finales en standard output (6)
  read (5,'(11a12)') enesal, velsal, r2sal, dipsal, diasal,
  & corsal, vsal, hissal, cormin, velmin, fdrsal
  open (1, file = enesal, status = 'unknown')
  open (2, file = velsal, status = 'unknown')
  open (3, file = r2sal, status = 'unknown')
  open (4, file = dipsal, status = 'unknown')
  open (7, file = diasal, status = 'unknown')
  open (10, file = hissal, status = 'unknown')

```

C genera los vectores ini y fin para saber donde empieza y donde acaba cada molecula

```

  j = 1
  ini(1) = 1
  do i = 2, numatom
    if (num(i).ne.num(i-1)) then
      fin(j) = i - 1
      j = j + 1
      ini(j) = i
    endif
  enddo
  fin(j) = numatom
  nmol = j

```

C GUARDA LAS POSICIONES INICIALES EN LOS VECTORES QVN Y QVC

```

  do i = 1, numatom
    do j = 1, 3
      qvn(i,j) = qn(i,j)
      qi(i,j) = qn(i,j) ! para la autocorrelacion de posiciones
      qvc(i,j) = qc(i,j)
    enddo
  enddo

```

C genera un vector booleano para identificar los atomos con movimiento restringido

```

  do j = 1, numatom
    lim(j) = .false.
  enddo
  do i = 1, nlim
    do j = 1, numatom
      if (tipo(j).eq.splim(i)) lim(j) = .true.
    enddo
  enddo

```

C calcula la energia potencial de los monomeros
call calmon

C ACTUALIZA LAS POSICIONES DE LAS CARGAS MOVILES
call actcar

C CALCULA EL MOMENTO DIPOLAR PARA LA CONFIGURACION INICIAL

```
acmom = 0.0d+00
do i = 1, nmol
  do n = 1, 3
    mu0(i,n) = 0.0d+00
  enddo
  do j = ini(i), fin(i)
    do n = 1, 3
      mu0(i,n) = mu0(i,n) + qc(j,n) * car(tipo(j))
    enddo
  enddo
  acmom = acmom + mu0(i,1)*mu0(i,1) + mu0(i,2)*mu0(i,2)
  + mu0(i,3)*mu0(i,3)
enddo
acmom = acmom/nmol
write (4,20) acmom
20 format (e15.8)
```

C CALCULA LA MASA TOTAL DEL SISTEMA

```
mat = 0.0d+00
do i = 1, numatom
  mat = mat + mas(tipo(i))
enddo
```

C DISTRIBUYE VELOCIDADES

```
if (reinicio) then
  write (1,*) 'comienza con una matriz de velocidades'
  open (9, file = vsal, status = 'unknown')
  do i = 1, numatom
    read (9,*) (v(i,j), j = 1, 3)
    if (lim(i)) then
      do j = 1, 3
        v(i,j) = 0.0d+00
      enddo
    endif
  enddo
  close (9)
else
  write (1,*) 'se asignan velocidades nuevas'
  call divel
endif

k = 0
sta = 0d+00
epa = 0d+00
eca = 0d+00
```

C calcula la energia cinetica para la primera configuracion
call ecin

C calcula la energia potencial para la primera configuracion
call epot

C rescala a kcal/mol
ep = ep / conver

C se resta la energia de los monomeros aislados
do i = 1, ntipmol
ep = ep - nummol(i) * eref(i)
enddo

```

et = ep + (ec * nmol)
write (1,30) et
30 format ('ENERGIA TOTAL : ',f11.4)
write (1,40) ep
40 format ('ENERGIA POTENCIAL: ',f11.4)
write (1,50) ec * nmol
50 format ('ENERGIA CINETICA : ',f11.4)
write (1,*)

write (1,*) 'PROMEDIOS:'
write (1,55) ep/nmol + ec, ep/nmol, ec
55 format ('total:',f11.4,' potencial:',f11.4,' cinetica:',f11.4)
write (1,*)

emin = ep / nmol

C asigna las velocidades ya escaladas a vi
do j = 1, 3
do i = 1, numatom
vi(i,j) = v(i,j)
enddo
enddo

C calcula el histograma de la distribucion de velocidades
call hivel

C calcula <vi.vi> para la autocorrelacion de velocidades
do i = 1, numtip
cont(i) = 0
acvel(i) = 0.0d+00
enddo
do i = 1, numatom
acvel(tipo(i)) = acvel(tipo(i)) + vi(i,1)*vi(i,1) +
e vi(i,2)*vi(i,2) + vi(i,3)*vi(i,3)
cont(tipo(i)) = cont(tipo(i)) + 1
enddo
do i = 1, numtip
acvel(i) = acvel(i) / cont(i)
enddo
write (2,10) (acvel(i), i = 1, numtip)

do i = 1, maxdis
rmean(i) = 0.0d+00
enddo

C inicializa los contadores para la funcion de distribucion radial
do i = 1, maxin
do j = 1, maxpin
fdr(i,j) = 0
enddo
enddo

C COMIENZA LA DINAMICA MOLECULAR
do k = 1, numpas
do n = 1, 3
vcm(n) = 0.0d+00
enddo
do h = 1, nmol
call box(h)

```

```

do j = ini(h), fin(h)
  do n = 1, 3
    f(j,n) = 0.0d+00
  enddo
  if (.not.lim(j)) then
    call fuerza(h,j)
  endif
  call mueve(j)
enddo ! del atomo j
enddo ! de la molecula h
do n = 1, 3
  vcm(n) = vcm(n)/mat
enddo

```

```

C vuelve a poner la velocidad del centro de masa del sistema en cero
do n = 1, 3
  do j = 1, numatom
    if (.not.lim(j)) then
      v(j,n) = v(j,n) - vcm(n)
    endif
  enddo
enddo

```

```

C CALCULO DE ENERGIA CINETICA Y TEMPERATURA
call ecin

```

```

C ACTUALIZA LAS POSICIONES DE LOS NUCLEOS
do j = 1, 3
  do i = 1, numatom
    qan(i,j) = qvn(i,j)
    qvn(i,j) = qn(i,j)
  enddo
enddo

```

```

C ACTUALIZA LAS POSICIONES DE LAS CARGAS MOVILES
call actcar

```

```

C CALCULO DE ENERGIA POTENCIAL
C call calmon
call epot

```

```

C reescala a kcal/mol
ep = ep / conver

```

```

C se resta la energia de los monomeros aislados
do i = 1, ntipmol
  ep = ep - nummol(i) * eref(i)
enddo
ep = ep / nmol

```

```

C si la configuracion tiene la energia mas baja encontrada,
C guarda las posiciones y las velocidades
if (ep.lt.emin) then
  emin = ep
  open (ll, file = cormin, status = 'unknown')
  write (ll,*) emin
  do i = 1, numatom
    write (ll,15) (qn(i,j), j= 1,3), (qc(i,j), j= 1,3),
    * tipo(i), num(i)
  enddo

```

```

rewind (11)
close (11)
open (12, file = velmin, status = 'unknown')
do i = 1, numatom
  write (12,16) (v(i,j), j= 1,3), tipo(i)
enddo
rewind (12)
close (12)
endif

```

```
et = ep + ec
```

```

eta = eta + et
epa = epa + ep
eca = eca + ec

```

C CALCULA LAS DISTANCIAS INTRAMOLECULARES

```

do h = 1, nmol
  l = 0
  do i = ini(h), fin(h)
    do j = i + 1, fin(h)
      l = l + 1
      ri(l) = sqrt((qn(i,1)-qn(j,1))**2 + (qn(i,2)-qn(j,2))**2
        + (qn(i,3)-qn(j,3))**2)
      rmean(l) = rmean(l) + ri(l)
    enddo
  enddo

```

```

C write (7,60) h, (ri(j), j = 1, l)
C 60 format (i5,3f11.8)
enddo

```

```

if (mod(k,nesc).eq.0) then
  write (1,30) et
  write (1,40) ep
  write (1,50) ec
  write (1,*)
  write (1,*) 'PROMEDIOS:'
  write (1,55) eta/k, epa/k, eca/k
  write (1,*)

```

C escribe las sumas parciales de fdr

```

open (13, file = fdrsal, status = 'unknown', form = 'unformatted')
write (13) fdr
rewind (13)
close (13)

```

C GUARDA LAS COORDENADAS

```

open (8, file = corsal, status = 'unknown')
do i = 1, numatom
  write (8,15) (qn(i,j), j= 1,3), (qc(i,j), j= 1,3),
    @ tipo(i), num(i)
15 format(6f10.5,i3,i5)
enddo
rewind (8)
close (8)

```

C GUARDA LAS VELOCIDADES PARA REINICIO

```

open (9, file = vsal, status = 'unknown')
do i = 1, numatom
  write (9,16) (v(i,j), j= 1,3), tipo(i)

```

```

16.   format(3e16.8,12)
      enddo
      rewind (9)
      close (9)

      if (escribe) then
C calcula el histograma de la distribucion de velocidades
      call hiasvel

C CALCULO DE LOS DESPLAZAMIENTOS CUADRATICOS MEDIOS PARA CADA ESPECIE ATOMICA
      do i = 1, numtip
        cont(i) = 0
        deltar2(i) = 0.0d+00
      enddo
      do i = 1, numatom
        deltar2(tipo(i)) = deltar2(tipo(i)) + (qi(i,1)-qn(i,1))**2 +
          (qi(i,2)-qn(i,2))**2 + (qi(i,3)-qn(i,3))**2
        cont(tipo(i)) = cont(tipo(i)) + 1
      enddo
      do i = 1, numtip
        deltar2(i) = deltar2(i) / cont(i)
      enddo
      write (3,10) (deltar2(i), i = 1, numtip)
10   format (3e16.8)

C CALCULA LA AUTOCORRELACION DE VELOCIDADES PARA CADA ESPECIE ATOMICA
      do i = 1, numtip
        cont(i) = 0
        acvel(i) = 0.0d+00
      enddo
      do i = 1, numatom
        acvel(tipo(i)) = acvel(tipo(i)) + vi(i,1)*v(i,1) +
          vi(i,2)*v(i,2) + vi(i,3)*v(i,3)
        cont(tipo(i)) = cont(tipo(i)) + 1
      enddo
      do i = 1, numtip
        acvel(i) = acvel(i) / cont(i)
      enddo
      write (2,10) (acvel(i), i = 1, numtip)

C CALCULA EL MOMENTO DIPOLAR Y LA AUTOCORRELACION DE DIPOLOS
      acmom = 0.0d+00
      do i = 1, nmol
        do n = 1, 3
          mu(i,n) = 0.0d+00
        enddo
        do j = ini(i), fin(i)
          do n = 1, 3
            mu(i,n) = mu(i,n) + qc(j,n) * car(tipo(j))
          enddo
        enddo
        acmom=acmom+mu0(i,1)*mu(i,1)+mu0(i,2)*mu(i,2)+mu0(i,3)*mu(i,3)
      enddo
      acmom = acmom/nmol
      write (4,20) acmom
    endif ! de escribe

    endif ! de escribir resultados de funciones accesorias

  enddo ! de k = 1, numpas

```

```
C por ahora, suponiendo un solo tipo molecular
write (7,56) (rmean(j)/(numpas*nmol), j = 1, maxdis)
56 format ('promedios:',3f11.8)
close(1)
close(2)
close(3)
close(4)
close(7)
close (10)

END
```


subroutine actcar
C ACTUALIZA LAS POSICIONES DE LAS CARGAS MOVILES

```
implicit none  
include 'dmpars'  
integer i, j, n, h, nmol  
integer ini(maxmol), fin(maxmol)  
real*8 drand  
real*8 qvc(maxatom,3), qct(maxatom,3)  
real*8 qvn(maxatom,3), qnt(maxatom,3)  
real*8 qan(maxatom,3), qac(maxatom,3)  
real*8 qn(maxatom,3), qc(maxatom,3)
```

```
common/sis/ nmol, ini, fin  
common/cov/ qvn, qvc  
common/coi/ qn, qc  
common/box/ qnt, qct  
common/ver/ qan, qac
```

C escoge al azar la molecula con la que va a empezar

```
h = int(nmol * drand(0)) + 1  
do i = h, nmol  
do n = 1, 3  
do j = ini(i), fin(i)  
qvc(j,n) = qc(j,n)  
enddo  
enddo  
call box(i)  
call osmol(i)
```

C de las coordenadas que devuelve osmol, asignarlas a qvc

```
do j = ini(i), fin(i)  
do n = 1, 3  
qac(j,n) = qvc(j,n)  
qvc(j,n) = qct(j,n)  
enddo  
enddo  
do i = 1, h - 1  
do n = 1, 3  
do j = ini(i), fin(i)  
qvc(j,n) = qc(j,n)  
enddo  
enddo  
call box(i)  
call osmol(i)
```

C de las coordenadas que devuelve osmol, asignarlas a qvc

```
do j = ini(i), fin(i)  
do n = 1, 3  
qac(j,n) = qvc(j,n)  
qvc(j,n) = qct(j,n)  
enddo  
enddo
```

```
return  
END
```

```
subroutine box(i)
C encuentra las imagenes de moleculas en celdas vecinas
C genera las matrices qnt y qct a partir de qvn y qcn
```

```
implicit none
include 'dmpars'
```

```
C variables locales
integer i, j, k, l, m, n
real*8 d
```

```
C variables de los commons
integer nmol, numatom, ini(maxmol), fin(maxmol)
real*8 caja(3), caja2(3), qvn(maxatom,3), qvc(maxatom,3)
real*8 qnt(maxatom,3), qct(maxatom,3)
```

```
common/num/ numatom
common/sis/ nmol, ini, fin
common/cov/ qvn, qvc
common/box/ qnt, qct
common/bol/ caja, caja2
```

```
do n = 1, 3
do l = 1, numatom
qnt(l,n) = qvn(l,n)
qct(l,n) = qvc(l,n)
enddo
do l = ini(i), fin(i)
do j = 1, nmol
do k = ini(j), fin(j)
d = qvn(k,n) - qvn(l,n)
if (d.gt.caja(n)) then
do m = ini(j), fin(j)
qnt(m,n) = qvn(m,n) - caja2(n)
qct(m,n) = qvc(m,n) - caja2(n)
enddo
else if (d.lt.-caja(n)) then
do m = ini(j), fin(j)
qnt(m,n) = qvn(m,n) + caja2(n)
qct(m,n) = qvc(m,n) + caja2(n)
enddo
endif
enddo
enddo
enddo

return
END
```

subroutine calson

C calcula la energia potencial y las fuerzas sobre los monomeros de cada
C especie molecular

```
implicit none
include 'dparms'
integer intercar(maxtip,maxtip), internuc(maxtip,maxtip)
integer intracar(maxtip,maxtip), intranuc(maxtip,maxtip)
integer tipo(maxatom), numpart, ntipmol, natmol(maxtmol)
integer tipom(maxtmol,maxat), pos, i, j, k, l, n, itmax
real*8 dosk(maxtip), kin(maxtip), ep, f(maxatom,3)
real*8 fu(maxpot,maxpart), eref(maxtmol), req(maxtip,maxtip)
real*8 qna(maxtmol,maxat,3), qem(maxtmol,maxat,3)
real*8 rn, rc, deltar, cf2, rmin, rmax, rinc, rmaxc, rminc, rmaxm
real*8 spn, epc, fun(3), fue(3), u(maxpot,maxpart)
real*8 sig
```

C variables internas para osmol (version de Jorge)

```
logical conv
integer in, it, is, js
real*8 ro
real*8 xo(maxatom), x1, x10, x21
real*8 yo(maxatom), y1, y10, y21
real*8 zo(maxatom), z1, z10, z21
real*8 ft, sx, sy, sz, ss
real*8 xu(maxatom), yu(maxatom), zu(maxatom), r, r2

common/min/intercar, intracar, internuc, intranuc
common/tab/ tipo, numpart
common/ntab/ dosk
common/cut/ cf2, rmin, rmax, rinc, rmaxc, rminc, rmaxm
common/epo/ ep, u, kin
common/fue/ f, fu
common/qna/ qna, qem, eref, ntipmol, natmol, tipom
common/ref/ req
common/osm/ itmax, sig
```

```
do i = 1, ntipmol
```

C hacer osmol al monomero

```
conv = .false.
it = 0
do while (.not.conv.and.it.lt.itmax)
conv=.true.
it = it + 1
```

C calcula la posicion de cada carga

```
do l = 1, natmol(i)
is = tipom(i,l)
sx = - dosk(is) * qna(i,l,1) * conver
sy = - dosk(is) * qna(i,l,2) * conver
sz = - dosk(is) * qna(i,l,3) * conver
ss = - dosk(is) * conver
do j = 1, natmol(i)
if (j.ne.1) then
js = tipom(i,j)
r2 = (qem(i,j,1) - qem(i,l,1))**2
+ (qem(i,j,2) - qem(i,l,2))**2
+ (qem(i,j,3) - qem(i,l,3))**2
r = sqrt(r2)
pos = 1 + int(numpart*(r - rminc) / (rmaxc - rminc))
in = intracar(is,js)
```

```

ft = fu(in,pos)
sx = sx + qem(i,j,1) * ft
sy = sy + qem(i,j,2) * ft
sz = sz + qem(i,j,3) * ft
ss = ss + ft
endif
enddo
xu(1) = sx / ss
yu(1) = sy / ss
zu(1) = sz / ss
enddo

```

C actualiza posiciones (usa shanks si es posible)

```

if (it.ge.3) then
do k = 1, natmol(i)
if (abs(xu(k)-qem(i,k,1)).gt.sig) then
conv = .false.
x1 = qem(i,k,1)
x21 = xu(k) - x1
x10 = x1 - xo(k)
if (abs(x10).gt.1e-09) then
ro = x21 / x10
if (abs(ro).lt.0.8d+00 .and. ro.gt.0.0d+00) then
qem(i,k,1) = xo(k) - x10**2 / (x21 - x10)
else
qem(i,k,1) = xu(k)
endif
else
qem(i,k,1) = xu(k)
endif
xo(k) = x1
else
qem(i,k,1) = xu(k)
endif

if (abs(yu(k)-qem(i,k,2)).gt.sig) then
conv = .false.
y1 = qem(i,k,2)
y21 = yu(k) - y1
y10 = y1 - yo(k)
if (abs(y10).gt.1e-09) then
ro = y21 / y10
if (abs(ro).lt.0.8d+00 .and. ro.gt.0.0d+00) then
qem(i,k,2) = yo(k) - y10**2 / (y21 - y10)
else
qem(i,k,2) = yu(k)
endif
else
qem(i,k,2) = yu(k)
endif
yo(k) = y1
else
qem(i,k,2) = yu(k)
endif

if (abs(zu(k)-qem(i,k,3)).gt.sig) then
conv = .false.
z1 = qem(i,k,3)
z21 = zu(k) - z1
z10 = z1 - zo(k)

```

```

if (abs(s10).gt.1e-09) then
  ro = s21 / s10
  if (abs(ro).lt.0.8d+00 .and. ro.gt.0.0d+00) then
    qem(i,k,3) = so(k) - s10**2 / (x21 - s10)
  else
    qem(i,k,3) = su(k)
  endif
  else
    qem(i,k,3) = su(k)
  endif
  zo(k) = s1
  else
    qem(i,k,3) = su(k)
  endif
enddo
else
do k = 1, natmol(i)
  xo(k) = qem(i,k,1)
  yo(k) = qem(i,k,2)
  zo(k) = qem(i,k,3)
  if (abs(xu(k) - qem(i,k,1)).gt.sig) conv = .false.
  if (abs(yu(k) - qem(i,k,2)).gt.sig) conv = .false.
  if (abs(zu(k) - qem(i,k,3)).gt.sig) conv = .false.
  qem(i,k,1) = xu(k)
  qem(i,k,2) = yu(k)
  qem(i,k,3) = zu(k)
enddo
endif
enddo

if (it.ge.itmax) then
  write (6,*) 'fuerza no converge en el monomero'
  STOP
endif

```

C calculo de energias potenciales de referencia

```

epn = 0.0d+00
epc = 0.0d+00
ep = 0.0d+00
do j = 1, natmol(i)
  do k = j + 1, natmol(i)
    rn = sqrt((qnm(i,j,1)-qnm(i,k,1))**2+(qnm(i,j,2)-qnm(i,k,2))**2
    + (qnm(i,j,3)-qnm(i,k,3))**2) - req(tipom(i,j),tipom(i,k))
    rc = sqrt((qem(i,j,1)-qem(i,k,1))**2+(qem(i,j,2)-qem(i,k,2))**2
    + (qem(i,j,3)-qem(i,k,3))**2)
    pos = 1 + int(numpart * ((rn - rminn) / (rmaxn - rminn)))
    epn = epn + u(intranuc(tipom(i,j),tipom(i,k)),pos)
    write (6,*) u(intranuc(tipom(i,j),tipom(i,k)),pos)/conver, j, k
    write (6,*) pos
    pos = 1 + int(numpart * ((rc - rminc) / (rmaxc - rminc)))
    epc = epc + u(intracar(tipom(i,j),tipom(i,k)),pos)
  enddo
  rn = (qnm(i,j,1)-qem(i,j,1))**2 + (qnm(i,j,2)-qem(i,j,2))**2
  + (qnm(i,j,3)-qem(i,j,3))**2
  ep = ep + kin(tipom(i,j)) * rn * conver
enddo
eref(i) = (epn + epc + ep) / conver
write (1,20) i, eref(i)
format ('energia potencial del monomero tipo',i2,' = ',f11.4)
write (1,30) epc / conver

```

```

30  format ('energía potencial de las cargas      = ',f11.4)
write (1,40)  apn / conver
40  format ('energía potencial de los núcleos    = ',f11.4)
write (1,50)  ap / conver
50  format ('energía potencial de los resortes   = ',f11.4)
enddo
write (1,*)

```

C calculo de las fuerzas sobre cada atomo

```

do i = 1, ntipmol
write (1,*) 'molecula numero ', i
do j = 1, natmol(i)
do n = 1, 3
fun(n) = 0.0d+00
fue(n) = 0.0d+00
enddo
do k = 1, natmol(i)
if (j.ne.k) then
rn=sqrt((qnm(i,j,1)-qnm(i,k,1))**2+(qnm(i,j,2)-qnm(i,k,2))**2
+ (qnm(i,j,3)-qnm(i,k,3))**2)
e  deltar = rn - req(tipom(i,j),tipom(i,k))
rc=sqrt((qem(i,j,1)-qem(i,k,1))**2+(qem(i,j,2)-qem(i,k,2))**2
e  + (qem(i,j,3)-qem(i,k,3))**2)
pos = 1 + int(numpart * ((deltar - rminm) / (rmaxm - rminm)))
do n = 1, 3
e  fun(n) = fun(n) + fu(intranuc(tipom(i,j),tipom(i,k)),pos)
* ((qnm(i,j,n) - qnm(i,k,n)) / rn)
enddo
pos = 1 + int(numpart * ((rc - rminc) / (rmaxc - rminc)))
do n = 1, 3
e  fue(n) = fue(n) + fu(intracar(tipom(i,j),tipom(i,k)),pos)
* ((qem(i,j,n) - qem(i,k,n)))
enddo
endif
enddo

write (1,*) 'fuerzas sobre el atomo ',j,' : '
write (1,10) ((fue(n) + fun(n)) / conver, n = 1, 3)
10 format (3f11.4)
enddo
write (1,*)
enddo

return
END

```

```
subroutine divel
C DISTRIBUYE VELOCIDADES A PARTIR DE LA VELOCIDAD INICIAL VO DE ACUERDO A
C UNA DISTRIBUCION DE MAXWELL
```

```
implicit none
include 'dmpars'
C variables de los commons
integer tipo(maxatom), numpart, numatom
real*8 v(maxatom,3), vcm(3), mat, mas(maxtip)
real*8 vo
logical lim(maxatom)
```

```
C variables locales
integer n, i, j
real*8 r1, r2, drand, vaux, pi2, a
```

```
common/tab/ tipo, numpart
common/num/ numatom
common/vel/ v, mas, vcm, mat
common/sca/ vo
common/lim/ lim
```

```
pi2 = 2.0 * pi
a = sqrt(mas(2))
```

```
do n = 1, 3
vcm(n) = 0.0d+00
enddo
do n = 1, 3
do j = 1, numatom
r1 = drand(0)
r2 = drand(0)
vaux = vo*sqrt(-log(r1))*cos(pi2*r2)
if (lim(j)) then
v(j,n) = 0.0d+00
else
if (tipo(j).eq.2) then
v(j,n) = vaux / a
else
v(j,n) = vaux
endif
vcm(n) = vcm(n) + mas(tipo(j))*v(j,n)
endif
enddo
vcm(n) = vcm(n)/mat
enddo
```

```
C hace que la velocidad del centro de masa del sistema sea cero
```

```
do i = 1, 3
do j = 1, numatom
if (.not.lim(j)) then
v(j,i) = v(j,i) - vcm(i)
endif
enddo
enddo
return
END
```

subroutine ecin
C calcula la energia cinetica

```
implicit none  
include 'dapars'  
integer i, j, numatom, tipo(maxatom), numpart, k, nesc, nmol  
integer ini(maxmol), fin(maxmol), ns, nb, n  
real*8 v(maxatom,3), vcm(3)  
real*8 mat, mas(maxtip)  
real*8 vt, ec, t, err, to, lambda, ta  
real*8 f(maxatom,3), fu(maxpot,maxpart)  
logical lis(maxatom), rescale  
save ns, nb, ta
```

```
common/sis/ nmol, ini, fin  
common/esc/ k, nesc  
common/tab/ tipo, numpart  
common/num/ numatom  
common/eci/ ec, err, to, rescale  
common/val/ v, mas, vcm, mat  
common/tam/ t  
common/fue/ f, fu  
common/lis/ lis
```

```
ec = 0.0  
do i = 1, numatom  
vt = v(i,1)**2 + v(i,2)**2 + v(i,3)**2  
ec = ec + mas(tipo(i))*vt/2  
enddo  
  
t = (2*ec)/(3*kb*numatom)  
  
if (k.ne.0) then  
ta = ta + t  
endif
```

C reescala a Kcal/mol

```
ec = ec / (conver * nmol)  
  
if (mod(k,nesc).eq.0) then  
write (1,10) k, t, ns, nb  
10 format ('paso: ',i7,' temperatura: ',f11.4,' sube: ',i6,  
& ' baja: ', i6)  
if (k.gt.0) then  
write (1,15) ta/k  
15 format ('temperatura promedio:',f11.4)  
endif  
endif  
  
if ((abs(t-to)).gt.err.and.rescale) then  
if (k.eq.0) then  
ns = 0  
nb = 0  
else  
if (t.gt.to) then  
nb = nb + 1  
else  
ns = ns + 1  
endif  
endif
```



```

lambda = 1 / sqrt(t / to)
do j = 1, 3
  do i = 1, numatom
    v(i,j) = lambda * v(i,j)
  enddo
enddo
if (k.eq.0) then
  do i = 1, nmol
    call box(i)
    do j = ini(i), fin(i)
      do n = 1, 3
        f(j,n) = 0.0d+00
      enddo
      if (.not. lim(j)) then
        call fuerza(i,j)
      endif
      call regresa(j)
    enddo
  enddo
else
  do j = 1, numatom
    call regresa(j)
  enddo
endif

ec = 0.0
do i = 1, numatom
  vt = v(i,1)**2 + v(i,2)**2 + v(i,3)**2
  ec = ec + mas(tipo(i))*vt/2
enddo

t = (2*ec)/(3*kb*numatom)
C reescala a Kcal/mol
ec = ec / (conver * nmol)

if (mod(k,nesc).eq.0) then
  write(1,20) k, t, ns, nb
20 format ('paso: ',i7,' temperatura escalada: ',f11.4,' sube: ',i6,
  ' baja: ',i6)
endif
endif
return
END

```

```

subroutine spot
C calcula la energia potencial

implicit none
include 'dmpars'
C variables de los commons
integer nmol, ini(maxmol), fin(maxmol), numpart, tipo(maxatom)
integer intercar(maxtip,maxtip), intracar(maxtip,maxtip)
integer internuc(maxtip,maxtip), intranuc(maxtip,maxtip)
integer fdr(maxin,maxpin)
real*8 rmin, rmax, rminc, rmaxc, rmaxa, rminm, cf2
real*8 qnt(maxatom,3), qct(maxatom,3)
real*8 kin(maxtip), u(maxspot,maxpart), sp
real*8 rminf, rmaxf

C variables locales
integer h, j, nvec, m, i, n, l, tpot(maxvec), pos, pldr
real*8 rn, rc, r(maxvec), req(maxtip,maxtip)

common/min/intercar, intracar, internuc, intranuc
common/mis/ nmol, ini, fin
common/tab/ tipo, numpart
common/box/ qnt, qct
common/cut/ cf2, rmin, rmax, rminc, rmaxc, rminm, rmaxm
common/epo/ ep, u, kin
common/ref/ req
common/fdr/ fdr, rminf, rmaxf

C LOS VECINOS SE DECIDEN CON MOLECULAS COMPLETAS
ep = 0.0d+00
do h = 1, nmol
call box(h)
C determinacion de vecinos para la parte intermolecular
do j = ini(h), fin(h)
nvec = 0
do 30 m = h+1, nmol
do i = ini(m), fin(m)
rn = (qnt(i,1)-qnt(j,1))**2 + (qnt(i,2)-qnt(j,2))**2
+ (qnt(i,3)-qnt(j,3))**2
e if (rn.lt.cf2) then
do l = ini(m), fin(m)
rn = sqrt((qnt(l,1)-qnt(j,1))**2 + (qnt(l,2)-qnt(j,2))**2
+ (qnt(l,3)-qnt(j,3))**2)
e nvec = nvec + 1
r(nvec) = rn
tpot(nvec) = internuc(tipo(j),tipo(l))
rc = sqrt((qct(l,1)-qct(j,1))**2 + (qct(l,2)-qct(j,2))**2
+ (qct(l,3)-qct(j,3))**2)
e nvec = nvec + 1
r(nvec) = rc
tpot(nvec) = intercar(tipo(j),tipo(l))
enddo
goto 30
endif
enddo
30 continue

C contribucion intermolecular
do l = 1, nvec
pos = 1 + int(numpart*((r(l) - rmin) / (rmax - rmin)))

```

```

if (pos.lt.1.or.pos.gt.nupart) then
  write (6,*) 'spot, intermolecular', r(1), 1, pos
  STOP
endif
ep = ep + u(tpot(1),pos)
if (tpot(1).le.maxin) then
  pfdr = 1 + int(maxpin*((r(1) - rminf) / (rmaxf - rminf)))
  fdr(tpot(1),pfdr) = fdr(tpot(1),pfdr) + 1
endif
enddo

```

C contribucion intramolecular

```

do m = ini(h), fin(h)
  if (m.gt.j) then
    rn = sqrt((qnt(m,1)-qnt(j,1))**2 + (qnt(m,2)-qnt(j,2))**2
    + (qnt(m,3)-qnt(j,3))**2) - req(tipo(m),tipo(j))
    rc = sqrt((qct(m,1)-qct(j,1))**2 + (qct(m,2)-qct(j,2))**2
    + (qct(m,3)-qct(j,3))**2)
    pos = 1 + int(nupart*((rn - rminn) / (rmaxn - rminn)))
    if (pos.lt.1.or.pos.gt.nupart) then
      write (6,*) 'spot, intramolecular', rn, j, m, pos
      STOP
    endif
    ep = ep + u(intranuc(tipo(j),tipo(m)),pos)
    pos = 1 + int(nupart*((rc - rminc) / (rmaxc - rminc)))
    if (pos.lt.1.or.pos.gt.nupart) then
      write (6,*) 'spot, intramolecular', rc, j, m, pos
      STOP
    endif
    ep = ep + u(intracar(tipo(j),tipo(m)),pos)
  endif
enddo

```

C parte intraatomica, no tabulada

```

rn = (qnt(j,1)-qct(j,1))**2 + (qnt(j,2)-qct(j,2))**2
+ (qnt(j,3)-qct(j,3))**2
ep = ep + kin(tipo(j)) * rn *conver
enddo ! del atomo j
enddo ! de la molecula h

return
END

```

```

subroutine fuerza(h,j)
C calcula la fuerza que sienta cada atomo

implicit none
include 'dmpars'
C variables locales
integer h, j, nvec, m, i, l, tpot(maxvec), n, pos
real*8 rn,rc,r(maxvec),seg(maxvec,3),deltar,req(maxtip,maxtip)

C variables de los commons
integer nmol, numpart, ini(maxmol), fin(maxmol), tipo(maxatom)
integer intercar(maxtip,maxtip), intracar(maxtip,maxtip)
integer internuc(maxtip,maxtip), intranuc(maxtip,maxtip)
real*8 rmin, rmax, rminc, rmaxc, cf2, rminm, rmaxm
real*8 fu(maxpot,maxpart), f(maxatom,3)
real*8 qnt(maxatom,3), qct(maxatom,3)

common/min/intercar, intracar, internuc, intranuc
common/sis/ nmol, ini, fin
common/tab/ tipo, numpart
common/box/ qnt, qct
common/cut/ cf2, rmin, rmax, rminc, rmaxc, rminm, rmaxm
common/fue/ f, fu
common/ref/ req

nvec = 0
do 20 m = 1, nmol
  if (m.ne.h) then
    do l = ini(m), fin(m)
C busca atomos vecinos con los nucleos
      rn = (qnt(l,1)-qnt(j,1))**2 + (qnt(l,2)-qnt(j,2))**2
      + (qnt(l,3)-qnt(j,3))**2
      e
      if (rn.lt.cf2) then
        do l = ini(m), fin(m)
          rn = sqrt((qnt(l,1)-qnt(j,1))**2 + (qnt(l,2)-qnt(j,2))**2
          + (qnt(l,3)-qnt(j,3))**2)
          e
          nvec = nvec + 1
          r(nvec) = rn
          seg(nvec,1) = qnt(j,1) - qnt(l,1)
          seg(nvec,2) = qnt(j,2) - qnt(l,2)
          seg(nvec,3) = qnt(j,3) - qnt(l,3)
          tpot(nvec) = internuc(tipo(j),tipo(l))
          rc = sqrt((qct(l,1)-qct(j,1))**2 + (qct(l,2)-qct(j,2))**2
          + (qct(l,3)-qct(j,3))**2)
          e
          nvec = nvec + 1
          r(nvec) = rc
          seg(nvec,1) = qct(j,1) - qct(l,1)
          seg(nvec,2) = qct(j,2) - qct(l,2)
          seg(nvec,3) = qct(j,3) - qct(l,3)
          tpot(nvec) = intercar(tipo(j),tipo(l))
        enddo
      goto 20
    endif
  enddo
endif
20 continua

C suma de las fuerzas intermoleculares
do n = 1, 3
  do l = 1, nvec

```

```

pos = 1 + int(nupart*((r(l) - rmin) / (rmax - rmin)))
if (pos.lt.1.or.pos.gt.nupart) then
  write (6,*) 'fuerza, intermolecular', h, pos, j, l
  STOP
endif
f(j,n) = f(j,n) + fu(tpot(l),pos) * seq(l,n)
enddo
enddo

```

C suma de las fuerzas intramoleculares

```

do m = ini(h), fin(h)
  if (m.ne.j) then
    rn = sqrt((qnt(m,1)-qnt(j,1))**2 + (qnt(m,2)-qnt(j,2))**2
    + (qnt(m,3)-qnt(j,3))**2)
    *
    deltar = rn - req(tipo(m),tipo(j))
    rc = sqrt((qct(m,1)-qct(j,1))**2 + (qct(m,2)-qct(j,2))**2
    *
    + (qct(m,3)-qct(j,3))**2)
    pos = 1 + int(nupart * ((deltar - rmin) / (rmax - rmin)))
    if (pos.lt.1.or.pos.gt.nupart) then
      write (6,*) 'fuerza, intranuc', pos, h, tipo(j), tipo(m), rn, deltar
      STOP
    endif
    do n = 1, 3
      f(j,n) = f(j,n) + fu(intranuc(tipo(j),tipo(m)),pos) *
      (qnt(j,n)-qnt(m,n))/rn
    *
    enddo
    pos = 1 + int(nupart * ((rc - rmin) / (rmax - rmin)))
    if (pos.lt.1.or.pos.gt.nupart) then
      write (6,*) 'fuerza, intracar', pos, h, tipo(j), tipo(m), rc
      STOP
    endif
    do n = 1, 3
      f(j,n) = f(j,n) + fu(intracar(tipo(j),tipo(m)),pos) *
      (qct(j,n)-qct(m,n))
    *
    enddo
  endif
enddo

return
END

```

subrutina hisvel
C genera un histograma a partir de la distribucion de velocidades
C y calcula energias cineticas para cada especie atomica

```
implicit none
include 'dapara'

integer i, j, k, n, numint, numatom, numtip, numpart
integer tipo(maxatom), nvel(maxtip,maxint), cont(maxtip)
real*8 vini, deltav, vmag(maxatom), vmax(maxtip), vmin(maxtip)
real*8 vm(maxtip,maxatom), vinin, mas(maxtip), t, ec
real*8 v(maxatom,3), vcm(3), mat
```

```
common/tab/ tipo, numpart
common/num/ numatom
common/his/ numint, numtip
common/vel/ v, mas, vcm, mat
common/tem/ t
```

C calcula las magnitudes de las velocidades

```
do i = 1, numatom
  vmag(i) = sqrt(v(i,1)*v(i,1)+v(i,2)*v(i,2)+v(i,3)*v(i,3))
enddo
```

C separa por tipo atomico

```
do j = 1, numtip
  k = 0
  do i = 1, numatom
    if (tipo(i).eq.j) then
      k = k + 1
      vm(j,k) = vmag(i)
    endif
  enddo
  cont(j) = k
enddo
```

C calcula energias cineticas para cada especie

```
do j = 1, numtip
  ec = 0.0d+00
  do i = 1, cont(j)
    ec = ec + mas(j) * vm(j,i) * vm(j,i) / 2
  enddo
  ec = ec / (conver * cont(j))
  write (10,*) 'energia cinetica de la especie ',j,' = ',ec
enddo
```

C encuentra maximos y minimos

```
do j = 1, numtip
  vmax(j) = vm(j,1)
  vmin(j) = vm(j,1)
  do i = 2, cont(j)
    if (vm(j,i).gt.vmax(j)) then
      vmax(j) = vm(j,i)
    else if (vm(j,i).lt.vmin(j)) then
      vmin(j) = vm(j,i)
    endif
  enddo
enddo
```

C cuenta

```

do j = 1, numtip
  deltav = (vmax(j) - vmin(j)) / numint
  vini = vmin(j)
  vinin = vini + deltav
  do k = 1, numint
    n = 0
    do i = 1, cont(j)
      if (vm(j,i).le.vinin.and.vm(j,i).ge.vini) then
        n = n + 1
      endif
    enddo
    nvel(j,k) = n
    vini = vini + deltav
    vinin = vini + deltav
  enddo
enddo

```

C escribe

```

do j = 1, numtip
  write (10,*) 'tipo atomico: ', j
  write (10,*) 'vmax = ', vmax(j), ' vmin = ', vmin(j)
  do i = 1, numint
    write (10,10) i, nvel(j,i)
  10   format (2i5)
  enddo
  write (10,*)
enddo

return
END

```

subroutine nueva(j)
C calcula las posiciones y velocidades nuevas de las partículas
C algoritmo de Verlet

```
implicit none
include 'dmpars'
integer n, j, numpart, tipo(maxatom)
real*8 v(maxatom,3)
real*8 qvn(maxatom,3),qvc(maxatom,3),qn(maxatom,3),qc(maxatom,3)
real*8 deltat, mat, f(maxatom,3), vcm(3), mas(maxtip)
real*8 qan(maxatom,3),qac(maxatom,3), fu(maxpot,maxpart)

common/tab/ tipo, numpart
common/col/ qn, qc
common/cov/ qvn, qvc
common/fue/ f, fu
common/vel/ v, mas, vcm, mat
common/tim/ deltat
common/ver/ qan, qac

do n = 1, 3
  qn(j,n) = 2*qvn(j,n)-qan(j,n)+ f(j,n)*deltat*deltat/mas(tipo(j))
  qc(j,n) = 2*qvc(j,n)-qac(j,n)+ f(j,n)*deltat*deltat/mas(tipo(j))
  v(j,n) = (qn(j,n) - qan(j,n))/(2*deltat)
  vcm(n) = vcm(n) + mas(tipo(j))*v(j,n)
enddo

return
END
```


C subrutina OSMOL de Jorge Hernandez, adaptada
C tabulado

```
subroutine osmol(ma)
```

```
implicit none  
include 'dmpars'  
integer ma
```

C variables externas, de las rutinas de Jorge y de datos generales

```
integer nmol, numpart, itmax  
integer ini(maxmol), fin(maxmol), sp(maxatom)  
integer intranuc(maxtip,maxtip), intracar(maxtip,maxtip)  
integer intercar(maxtip,maxtip), internuc(maxtip,maxtip)  
real*8 xnt(maxatom), ynt(maxatom), znt(maxatom)  
real*8 xct(maxatom), yct(maxatom), zct(maxatom)  
real*8 rmin, rmax, rminc, rmaxc, cf2  
real*8 dosk(maxtip), sig  
real*8 f(maxatom,3), fu(maxpot,maxpart)
```

```
common/min/intercar, intracar, internuc, intranuc  
common/sis/ nmol, ini, fin  
common/tab/ sp, numpart  
common/osm/ itmax, sig  
common/ntab/ dosk  
common/box/ xnt, ynt, znt, xct, yct, zct  
common/cut/ cf2, rmin, rmax, rminc, rmaxc  
common/fue/ f, fu
```

C variables internas

```
logical conv  
integer i, j, k, ia, in, it, is, js, na, nv, pos, iv(maxatom)  
real*8 ro  
real*8 xo(maxatom), x1, x10, x21  
real*8 yo(maxatom), y1, y10, y21  
real*8 zo(maxatom), z1, z10, z21  
real*8 ft, sx, sy, sz, ss  
real*8 xu(maxatom), yu(maxatom), zu(maxatom), r, r2
```

```
ia = ini(ma)  
na = fin(ma)
```

C calcula las distancias y define vecinos (sobre todos los atomos)

```
nv = 0  
do 10 i = 1, nmol  
  if (i.eq.ma) goto 10  
  do j = ini(i), fin(i)  
    do k = ia, na  
      r = (xct(j) - xct(k))**2  
e      + (yct(j) - yct(k))**2  
e      + (zct(j) - zct(k))**2  
      if (r.lt.cf2) then  
        nv = nv + 1  
        iv(nv) = i  
        goto 10  
      endif  
    enddo  
  enddo  
10 continue  
conv = .false.
```

```

it = 0
do while (.not.conv.and.it.lt.itmax)
conv=.true.
it = it + 1
C calcula la posicion de cada carga
do i = ia, na
is = sp(i)
sx = 0.0d+00
sy = 0.0d+00
sz = 0.0d+00
ss = 0.0d+00
do j = 1, nv
do k = ini(iv(j)), fin(iv(j))
js = sp(k)
r2 = (xct(i) - xct(k))**2
+ (yct(i) - yct(k))**2
+ (zct(i) - zct(k))**2
r = sqrt(r2)
pos = 1 + int(numpart*((r - rmin) / (rmax - rmin)))
if (pos.lt.1.or.pos.gt.numpart) then
write (6,*) 'osmol, inter', i, k, sp(i), sp(k), r, it
write (6,*) 'osmol, coord', xct(i), ynt(i), znt(i)
write (6,*) 'osmol, coord', xct(k), ynt(k), znt(k)
write (6,*) 'osmol, coord', xct(i), yct(i), zct(i)
write (6,*) 'osmol, coord', xct(k), yct(k), zct(k)
STOP
endif
in = intercar(is,js)
ft = fu(in,pos)
sx = sx + xct(k) * ft
sy = sy + yct(k) * ft
sz = sz + zct(k) * ft
ss = ss + ft
enddo
enddo
sx = sx - dosk(is) * xnt(i) * conver
sy = sy - dosk(is) * ynt(i) * conver
sz = sz - dosk(is) * znt(i) * conver
ss = ss - dosk(is) * conver
do j = ia, na
if (j.ne.i) then
js = sp(j)
r2 = (xct(j) - xct(i))**2
+ (yct(j) - yct(i))**2
+ (zct(j) - zct(i))**2
r = sqrt(r2)
pos = 1 + int(numpart*((r - rminc) / (rmaxc - rminc)))
if (pos.lt.1.or.pos.gt.numpart) then
write (6,*) 'osmol, intra', i, j, it, r, pos
STOP
endif
in = intracar(is,js)
ft = fu(in,pos)
sx = sx + xct(j) * ft
sy = sy + yct(j) * ft
sz = sz + zct(j) * ft
ss = ss + ft
endif
enddo
xu(i) = sx / ss

```

```
yu(i) = sy / sa
su(i) = sz / sa
enddo
```

C actualiza posiciones (usa shanks si es posible)

```
if (it.ge.3) then
do k = ia, na
if (abs(xu(k)-xct(k)).gt.sig) then
conv = .false.
x1 = xct(k)
x21 = xu(k) - x1
x10 = x1 - xo(k)
if (abs(x10).gt.1e-09) then
ro = x21 / x10
if (abs(ro).lt.0.8d+00 .and. ro.gt.0.0d+00) then
xct(k) = xo(k) - x10**2 / (x21 - x10)
else
xct(k) = xu(k)
endif
else
xct(k) = xu(k)
endif
xo(k) = x1
else
xct(k) = xu(k)
endif

if (abs(yu(k)-yct(k)).gt.sig) then
conv = .false.
y1 = yct(k)
y21 = yu(k) - y1
y10 = y1 - yo(k)
if (abs(y10).gt.1e-09) then
ro = y21 / y10
if (abs(ro).lt.0.8d+00 .and. ro.gt.0.0d+00) then
yct(k) = yo(k) - y10**2 / (y21 - y10)
else
yct(k) = yu(k)
endif
else
yct(k) = yu(k)
endif
yo(k) = y1
else
yct(k) = yu(k)
endif

if (abs(zu(k)-zct(k)).gt.sig) then
conv = .false.
z1 = zct(k)
z21 = zu(k) - z1
z10 = z1 - zo(k)
if (abs(z10).gt.1e-09) then
ro = z21 / z10
if (abs(ro).lt.0.8d+00 .and. ro.gt.0.0d+00) then
zct(k) = zo(k) - z10**2 / (z21 - z10)
else
zct(k) = zu(k)
endif
else
```

```

      sct(k) = su(k)
    endif
    so(k) = s1
  else
    sct(k) = su(k)
  endif
enddo
else
  do k = ia, na
    xo(k) = xct(k)
    yo(k) = yct(k)
    zo(k) = zct(k)
    if (abs(xu(k) - xct(k)).gt.sig) conv = .false.
    if (abs(yu(k) - yct(k)).gt.sig) conv = .false.
    if (abs(su(k) - sct(k)).gt.sig) conv = .false.
    xct(k) = xu(k)
    yct(k) = yu(k)
    sct(k) = su(k)
  enddo
endif
enddo
C write (6,*) 'numero de iteraciones: ',it,' para la molecula:',ma
C write (6,*) 'fuerzas finales:'
C do k = ia , na
C   write (6,*) k, xu(k), yu(k), zu(k)
C enddo
C write (6,*)
C if (it.ge.itmax) then
C   write (6,*) 'fuerza no converge en :', ma
C   STOP
C endif

return
END

```

subroutine regresa(j)
C calcula las posiciones anteriores de las partículas

```
implicit none
include 'dmpars'
integer n, j, numpart, tipo(maxatom)
real*8 v(maxatom,3), fu(maxpot,maxpart)
real*8 qvn(maxatom,3),qvc(maxatom,3)
real*8 deltat, mat, f(maxatom,3), vcm(3), mas(maxtip)
real*8 qan(maxatom,3),qac(maxatom,3)

common/tab/ tipo, numpart
common/cov/ qvn, qvc
common/fue/ f, fu
common/vel/ v, mas, vcm, mat
common/tim/ deltat
common/ver/ qan, qac

do n = 1, 3
  qan(j,n) = qvn(j,n) - v(j,n)*deltat + f(j,n)*deltat*deltat/  
             (2*mas(tipo(j)))
  qac(j,n) = qvc(j,n) - v(j,n)*deltat + f(j,n)*deltat*deltat/  
             (2*mas(tipo(j)))
enddo

return
END
```