



03081  
10  
2ej

**UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO**

Unidad Académica de los Ciclos  
Profesional y de Posgrado

**COLEGIO DE CIENCIAS Y HUMANIDADES**

**CALCULOS MOLECULARES Y SIMULACIONES  
NUMERICAS DE LA REACCION DE  
HIDROLISIS DEL PIROFOSFATO**

**T E S I S**

Que para obtener el Grado de:

**DOCTOR EN INVESTIGACION BIOMEDICA BASICA**

Presenta el

**MIBB HUMBERTO | SAINT MARTIN POSADA**

Cuernavaca, Mor.

TESIS CON  
FALLA DE ORIGEN

1992.

**TESIS CON  
FALLA DE ORIGEN**



Universidad Nacional  
Autónoma de México



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

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

Todo el material contenido en esta tesis 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.

## PREFACIO

---

Comúnmente los reportes científicos se escriben en forma impersonal, lo cual origina una serie de construcciones gramaticales que resultan muy extrañas y hasta chocantes. En cambio, es privilegio del autor de una tesis el dedicar algunas páginas para la escritura de agradecimientos, dedicatorias y una que otra frase que haya considerado pertinente. De esta forma, cada tesista da un sello personal a su reporte escrito.

Yo he preferido usar este espacio para hacer un prefacio en primera persona, de manera que pueda reunir los agradecimientos, las dedicatorias y las frases ridículas en una misma sección. Así pues, comenzaré por los agradecimientos, que ya las frases ridículas aparecerán sin mucho esfuerzo.

En primer lugar, mi gratitud hacia mis padres, Emilio Saint Martín y Blanca Posada, sin cuyo apoyo moral, logístico y financiero, *nada* hubiera sido posible.

A los cuatro miembros de mi Comité Tutorial, Dr. Armando Gómez Puyou, Dr. Georges Dreyfus, Dr. Francisco Lara y Dr. Iván Ortega. Cada examen tutorial fue muy formativo, ya que pude observar cómo investigadores establecidos en líneas bien definidas, mostraban una gran flexibilidad mental para adaptarse a conocimientos que en un principio les eran extraños, y un espíritu crítico, que enriqueció mi concepción de la Ciencia. Escribí la tesis pensando en las respuestas a sus posibles preguntas, particularmente las formuladas por el Dr. Lara.

El Dr. Ortega participó como mi asesor directo, y supo guiarme hacia las soluciones de los problemas que fueron presentándose durante el desarrollo del proyecto. Por sobre todo, derrochó ingenio para colocarme en situaciones que templaran mi carácter.

El trabajo de programación "en la trinchera" lo compartí con Jorge Hernández, cuyas aportaciones fueron innumerables y siempre oportunas. También me siento afortunado por haber contado con la gran capacidad de Nina Pastor para el análisis de la información. Las discusiones con ambos, JH y NP, me permitieron comprender las teorías y los experimentos relacionados con este proyecto.

El Dr. Ludwik Adamowicz y el Dr. Andrzej Leś tuvieron una participación muy importante en la obtención y el procesamiento de los datos de los cálculos *ab initio*.

Quiero expresar mi reconocimiento a los investigadores que aceptaron formar el jurado para mi examen doctoral: Dra. Marieta Tuena de Gómez Puyou, Dr. Octavio Novaro, Dr. Heliodoro Celis, Dr. Mariano López de Haro, Dr. Luis Mier y Terán y Dr. Eduardo Bárzana. Sus observaciones contribuyeron a mejorar mi entendimiento de los diversos temas relacionados con esta tesis.

Agradezco a Elena Dreser por revisar el escrito. Al corregirlo para que fuera inteligible en español, me hizo notar la importancia que tiene una buena redacción para comunicar el conocimiento científico.

El trabajo fue efectuado en el Laboratorio de Biofísica de la Subdependencia de Cuernavaca, del IFUNAM. Una parte de los cálculos fueron hechos en el Departamento de Química de la Universidad de Arizona.

El apoyo financiero fue proporcionado por el CONACyT, con la beca 47868, y por PADEP, con cuya aportación pude hacer las estancias necesarias en Tucson.

Ahora van las dedicatorias:

A mi novia, Anahí.

A mis hermanos, Emilio, Silvia, Gabriel y Blanca.

A mis amigos del Laboratorio de Biofísica, Mabel, María Luisa, Juan Manuel, Felipe y Raúl. También a mis alumnos de Física Moderna del CAD: Mauricio, Edgar, 2 \* Daniel, Iker, Margareta, Alexis, Pedro, Rafael, Ruy y Gabriela.

Mayo de 1992

HSMP

## RESUMEN

---

El objetivo principal de esta tesis es el de contribuir al desarrollo de un método para la descripción de fenómenos bioquímicos, a partir de principios físicos fundamentales. El grado de refinamiento alcanzado hasta ahora era reducido, dada la inexactitud en varias aproximaciones de los modelos usados. Aquí se emplea el esquema que parte de cálculos *ab initio*, con los que se diseñan potenciales de interacción intermolecular para efectuar simulaciones numéricas de los sistemas de interés. El aporte de esta tesis, representa un avance cualitativo y cuantitativo en el potencial de interacción, y en los programas de simulación y de procesamiento de resultados.

Se diseñó un potencial analítico, basado en cálculos *ab initio*, para incluir efectos no aditivos de muchos cuerpos: cada átomo se modela con un núcleo y una carga móvil. Esta última responde a las fuerzas producidas por su núcleo y por las demás cargas en el sistema considerado. Con este potencial se ha logrado la mejor reproducción, reportada hasta la fecha, de los datos experimentales sobre la energía intermolecular y la estructura del agua líquida. También se ha aplicado a sistemas que incluyen otras moléculas.

Cada vez que se mueve una molécula, las características del modelo requieren ajustar las posiciones de las cargas a las nuevas interacciones, por lo que el esquema de programación debe permitir el seguimiento de las posiciones de los núcleos y de las cargas, así como de las energías del sistema. De este modo se pueden calcular diversos parámetros energéticos y estructurales, ya bien de todo el sistema o de regiones específicas.

El refinamiento alcanzado, hace posible obtener datos relevantes para comprender la hidrólisis del pirofosfato: con cálculos *ab initio* al nivel  $MBPT(2)/6-31G^{**}$  se logró obtener una estimación confiable de la entalpía de la reacción en fase gaseosa,  $\Delta H_g = -3.8 \pm 1.0 \text{ kcal/mol}$ . Con simulaciones MC, se calcularon las contribuciones a  $\Delta G_{aq}$  que provienen de la energía de hidratación,  $\Delta H_{hid} = -7.3 \text{ kcal/mol}$ , y del término entrópico,  $T\Delta S_{aq} = -0.6 \text{ kcal/mol}$ . Estos resultados son preliminares y están sujetos a perfeccionamiento. Sin embargo, al combinarlos con datos experimentales, proporcionan evidencia suficiente para plantear que tres fenómenos contribuyen en igual medida a  $\Delta G_{aq}$ : los efectos intramoleculares, la energía de hidratación y el efecto entrópico sobre el solvente.

El esquema aquí presentado permite estudiar cada fenómeno por separado, analizar inexactitudes y proponer refinamientos.

## ABSTRACT

---

The main goal in this thesis is to contribute to the development of a method for the description of biochemical phenomena, starting from first principles. The sophistication attained up to date, was constrained, because of the inaccuracies in various approximations of the models in use. The scheme employed here starts from *ab initio* calculations, with which intermolecular interaction potentials are designed, to perform numerical simulations of the systems under study. This thesis provides an advancement, both qualitative and quantitative, to the interaction potential, and to the software for the simulations and the processing of results.

An analytical potential was designed, based on *ab initio* calculations, to include many-body nonadditivity: each atom is modelled by a nucleus and a mobile charge, which responds to the forces produced by its nucleus and by the other charges in the system under consideration. The best reported agreement with experimental data on the intermolecular energy, and on the structure of liquid water, has been obtained from this potential. It has also been applied to systems which include other molecules.

Since the features of the model require the positions of the charges to adjust to the new interactions, each time a molecule is moved, the code was designed to keep track of the positions of both, nuclei and charges, and of the system's energies. Hence, various energetic and structural parameters can be computed, either for the whole system or for selected regions.

The improvement attained made it possible to get results which are relevant to the understanding of the pyrophosphate hydrolysis: from *ab initio* calculations at the *MBPT(2)/6 - 31G\*\** level, a reliable estimate for the gas phase enthalpy of the reaction was obtained,  $\Delta H_g = -3.8 \pm 1.0 \text{ kcal/mol}$ . From MC simulations, the contributions to  $\Delta G_{aq}$  of the hydration energy and of the entropic term were estimated to be  $\Delta H_{hyd} = -7.3 \text{ kcal/mol}$ , and  $T\Delta S_{aq} = -0.6 \text{ kcal/mol}$ , respectively. These results are preliminary and subject to improvement. Nevertheless, when combined with experimental data, they provide enough evidence to propose that three phenomena contribute equally to  $\Delta G_{aq}$ : the intramolecular effects, the hydration energy and the entropic effect on the solvent.

The scheme here presented allows for a separate study of each contribution, for the analysis of inaccuracies, and for the proposal of refinements.

## INDICE DE CONTENIDOS

<b>Introducción</b>	<b>1</b>
<b>1 Catálisis enzimática de los compuestos fosforilados de <i>alta energía</i></b>	<b>4</b>
1.1 Los efectos de los solventes en la bioquímica .....	5
1.2 Catálisis enzimática en distintos solventes .....	5
1.3 Catálisis enzimática de la reacción de hidrólisis del pirofosfato .....	7
1.4 ¿De dónde proviene la <i>alta energía</i> de los anhídridos de fosfato? .....	8
1.5 La reacción de hidrólisis del pirofosfato en diversos solventes .....	9
1.6 Referencias del capítulo 1 .....	12
<b>2 Aplicación de los cálculos cuánticos al estudio de las reacciones químicas</b>	<b>15</b>
2.1 Los cálculos <i>ab initio</i> .....	16
2.2 Clasificación de bases para cálculos <i>ab initio</i> .....	20
2.3 Aplicación a la determinación de estructuras moleculares .....	21
2.4 Aplicación a la descripción de las reacciones químicas .....	22
2.5 Clasificación de las reacciones químicas, con base en la energía de correlación	24
2.5.1 La cancelación de la energía de correlación .....	24
2.5.2 Reacciones homodésmicas .....	25
2.5.3 Reacciones isodésmicas .....	25
2.5.4 Reacciones anisodésmicas .....	26
2.6 Aproximación al límite de Hartree - Fock .....	26
2.7 La correlación electrónica .....	27
2.7.1 Los métodos de interacción de configuraciones CI .....	27
2.7.2 Efectos no dinámicos de correlación .....	28
2.7.3 Efectos dinámicos de correlación .....	30
2.8 La energía de punto cero (ZPE) y los efectos térmicos .....	32
2.9 Cálculos cuánticos de la reacción de hidrólisis del pirofosfato .....	33

2.10	Referencias del capítulo 2 .....	35
<b>3</b>	<b>Simulaciones de los líquidos moleculares</b>	<b>40</b>
3.1	La mecánica estadística y la función de partición .....	41
3.2	La función de partición molecular .....	42
3.3	La función de partición cinética y la configuracional .....	44
3.4	Métodos de simulación numérica .....	46
3.5	La información que se obtiene de las simulaciones numéricas .....	47
3.6	La energía de hidratación .....	49
3.7	Referencias del capítulo 3 .....	51
<b>4</b>	<b>La reacción de hidrólisis del pirofosfato en la fase gaseosa</b>	<b>54</b>
4.1	Antecedentes .....	55
4.2	Descripción de los métodos empleados .....	55
4.3	Estructuras de las moléculas de ortofosfato y de pirofosfato .....	56
4.4	La energía de la reacción de hidrólisis del pirofosfato .....	59
4.4.1	Cálculos al nivel SCF .....	59
4.4.2	Inclusión de la energía de correlación y de la ZPE .....	62
4.5	La entalpía de la reacción de hidrólisis del pirofosfato en la fase gaseosa .....	63
4.6	Referencias del capítulo 4 .....	65
<b>5</b>	<b>La reacción de hidrólisis del pirofosfato en la solución acuosa</b>	<b>67</b>
5.1	La estimación de las energías de hidratación .....	68
5.1.1	Las aproximaciones termodinámicas .....	68
5.1.2	La aproximación por cálculos moleculares .....	68
5.1.3	La aproximación por simulaciones numéricas .....	68
5.2	Descripción del método desarrollado en esta tesis .....	69
5.3	La interacción del ortofosfato con el agua .....	71
5.4	La interacción del pirofosfato con el agua .....	79
5.5	La descripción del agua líquida con el potencial MCHO .....	85

5.5.1	La convergencia de los resultados de las simulaciones .....	85
5.5.2	La energía intermolecular del agua líquida .....	88
5.5.3	La estructura del agua líquida a temperatura ambiente ( $T = 298 K$ ) .....	89
5.6	Las energías de hidratación del ortofosfato y del pirofosfato .....	95
5.7	Referencias del capítulo 5 .....	99
<b>6</b>	<b>Discusión, conclusiones y perspectivas</b>	<b>102</b>
6.1	Discusión y conclusiones .....	103
6.2	Perspectivas .....	105
6.3	Referencias del capítulo 6 .....	106
	<b>Apéndice A. Algunos conceptos fisicoquímicos usados en la bioquímica</b>	<b>107</b>
A.1	Los parámetros termodinámicos .....	108
A.2	Los potenciales termodinámicos en las reacciones químicas .....	111
A.3	El potencial químico, la fugacidad y la actividad .....	114
A.4	El equilibrio y las velocidades de las reacciones .....	118
A.5	Los mecanismos de reacción con base en la teoría de colisiones .....	120
A.6	Referencias del apéndice A .....	123
	<b>Apéndice B. Superficies de interacción del agua con los fosfatos</b>	<b>124</b>
	<b>Apéndice C. Artículos relacionados con esta tesis</b>	<b>140</b>

## INTRODUCCIÓN

---

En la bioquímica actual se están desarrollando nuevas técnicas experimentales y nuevos métodos, que han originado la necesidad de cambiar la descripción de *sustancias* por la descripción de *moléculas*. Este nuevo punto de vista requiere de una mayor precisión en los datos experimentales, y de un mejor entendimiento de las interacciones intermoleculares: se trata de refinar un panorama termodinámico de conjunto, y obtener una visión molecular detallada.

En la mayoría de los casos, las técnicas experimentales que pueden alcanzar la precisión requerida, son costosas. Y la interpretación de los resultados está sujeta al uso de parámetros diseñados *ad hoc*.

Por otra parte, la descripción de los fenómenos moleculares, a partir de la mecánica cuántica, se ha visto limitada a sistemas *chicos* (que contienen menos de diez átomos), por la dificultad matemática de la ecuación de Schrödinger. Las soluciones de esta ecuación, caracterizan las propiedades más relevantes en el intervalo de energías propias de los enlaces químicos. Su aplicación a sistemas *grandes*, requiere de diversas aproximaciones, algunas de las cuales pueden ser muy inexactas.

La intención de esta tesis, es mostrar la factibilidad de una descripción molecular detallada y útil de los fenómenos bioquímicos, basada en primeros principios. Aunque el uso de modelos moleculares ya ha sido aplicado a diversos sistemas, en la mayoría de esos modelos se han empleado parámetros empíricos. Por otro lado, existe el esquema que parte de cálculos cuánticos *ab initio*, desarrolla potenciales analíticos de interacción intermolecular y efectúa simulaciones numéricas. Con los cálculos se puede describir la fase gaseosa, y por medio de las simulaciones se incluyen los efectos térmicos y colectivos. En este esquema no se introduce ningún parámetro empírico, lo que proporciona la ventaja de discernir el efecto de cada aproximación usada.

Hace quince años se tenía una descripción del agua líquida muy poco precisa, pero gracias a los estudios con cálculos *ab initio* y con simulaciones, actualmente se ha llegado al punto de forzar la revisión de los datos experimentales. El trabajo que se presenta aquí, sigue el patrón de las simulaciones basadas en primeros principios. Para alcanzar el refinamiento que permitiera obtener datos relevantes al entendimiento de un fenómeno bioquímico, fue necesario desarrollar un potencial analítico, que incluyera efectos no aditivos de muchos cuerpos: el potencial MCHO. El esquema de programación tuvo que ajustarse a las características del modelo.

El método propuesto en el párrafo anterior, se aplicó al estudio de la hidrólisis del pirofosfato. En el primer capítulo se plantea el problema específico sobre el origen de la energía liberada en la reacción. En dicho capítulo se presenta un panorama general sobre las teorías de la catálisis enzimática, en relación a las hidrólisis y a las síntesis de los compuestos fosforilados de *alta energía*. También se hace una revisión de los resultados experimentales y teóricos al respecto.

En los planteamientos de las explicaciones sobre la *alta energía* de los anhídridos de fosfato, se maneja la idea de que la reacción es termoneutra en la fase gaseosa; pero no se cuenta con datos experimentales en esa fase, y los cálculos ya hechos dan resultados contradictorios. Esto se debe a que las soluciones de la ecuación de Schrödinger se obtienen de forma numérica y con una serie de aproximaciones; en consecuencia, la calidad de los resultados depende de la capacidad de cómputo disponible. En el segundo capítulo se presenta una revisión de las aproximaciones, así como de los criterios que se han desarrollado para juzgar la calidad de los cálculos, y de los problemas que plantea su aplicación al estudio de sistemas *grandes* (que contienen más de diez átomos). El uso de los criterios ya mencionados, permite que en este trabajo se llegue a un valor confiable de la entalpía de la hidrólisis del pirofosfato en la fase gaseosa.

En el tercer capítulo se presenta un panorama general sobre las simulaciones numéricas y su relación con la termodinámica, a través de la mecánica estadística, haciendo referencia al tipo de análisis que puede efectuarse con los datos de las simulaciones; en particular respecto a las energías de la hidratación y las relaciones estructurales de unas moléculas con otras.

El cuarto capítulo se dedica a la reacción de hidrólisis del pirofosfato en la fase gaseosa, planteando la utilidad de los cálculos *ab initio*. En ese capítulo se describe en detalle el método sistemático que se desarrolló para el estudio de los calores de las reacciones químicas en la fase gaseosa. Se muestran los resultados sobre la estructura de las moléculas de interés, y sobre la energética de la reacción de hidrólisis del pirofosfato.

En el quinto capítulo se propone el estudio de la reacción de hidrólisis en solución acuosa, por medio de simulaciones numéricas. También se describen las aproximaciones y los métodos empleados, así como los resultados obtenidos sobre la descripción del agua líquida, y sobre las energías de hidratación de los fosfatos. De los capítulos cuarto y quinto, se obtiene una descripción que ayuda a aclarar las propuestas sobre la *alta energía* de los compuestos fosforilados.

Las conclusiones generales se presentan en el sexto capítulo. También se plantean las perspectivas del método desarrollado, tanto para la continuación del estudio de la bioenergética del pirofosfato, como para su extensión a otros sistemas.

Se incluyen en la tesis tres apéndices. El primero de ellos consiste en un resumen de diversos conceptos fisicoquímicos que están directamente involucrados con los temas tratados en esta tesis, como son los potenciales termodinámicos y la actividad del solvente. Aunque se trata de un conocimiento general ya establecido, y accesible en cualquier texto de fisicoquímica, se consideró adecuado incluirlo para tener una referencia inmediata que permitiera una lectura más ágil de la tesis.

El apéndice B contiene las coordenadas de las moléculas de agua, de ortofosfato y de pirofosfato, que se usaron en los cálculos de las superficies de energías por pares, y de las contribuciones no aditivas de tríos; a las cuales se ajustaron los parámetros de los potenciales de interacción que se emplearon en las simulaciones.

El apéndice C consta de tres artículos del autor de esta tesis. El primero describe el potencial analítico usado en las simulaciones MC, del trabajo aquí presentado, y demuestra la gran capacidad del potencial para describir los parámetros estructurales del agua; en otras palabras, justifica el uso del potencial MCHO para obtener datos confiables sobre la energía y sobre la estructura de una solución acuosa.

El segundo artículo del apéndice C es una revisión de resultados, tanto de simulaciones como experimentales, sobre los efectos estructurales y energéticos que tienen las membranas sobre el agua líquida. La relevancia de esta revisión con respecto al tema de la tesis, reside en que tanto las energías de hidratación, como la entropía del sistema, podrían verse afectadas por la estructura del agua en el ambiente cercano a la enzima. En ese artículo se usaron los programas que se desarrollaron para la obtención de parámetros energéticos y estructurales de líquidos, a partir de simulaciones con el potencial MCHO.

El tercer artículo en el apéndice C, es el estudio sobre la reacción de hidrólisis del pirofosfato en la fase gaseosa, que se realizó por medio de cálculos *ab initio*, como parte de esta tesis.

En resumen, se presenta un análisis de los conceptos teóricos para la descripción de algunos fenómenos bioquímicos, en términos de primeros principios. También se establece un método con el que se logra llevar a la práctica esta descripción, y se aplica a un caso específico. El logro más relevante, está en haber desarrollado un método práctico, que permite conectar la teoría básica con la aplicación específica. Además, gracias a la capacidad de estudiar las distintas contribuciones por separado, fue posible proponer una nueva explicación sobre la alta reactividad de los anhídridos de fosfato.

## CAPÍTULO 1

---

# CATÁLISIS ENZIMÁTICA DE LAS REACCIONES DE HIDRÓLISIS DE LOS COMPUESTOS FOSFORILADOS DE *ALTA ENERGÍA*

### 1.1 Los efectos de los solventes en la bioquímica

Las reacciones químicas que ocurren en solución, son distintas de las que ocurren en fase gaseosa; fundamentalmente debido a las interacciones de los reactantes y de los productos con el solvente. Pero además existen otros efectos estéricos y de acomodación que afectan la solubilidad de distintas sustancias.

Aunque generalmente se acepta que el líquido más importante en la bioquímica es el agua, hasta fechas relativamente recientes sólo se le había considerado como un medio inerte en el que ocurren la gran mayoría de las reacciones bioquímicas. En cambio, la visión actual considera la participación del solvente en la energética de esas reacciones: el agua es un solvente polar y forma puentes de hidrógeno, por lo que puede hidratar con facilidad solutos iónicos y polares, alrededor de los cuales se agrupa en capas de hidratación. Existen diversos estudios sobre la hidratación de distintas moléculas (e. g. [1-3]), mostrando que la energía de interacción de un ión pequeño con su primera capa de hidratación puede ser de varias decenas de kcal/mol (e. g. [2, 4, 5]) y decrece para la segunda capa. Incluso se ha encontrado que en las proteínas, algunas moléculas de agua están tan fuertemente ligadas, que forman parte de la estructura misma de la proteína [6]. Los efectos de la hidratación son, por lo tanto, de gran importancia en la bioquímica.

Por otra parte, las moléculas apolares no son solubles en agua, pero pueden adoptar diversas formas de agregación en mezclas con agua, que van desde las moléculas dispersas hasta las micelas y las bicapas, dependiendo de la concentración del soluto. Se piensa que la formación de estas estructuras se debe al efecto hidrofóbico, que puede definirse fenomenológicamente en términos de la baja solubilidad de las moléculas no polares en el agua [7]. También se cree que las interacciones físicas subyacentes que son responsables de este efecto están relacionadas con la ruptura (costosa en términos de energía) del arreglo de puentes de hidrógeno (HB) en el agua, producida por las moléculas del soluto que no pueden formar HB's ellas mismas [8]. De hecho, aún existe mucha discusión sobre el efecto hidrofóbico y su origen [8, 9]. Lo que queda claro es que el agua no es tan sólo un medio inerte en el que ocurren las reacciones químicas, sino que participa activamente en ellas, no sólo como reactante, sino afectando sus parámetros termodinámicos.

### 1.2 Catálisis enzimática en distintos solventes

La mayor parte de los procesos metabólicos en las células ocurren dentro de compartimentos acuosos, por lo que la teoría clásica de la cinética enzimática, se ha desarrollado con base en los resultados obtenidos con enzimas en soluciones acuosas. Sin embargo, en fechas recientes se ha planteado la duda sobre cuán diferente es el ambiente fisicoquímico intracelular con respecto a las soluciones acuosas que se manejan en los laboratorios [10]. Algunos autores afirman que el agua celular no es distinta del agua libre (*bulk*) [11], en tanto que otros proponen que el agua celular interactúa con diversos solutos y superficies membranales, lo cual da lugar a un ambiente acuoso con propiedades fisicoquímicas diferentes a las del agua pura [12].

Muchos de los métodos usados para estudiar las propiedades físicas del agua celular

tales como la densidad, la viscosidad, la capacidad calorífica, la movilidad traslacional y rotacional, y la constante dieléctrica, implican el uso de condiciones experimentales muy drásticas, que no son toleradas por la mayor parte de las células ni de los tejidos. Sin embargo, en quistes de un crustáceo llamado *Artemia*, ha sido posible hacer estudios de resonancia magnética nuclear (NMR) y de dispersión cuasielástica de neutrones (QND) sobre los tiempos de relajación y el coeficiente de difusión. Con estos estudios se encontró que la movilidad del agua en los quistes es sensiblemente menor que la del agua pura [13]. Otro tipo de estudios sobre la difusión de proteínas inyectadas al citoplasma de algunas células vivas, han mostrado diferencias en la viscosidad del agua en las células con respecto a la del agua pura [14, 15].

Ante este tipo de evidencias experimentales, se ha comenzado a manejar la idea de que la mayor parte del agua celular interactúa fuertemente con las membranas celulares, lo que da lugar a una población altamente estructurada, conocida como *agua de interfase* [16]. Incluso se ha propuesto que el agua localizada hasta a 500 Å desde la superficie de una membrana corresponde a agua de interfase [17], lo cual, combinado con la presencia de los organelos celulares, implicaría que prácticamente toda el agua citoplásmica está bajo la influencia de alguna superficie membranal. Existen, empero, una serie de resultados experimentales y de simulaciones numéricas que no muestran un efecto estructural de las membranas sobre el agua líquida, a más de 10 Å a partir de la membrana. Se realizó una revisión de estos datos [18] (artículo adjunto en el apéndice C), como parte de una fundamentación del problema de la estructura del agua, ya que para el caso de estudio de esta tesis, esa estructura podría tener un efecto importante sobre las energías de hidratación de las especies moleculares empleadas y, con ello, sobre la entalpía de la reacción, como se plantea en los párrafos siguientes, en especial en las secciones 1.4 a 1.6.

La idea de que la estructura del agua afecte a las reacciones bioquímicas, ha sido explorada por varios investigadores, y ya ha dado origen a una serie de especulaciones sobre posibles mecanismos para explicar la catálisis enzimática. Una de estas propuestas es la *hipótesis de desolvatación* de Dewar y Storch [19], según la cual las barreras de activación de las reacciones catalizadas por enzimas, se deben por completo al impedimento impuesto por el solvente, en particular debido a la necesidad de eliminar las capas de solvatación para que los reactantes puedan aproximarse. Dewar y Storch llaman a este tipo de reacciones *solvactivadas*. Bajo esta suposición, la misión de la enzima consiste en remover el agua de los sitios activos, para así permitir que los reactantes interactúen en un ambiente idéntico al de la fase gaseosa, en la cual no hay barreras de activación adicionales. La hipótesis de desolvatación ha sido criticada por Warshel *et al.* [20] en el sentido de que las especies moleculares que participan en las reacciones catalizadas por las enzimas, son iónicas; y se requiere de una gran cantidad de energía para llevar los iones de una solución acuosa a la fase gaseosa. Ellos dan como contrapropuesta una *hipótesis de resolvatación*, sugiriendo que los sitios activos de las enzimas proveen de ambientes polares específicos que no se parecen a la fase gaseosa, sino que están diseñados para la estabilización electrostática de los estados de transición iónicos y que *solvatan* a estos estados mejor de lo que lo hace el agua; es decir, que los sitios activos ofrecen a los sustratos un ambiente energéticamente más favorable que la solución acuosa. Esto puede verse como una variación de la teoría

del estado de transición [21], según la cual las enzimas tienen una mayor afinidad por los sustratos en sus estados activados (véase la sección 5 del apéndice A) y, por ende, contribuyen a aumentar la probabilidad de la colisión entre dos moléculas capaces de reaccionar.

Es claro que una explicación de la catálisis enzimática requiere de una descripción molecular. Un problema de la aproximación experimental es que obtener resolución molecular es muy difícil, y por lo general se trata de mediciones termodinámicas sobre una población. En cambio el ataque al problema, hecho a partir de cálculos moleculares y simulaciones numéricas, sí permite estudiar el comportamiento molecular en detalle. Este tipo de aproximación presenta una serie de condiciones tales que para dar resultados confiables se requiere de una gran capacidad de cómputo, por lo que su aplicación apenas comienza a producir resultados relevantes para elucidar los mecanismos de la catálisis enzimática. Así, por ejemplo, en un trabajo muy reciente, Karplus *et al.* [22] hicieron una simulación numérica de la ruta de reacción de la triosafosfato isomerasa, de la cual concluyen que las especies cargadas que se presentan a lo largo de la ruta de reacción, tienden a tener interacciones estabilizadoras significativas con la enzima, con una energía similar a la de la hidratación en solución acuosa; y por lo tanto, favorecen la *resolvatación* de Warshel sobre la *desolvatación* de Dewar. Son estas interacciones de resolvatación, las que permiten a la enzima alterar la energía libre de la barrera de una reacción hasta en varias decenas de kcal/mol. Dada la *barrera de comunicación* que suele existir entre los bioquímicos experimentales y los fisicoquímicos teóricos, vale la pena mencionar que este artículo apareció en la revista *Biochemistry*, que se dedica principalmente a los estudios experimentales. Resulta muy alentador el observar el abatimiento de la *barrera de comunicación* mencionada arriba.

### 1.3 Catálisis enzimática de la reacción de hidrólisis del pirofosfato

El caso bajo estudio en esta tesis es la reacción de hidrólisis del pirofosfato, tomada en principio como modelo de las hidrólisis de los fosfatos de *alta energía*, debido a que su energía libre es de  $\Delta G \sim -7\text{kcal/mol}$ , y a que se trata del anhídrido de fosfato más simple (las medidas calorimétricas de George *et al.* [23] y otros resultados experimentales se discutirán más adelante). Además, su relevancia en el metabolismo ha sido demostrada por varios investigadores [24-28]. De hecho, se ha encontrado que en varios organismos la síntesis de pirofosfato está acoplada a un gradiente electroquímico, o que su hidrólisis puede generar uno [29], así que la descripción que se haga del pirofosfato sirve como una buena base para formular una teoría general sobre los compuestos fosforilados de *alta energía*.

El problema de los compuestos fosforilados de *alta energía* consiste básicamente en entender el origen de ésta; es decir, encontrar a qué se debe que en la hidrólisis se liberen  $\sim 7\text{kcal/mol}$ . Además, en el caso del ATP y del pirofosfato, las enzimas que catalizan la síntesis son las mismas que catalizan la hidrólisis; aquí la pregunta es ¿cómo hacen las enzimas para catalizar un proceso cuesta arriba, si los catalizadores no afectan la energía libre de una reacción (ver la sección A.5)? El paradigma que se usa como respuesta a esta pregunta, es la hipótesis quimiosmótica de Mitchell [30]: la energía necesaria para la síntesis la proporciona un gradiente electroquímico, formado a través de una membrana.

Sin embargo, en la última década se ha acumulado información que indica la posible necesidad de modificar este planteamiento:

Con experimentos sobre el intercambio de oxígeno marcado, el grupo de Boyer [31-33] ha encontrado que es posible sintetizar ATP y pirofosfato sin la presencia de una membrana. Estos resultados han sido corroborados por los grupos de De Meis [34] y de Gómez-Puyou [35]. La explicación propuesta es que la síntesis de ATP (o de pirofosfato) es termoneutra en el ambiente en el que se desarrolla, y la energía del gradiente electroquímico se necesita para separar al producto del sitio activo de la enzima. A este respecto Blyumenfel'd y Tikhonov [36], han elaborado un modelo de *compuerta* sobre la ATP-sintetasa. Según este modelo, el *acto elemental* de la síntesis de ATP a partir del ADP y del Pi en las estructuras de las mitocondrias, de los cloroplastos y de los cromatóforos, no requiere ni la presencia de vesículas de membrana cerrada ni residuos de membrana ni la componente  $F_0$  de la  $H^+$  ATP-asa; sino que se realiza a través de una muy rápida deprotonación de algunos grupos ácidos del factor  $F_1$  de la  $H^+$  ATP-asa, que da por resultado que la enzima adopte un estado conformacional fuera del equilibrio. La liberación del ATP así formado, ocurre durante la relajación a partir del estado conformacional mencionado. Blyumenfel'd y Tikhonov concluyen que las vesículas de membrana cerrada, se necesitan para que la enzima regrese periódicamente al estado inicial, con sus grupos ácidos protonados; así se mantiene la posibilidad de la síntesis repetida de ATP, bajo condiciones estacionarias y cuasiestacionarias. Un grupo de resultados que apoya a este modelo, es el que en preparaciones de mitocondrias de hígado de rata y de corazón de res, se puede sintetizar ATP a partir de ADP y Pi, como respuesta a un incremento del pH del medio [37, 38]. Más aún, ya que la formación de ATP inducida por el aumento en pH, no es inhibida por cianuro ni por antimicina A, se concluye que la síntesis no requiere del funcionamiento de la cadena transportadora de electrones.

Todos estos datos y teorías llevan a pensar que la reacción de síntesis, y por lo tanto también la de hidrólisis del ATP y del pirofosfato, son termoneutras en el sitio activo de la enzima. Entonces resulta necesario encontrar el origen de las  $\sim 7kcal/mol$  que se liberan en su hidrólisis.

#### 1.4 ¿De dónde proviene la *alta energía* de los anhídridos de fosfato?

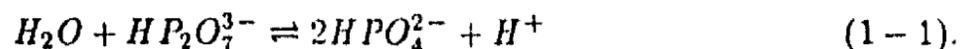
Desde que se descubrió la relevancia metabólica del sistema ADP - ATP, se propusieron explicaciones sobre la *alta energía* de la hidrólisis [39-42]. Estas explicaciones se basaban en considerar al ATP completamente deprotonado en solución,  $ATP^{4-}$ . Entonces, la repulsión electrostática entre los grupos cargados negativamente, habría de contribuir en gran medida a la liberación de energía al ser separados. Además, la distribución electrónica en los puentes  $P - O - P$  presentes en los enlaces de los anhídridos, es tal que la energía disminuye al separarlos; este efecto se conoce como *resonancia en oposición*, y consiste en considerar que en los residuos tales como la formulación  $> (O = P) - O - (P = O) <$  del  $\mu$ -oxo difosforilo, es imposible obtener estructuras de resonancia que involucren simultáneamente al puente  $P - O - P$  y a los dos oxígenos de los fosforilos. En cambio después de la hidrólisis, cada fósforo está conectado a dos oxígenos que no forman ningún puente, y se estima que es posible obtener estructuras de resonancia para los dos al mismo

tiempo, aumentando la estabilidad de los productos con respecto a los reactantes [41]. La conclusión a la que se llega, es que la hidrólisis del ATP y del pirofosfato, debería ser exergónica en la fase gaseosa, en clara contradicción con la explicación discutida arriba para los datos experimentales de Boyer, de De Meis y de Gómez-Puyou. Sin embargo, esta conclusión se obtiene de una descripción de la reacción en la que se omiten las condiciones del microambiente en el que tiene lugar, es decir, no se toman en cuenta los efectos del solvente ni los de la propia enzima.

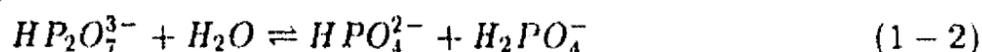
El primer trabajo en resaltar la importancia de los efectos del solvente en las reacciones de hidrólisis de los anhídridos de fosfato, fue el de George *et al.* [23]. En este artículo se presentan datos calorimétricos sobre las energías libres, las entalpías y las entropías de las hidrólisis de diversas especies moleculares. También se hace una estimación de las energías de hidratación de los reactantes y de los productos, usando un modelo de campo promedio y datos experimentales sobre la disolución de diversas sales. Los autores muestran que las diferencias entre las energías de hidratación de los reactantes y las de los productos, pueden ser comparables a las energías libres medidas para las reacciones en solución acuosa; de ahí concluyen que la descripción molecular de las reacciones bioenergéticas, debe incluir los efectos de la energía de interacción del solvente con los solutos, y especulan sobre la posibilidad de que las reacciones de hidrólisis sean termoneutras en la fase gaseosa. Desgraciadamente no existen datos experimentales sobre las reacciones en la fase gaseosa, principalmente por la dificultad de tener las moléculas de los fosfatos en dicha fase. Para sortear esta dificultad, los bioquímicos han decidido investigar los efectos del solvente por medio de la comparación de las reacciones efectuadas en distintos líquidos, en tanto que los fisicoquímicos han recurrido a los cálculos cuánticos, que justamente representan a la fase gaseosa.

### 1.5 La reacción de hidrólisis del pirofosfato en diversos solventes

Los resultados de George *et al.* [23] corresponden a medidas calorimétricas de las entalpías y las entropías de las reacciones de hidrólisis con diversas especies. Una forma alternativa para obtener la energía libre, consiste en medir las concentraciones de los reactantes y de los productos, determinar una constante de equilibrio  $K_{eq} = [\text{productos}]/[\text{reactantes}]$  (donde los paréntesis cuadrados denotan la concentración de la especie indicada), y con ella calcular  $\Delta G = -RT \ln K_{eq}$ . En principio, la ventaja de esta aproximación consiste en la gran precisión con la que se pueden medir las concentraciones de diversas sustancias. Así, por ejemplo, Flodgaard y Fleron [43] emplearon un método de sustitución isotópica para obtener una constante de equilibrio iónica  $K_{ion}$  para la reacción de hidrólisis



La energía libre que ellos proponen para las condiciones intracelulares, es de tan sólo  $\Delta G = -4.0 \text{ kcal/mol}$ , considerando  $\text{pH} = 7.4$ ,  $T = 298 \text{ K}$ , concentraciones de  $1 \text{ mM}$  de  $Mg^{2+}$  libre,  $150 \text{ mM}$  de  $K^+$  libre,  $I = 0.25$ ,  $2.42 \text{ mM}$  de  $P_i$  libre y concentraciones de  $PP_i$  correspondientes a las que se han medido en biopsias de hígado de ratón congelado. La reacción (1-1) no es reportada específicamente por George *et al.* [23], pero la medida de  $\Delta G^\circ$  que ellos reportan para la reacción



es de  $\Delta G^\circ = -7.1 \text{ kcal/mol}$ , a  $T = 298 \text{ K}$  y concentraciones 1 molal de los solutos. Las medidas calorimétricas de  $\Delta H^\circ$ , se hicieron con soluciones no amortiguadas (fuerza iónica  $I \leq 0.01$ ), bajo condiciones tales que la formación de complejos con iones metálicos era insignificante, en tanto que los valores de  $\Delta S^\circ$  se obtuvieron de la determinación de  $S^\circ$  para sales cristalinas de  $\text{Na}_2\text{H}_2\text{P}_2\text{O}_7$  y de  $\text{Na}_3\text{P}_3\text{O}_{10} \cdot 6\text{H}_2\text{O}$ ; de medidas de coeficientes de actividad y calores de solución, y de estudios de solubilidad.

Un detalle que merece atención, es que la energía de la "reacción de hidrólisis del  $PP_i$ " depende de lo que se entienda por  $P_i$  y por  $PP_i$ : las reacciones (1-1) y (1-2) no son iguales y no tendrían que liberar la misma energía, independientemente de que pudieran ocurrir en el mismo microambiente. Un problema de la comparación de unos resultados con otros, es referirlos a la misma reacción. En el caso del pirofosfato, De Meis [50] ha mostrado que existe una veintena de reacciones de hidrólisis que involucran distintas especies de  $P_i$  y de  $PP_i$  y que, por supuesto, tienen diferentes parámetros termodinámicos. Además, cabe hacer notar que en soluciones reales, la constante de equilibrio  $K_{eq}$  se calcula con las actividades, y no con las concentraciones de los solutos. Hay que recordar que la actividad es un índice asociado a la interacción del soluto con el solvente (véase la sección A.3), para denotar las desviaciones respecto de los casos ideales. Se trata, por lo tanto, de un parámetro empírico que no está basado en primeros principios. De hecho, las medidas de Flodgaard y Fleron [43] están corregidas con los coeficientes de actividad.

Una de las conclusiones más relevantes a las que llegan estos últimos autores, es que a pH 7.4 el aumento en la concentración de  $\text{Mg}^{2+}$  hace que la reacción de hidrólisis del pirofosfato sea menos favorable. Otros estudios han mostrado que el verdadero sustrato de la pirofosfatasa es el  $\text{Mg}(\text{H}_2\text{O})_4\text{P}_2\text{O}_7^{2-}$  [44-46], y que además se requiere de  $\text{Mg}^{2+}$  libre para la hidrólisis [47, 48]. Resultados similares se han obtenido para la hidrólisis del ATP [49]. Se han hecho otros estudios sobre la dependencia que tienen los parámetros termodinámicos, en función de otros cationes divalentes, como el  $\text{Ca}^{2+}$  [48, 50, 51], y del pH del medio [35, 50, 52]. Estas caracterizaciones se han realizado a partir de medidas de las concentraciones de diversos reactantes y productos en solución acuosa, y en soluciones con diversas mezclas de líquidos apróticos, con la finalidad de determinar qué papel juegan las interacciones con el solvente y la actividad del agua en la energía de hidrólisis de los anhídridos de fosfato [53]. A este grupo de experimentos pertenece la serie de estudios de De Meis [52], quien encuentra que la constante de equilibrio  $K_{obs}$  cambia varios órdenes de magnitud al variar la relación  $v/v$  de dimetilsulfóxido en agua o la relación  $w/v$  de polietilenoglicol en agua, de forma que la  $\Delta G_{obs}$  para la hidrólisis del pirofosfato va desde -5.0 kcal/mol en agua a pH 7.0, con 1 mM de  $\text{Mg}^{2+}$ , hasta -1.6 kcal/mol en una solución de DMSO 50% ( $v/v$ ) a pH 7.2, con 8 mM de  $\text{Mg}^{2+}$ , e incluso hasta +1.0 kcal/mol en PEG 50% ( $w/v$ ) a pH 8.0, con 0.9 mM de  $\text{Mg}^{2+}$ . En cambio, la hidrólisis del éster de fosfato en la fosfoserina, prácticamente no se ve modificada al ocurrir en solventes distintos del agua [54]. Los datos de De Meis también muestran que la concentración de  $\text{Mg}^{2+}$  y el pH pueden modificar la  $K_{obs}$  de la hidrólisis del pirofosfato.

La serie de datos acumulados por Boyer [31-33], por Wolfenden y Williams [53] y por De Meis [34, 50, 52, 54] muestra que la contribución del solvente a la energía libre de la

hidrólisis de los anhídridos de fosfato, es muy relevante. Las primeras propuestas para explicar el cambio en la  $\Delta G_{obs}$  liberada en la reacción, consideran que el sitio activo de la enzima permite a los reactantes interactuar como si estuvieran en fase gaseosa [55], usando una hipótesis similar a la *desolvatación* de Dewar y Storch [19]. Sin embargo, los datos de De Meis muestran que la contribución del solvente a la  $\Delta G^{\circ}$  de la reacción, no ocurre en el término entálpico  $\Delta H^{\circ}$ , sino en el entrópico,  $T\Delta S^{\circ}$  (ver *e. g.* [54]), lo cual no requiere que la hidrólisis sea termoneutra en la fase gaseosa. Más aún, De Meis muestra con una gráfica de van't Hoff, que la entalpía de la reacción debe ser negativa [50]. En tal caso, se propone que la contribución que regula la  $\Delta G^{\circ}$  de la reacción en solución es la *actividad* del agua.

En resumen, de los datos experimentales obtenidos para la reacción de hidrólisis del pirofosfato en diversos solventes, se pueden obtener las siguientes conclusiones:

- a) La reacción es termoneutra en el sitio activo de la enzima [31-35].
- b) El sitio activo es hidrofóbico [53].
- c) Se requiere la presencia de  $Mg^{2+}$  para catalizar la hidrólisis [47, 48].
- d) La contribución del término entrópico  $T\Delta S^{\circ}$  puede determinar la  $\Delta G^{\circ}$  de la reacción [54].

Como se puede notar, la situación es más compleja de lo que se pensaba en un principio, y no sólo involucra al solvente. En los párrafos anteriores se mencionó el problema de contar con diversas reacciones de hidrólisis del  $PP_i$ , lo cual dificulta la interpretación de los resultados, dada la multitud de variables que deben tomarse en cuenta. En principio, es posible controlar la reacción específica que ocurre en cada experimento, así como las condiciones de pH y de concentración de iones. Esto permite discriminar los efectos debidos a la modificación de cada condición. Aun así, no es posible generar una imagen molecular detallada de lo que ocurre con los solutos y los solventes, debido a la falta de información sobre la fase gaseosa y sobre el comportamiento de las moléculas aisladas.

Como se mencionó en la sección 1.4, una forma alternativa de estudiar una reacción química en la fase gaseosa, es la de efectuar cálculos moleculares del sistema. En el siguiente capítulo se presenta una descripción general de este tipo de cálculos, y en la sección 2.9 se hace una discusión sobre los que ya se han reportado para la reacción de hidrólisis del pirofosfato.

En la aproximación teórica, es posible analizar cada especie molecular en forma separada, y luego ir agregando las otras; lo cual permite estudiar la contribución de cada interacción. Además, de las simulaciones se pueden obtener parámetros estructurales de las esferas de hidratación alrededor de cada especie.

En esta tesis se muestra que la hidrólisis del pirofosfato neutro es exotérmica en la fase gaseosa, y que la diferencia de entalpías entre la hidratación de los productos y la de los reactantes, tiene gran relevancia. De hecho, se concluye que existen tres términos que son

casi igualmente relevantes en sus contribuciones a la  $\Delta G^\circ$  de la hidrólisis del pirofosfato: (1) la entalpía en la fase gaseosa: los efectos intramoleculares de rompimiento de algunos enlaces y formación de otros nuevos; (2) la diferencia en las energías de hidratación: las interacciones de los reactantes y de los productos con el solvente, así como las interacciones entre las mismas moléculas del solvente, en presencia de los solutos, y (3) el término entrópico: las diferentes configuraciones que puedan adoptar las moléculas del solvente alrededor de los solutos.

### 1.6 Referencias del capítulo 1

- [1].- M. M. Probst, T. Radnai, K. Heinzinger, P. Bopp y B. M. Rode, *J. Phys. Chem.* **89**, 753 (1985)
- [2].- P. Cieplak y P. A. Kollman, *J. Chem. Phys.* **92**, 6761 (1990)
- [3].- X. G. Zhao, A. González-Lafont, D. G. Truhlar y R. Steckler, *J. Chem. Phys.* **94**, 5544 (1991)
- [4].- E. Kochanski y E. Constantin, *J. Chem. Phys.* **87**, 1661 (1987)
- [5].- M. I. Bernal Uruchurtu, *comunicación personal*
- [6].- M. M. Teeter, *Annu. Rev. Biophys. Biophys. Chem.* **20**, 577 (1991)
- [7].- C. H. Tanford, *The hydrophobic effect*, John Wiley and Sons, Nueva York (1980)
- [8].- K. A. Sharp, E. Nicholls, R. Friedman y B. Honig, *Biochemistry* **30**, 9686 (1991)
- [9].- R. H. Wood y P. T. Thompson, *Proc. Natl. Acad. Sci. USA* **87**, 946 (1990)
- [10].- L. Escobar Pérez y E. Escamilla Marván, *Ciencia* **41**, 307 (1990)
- [11].- E. E. Burnell, M. E. Clarke, J. A. M. Hinke y N. R. Chapman, *Biophys. J.* **33**, 1 (1981)
- [12].- G. N. Ling, C. L. Walton y M. M. Oschenfeld, *J. Cell Physiol.* **106**, 385 (1981)
- [13].- C. F. Hazlewood, en *Cell Associated Water*, W. Drost-Hansen y J. S. Clegg, eds., Academic Press, Nueva York (1979)
- [14].- J. W. Wojcieszyn, R. A. Schlegel, E. S. Wu y K. A. Jacobson, *Proc. Natl. Acad. Sci. USA* **78**, 4407 (1981)
- [15].- J. E. Tanner, *Arch. Biochem. Biophys.* **224**, 416 (1983)

- [16].- W. Drost-Hansen y J. S. Clegg, eds., *Cell Associated Water*, Academic Press, Nueva York (1979)
- [17].- W. Drost-Hansen y J. S. Clegg, en *Biophysics of Water*, F. Franks y S. Mathias, eds., John Wiley and Sons, Nueva York (1982)
- [18].- H. Saint-Martin e I. Ortega-Blake, en *Biomolecules in Organic Solvents*, A. Gómez-Puyou, ed., CRC Press, Boca Ratón, FL (1992)
- [19].- M. J. S. Dewar y J. M. Storch, *Proc. Natl. Acad. Sci. USA* **82**, 2225 (1985)
- [20].- A. Warshel, J. Åqvist y S. Creighton, *Proc. Natl. Acad. Sci. USA* **86**, 5820 (1989)
- [21].- J. Kraut, *Science* **242**, 533 (1988)
- [22].- P. A. Bash, M. J. Field, R. C. Davenport, G. A. Petsko, D. Ringe y M. Karplus, *Biochemistry* **30**, 5826 (1991)
- [23].- P. George, R. J. Witonsky, M. Trachtman, C. Wu, W. Dorwart, L. Richman, W. Richman, F. Shurayh y B. Lentz, *Biochim, Biophys. Acta* **223**, 1 (1970)
- [24].- H. Baltscheffsky, L. V. Von Stedingk, H. W. Heldt y M. Klingenberg, *Science* **153**, 1120 (1966)
- [25].- R. E. Reeves, *TIBS*, p. 53, Marzo (1976)
- [26].- S. E. Mansurova, Yu. A. Shakhov e I. S. Kulaev, *FEBS Lett.* **74**, 31 (1977)
- [27].- M. Baltscheffky, en *The Photosynthetic Bacteria*, R. K. Clayton y W. R. Sistrom, eds., Plenum Press, Nueva York (1978)
- [28].- Yu. A. Shakhov, S. E. Mansurova e I. S. Kulaev, *Biochem. Internat.* **3**, 139 (1981)
- [29].- Yu. A. Shakhov, P. Nyren y M. Baltscheffsky, *FEBS Lett.* **46**, 177 (1982)
- [30].- P. Mitchell, *Chemiosmotic Coupling and Energy Transduction*, Glynn Research Laboratories, Bodmin, Cornwall, Inglaterra (1968)
- [31].- P. D. Boyer, B. Chance, L. Ernster, P. Mitchell, E. Racker y E. C. Slater, *Annu. Rev. Biochem.* **46**, 955 (1977)
- [32].- P. D. Boyer, en *Chemiosmotic Proton Circuits in Biological Membranes*, V. P. Skulachev y P. Hinckle, eds., Addison Wesley, Reading, MA (1981)
- [33].- P. D. Boyer, *TIBS* p. - , Enero (1988)

- [34].- L. De Meis, M. I. Behrens, J. H. Petretski y M. J. Politi, *Biochemistry* **24**, 7783 (1985)
- [35].- L. De Meis, M. I. Behrens, H. Celis, I. Romero, M. T. Gómez-Puyou y A. Gómez-Puyou, *Eur. J. Biochem.* **158**, 149 (1986)
- [36].- L. A. Blyumenfel'd y A. N. Thikhonov, *Biophysics* **32**, 865 (1987)
- [37].- I. V. Malenkova, *Dokl. Akad. Nauk. URSS* **252**, 743 (1982)
- [38].- I. V. Malenkova, *Biochim. Biophys. Acta* **682**, 179 (1982)
- [39].- H. M. Kalckar, *Chem. Rev.* **28**, 71 (1941)
- [40].- T. L. Hill y M. H. Morales, *J. Am. Chem. Soc.* **73**, 1656 (1951)
- [41].- A. Pullman y B. Pullman, *Quantum Biochemistry*, Interscience, Nueva York (1963)
- [42].- D. B. Boyd y W. N. Lipscomb, *J. Theor. Biol.* **25**, 403 (1969)
- [43].- H. Flodgaard y P. Fleron, *J. Biol. Chem.* **249**, 3465 (1974)
- [44].- M. Hørdér, *Biochim. Biophys. Acta* **358**, 319 (1974)
- [45].- W. B. Knight, S. W. Fitts y D. Dunaway-Mariano, *Biochemistry* **20**, 4079 (1981)
- [46].- Su J. Ting y D. Dunaway-Mariano, *FEBS Lett.* **165**, 251 (1984)
- [47].- B. Springs, K. M. Welsh y B. S. Cooperman, *Biochemistry* **20**, 6384 (1981)
- [48].- H. Celis e I. Romero, *J. Bioenerget. Biomembr.* **19**, 255 (1987)
- [49].- A. Gómez-Puyou, G. Ayala, U. Muller y M. T. Gómez-Puyou, *J. Biol. Chem.* **258**, 13673 (1983)
- [50].- L. De Meis, *J. Biol. Chem.* **259**, 6090 (1984)
- [51].- M. Montero-Lomelí, O. B. Martins y G. Dreyfus, *J. Biol. Chem.* **264**, 21014 (1989)
- [52].- L. De Meis, *Biochim, Biophys. Acta* **973**, 333 (1989)
- [53].- R. Wolfenden y R. Williams, *J. Am. Chem. Soc.* **107**, 4345 (1985)
- [54].- P. J. Romero y L. De Meis, *J. Biol. Chem.* **264**, 7896 (1989)
- [55].- Y. Dupont y R. Pugeois, *FEBS Lett.* **156**, 93 (1983)

## CAPÍTULO 2

---

### APLICACIÓN DE LOS CÁLCULOS CUÁNTICOS AL ESTUDIO DE LAS REACCIONES QUÍMICAS

## 2.1 Los cálculos *ab initio*

Al poco tiempo de haberse establecido la Mecánica Cuántica como la teoría adecuada para la descripción de los fenómenos atómicos y moleculares, y la ecuación de Schrödinger como una de sus expresiones matemáticas, se comenzó a tratar de aplicar esta ecuación al estudio de la Química, buscando las soluciones para diversos sistemas. Sin embargo, sólo se han obtenido soluciones analíticas en casos particulares, como en el átomo de hidrógeno [1]. El tratamiento general de los sistemas moleculares se hace con algunas aproximaciones bien definidas, como la separación de Born - Oppenheimer, el principio variacional y la teoría de perturbaciones [2,3]. Cuando en este tipo de soluciones no se usa ningún dato empírico ni parámetro ajustable, el método se define como *ab initio*. La mayor virtud de este método es que se conocen todas las aproximaciones empleadas, lo cual permite estimar cuantitativamente la precisión de sus soluciones, así como proponer la forma de mejorar esta precisión.

La descripción detallada de las aproximaciones mencionadas, puede encontrarse en cualquier texto sobre la teoría de los orbitales moleculares [*e. g.* 2, 3]; pero es adecuado hacer una breve descripción, para tener un criterio con el cual juzgar la confiabilidad de los resultados que se obtienen con los cálculos *ab initio*. Dicha descripción será relevante en esta tesis.

El estudio de las combinaciones de átomos para formar moléculas, puede hacerse al conocer las propiedades electrónicas de los sistemas moleculares y su energía  $E_R$ . Estas propiedades se obtienen de la solución de la parte electrónica de la ecuación de Schrödinger, para una configuración  $R$  dada de los núcleos:

$$H(\mathbf{r}, \mathbf{R})\Psi_R(\mathbf{r}) = E_R\Psi_R(\mathbf{r}) \quad (2-1).$$

La función de onda de  $n$  partículas  $\Psi_R(\mathbf{r})$  describe el movimiento de los electrones en el campo de los núcleos fijos; ésta es la aproximación de Born - Oppenheimer. En el hamiltoniano de la ec. (2-1) existe todavía un término de interacción electrónica que obliga al uso de otras aproximaciones para encontrar la solución. La búsqueda de estas aproximaciones, y el estudio de su aplicabilidad y su confiabilidad, ha dado origen a la química cuántica computacional.

Casi todos los métodos que se han desarrollado para resolver la ec. (2-1) se basan en la aproximación de Hartree - Fock (HF) de campo autoconsistente (SCF), en la que se supone que el movimiento de cada electrón en el sistema molecular, depende sólo del potencial promedio de todos los demás electrones. También se le conoce como el modelo de electrones independientes, ya que no toma en cuenta que en la función de onda exacta, el movimiento de cada electrón está correlacionado con el de los otros. A la diferencia de energías entre el resultado exacto y la aproximación HF, se le llama la *energía de correlación*. La inclusión de esta energía a un nivel razonable, requiere de un esfuerzo computacional mucho mayor que el usado en una solución SCF: para un mismo sistema se puede estimar un esfuerzo de un orden de magnitud mayor en recursos de cómputo [4] (*vide infra*, sección 2.7).

En la aproximación HF, se supone que la función de onda de  $n$  partículas puede

escribirse como un producto antisimetrizado de funciones de un electrón  $\psi_i(\mathbf{r}_i)$  (un determinante de Slater):

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \dots & \psi_n(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \dots & \psi_n(\mathbf{r}_2) \\ \dots & \dots & \dots & \dots \\ \psi_1(\mathbf{r}_n) & \psi_2(\mathbf{r}_n) & \dots & \psi_n(\mathbf{r}_n) \end{vmatrix} \quad (2-2).$$

Estas funciones  $\psi_i(\mathbf{r}_i)$  de un electrón se denominan orbitales. Al buscar el conjunto de  $n$  orbitales que para un sistema molecular produzca la energía más baja, en el sentido del principio variacional, se encuentra que tales orbitales quedan determinados por las ecuaciones:

$$\hat{F}(\mathbf{r}_i)\psi_i(\mathbf{r}_i) = E_i\psi_i(\mathbf{r}_i) \quad (2-3),$$

denominadas ecuaciones de Hartree Fock. Los orbitales  $\psi_i$  se llaman orbitales moleculares, y el operador de Fock,  $\hat{F}$ , está dado por:

$$\hat{F}(\mathbf{r}_1) = \hat{h}(\mathbf{r}_1) + \hat{g}(\mathbf{r}_1) \quad (2-4).$$

La parte monoeléctronica  $\hat{h}(\mathbf{r}_1)$  comprende al operador diferencial de la energía cinética y al potencial de los núcleos, en tanto que el término de interacción electrónica  $\hat{g}(\mathbf{r}_1)$  tiene la forma de un operador monoeléctronico efectivo, y describe el potencial promedio creado por los electrones del sistema. El término  $\hat{g}(\mathbf{r}_1)$  se compone de dos partes, la coulombiana y la de intercambio:

$$\hat{g}(\mathbf{r}_1) = \hat{j}(\mathbf{r}_1) - \hat{k}(\mathbf{r}_1) \quad (2-5a),$$

en que

$$\hat{j}(\mathbf{r}_1) = \sum_{k=1}^n \int \frac{|\psi_k(\mathbf{r}_2)|^2}{|\mathbf{r}_2 - \mathbf{r}_1|} d\mathbf{r}_2 \quad (2-5b)$$

y

$$\hat{k}(\mathbf{r}_1)\psi(\mathbf{r}_1) = \delta_{s,s'} \sum_{k=1}^n \psi_k(\mathbf{r}_1) \int \frac{\psi_k(\mathbf{r}_2)\psi(\mathbf{r}_2)}{|\mathbf{r}_2 - \mathbf{r}_1|} d\mathbf{r}_2 \quad (2-5c).$$

El factor  $\delta_{s,s'}$  surge de la suma sobre las coordenadas de spin, toma el valor 0 cuando las funciones de spin de los orbitales  $\psi$  y  $\psi_k$  son diferentes, y vale 1 cuando la misma función de spin se asigna tanto a  $\psi$  como a  $\psi_k$ . Los operadores  $\hat{j}(\mathbf{r}_1)$  y  $\hat{k}(\mathbf{r}_1)$  son integrales que dependen de las soluciones  $\psi_k$  de la ec. (2-3). Luego las ecuaciones HF son integrodiferenciales y deben resolverse iterativamente hasta que el potencial  $\hat{g}(\mathbf{r}_1)$  sea "autoconsistente". La solución de las ecuaciones HF por integración numérica, es factible sólo para potenciales con un alto grado de simetría. Se han obtenido soluciones numéricas únicamente para átomos y para algunas moléculas diatómicas [4]. Incluso la pequeña molécula de agua representa un obstáculo que aún no ha sido superado. Por consiguiente son necesarias algunas aproximaciones ulteriores, y la que se usa en casi todas las soluciones para moléculas es la

llamada *aproximación algebraica*: por medio de la expresión de los orbitales moleculares  $\psi_i$  como series finitas de una base de funciones  $\chi_i(\mathbf{r})$ ,

$$\psi_i(\mathbf{r}_1) = \chi(\mathbf{r}_1)\mathbf{c}_i = (\chi_1(\mathbf{r}_1), \chi_2(\mathbf{r}_1), \dots, \chi_m(\mathbf{r}_1)) \begin{pmatrix} c_{1i} \\ c_{2i} \\ \dots \\ c_{mi} \end{pmatrix} \quad (2-6)$$

las ecuaciones integrodiferenciales en (2-3) se transforman en ecuaciones matriciales (ecuaciones de Roothan [5]),

$$\mathbf{FC} = \mathbf{SCE} \quad (2-7)$$

con las matrices de Fock y de traslape

$$\begin{aligned} \mathbf{F} &\equiv \langle \chi | \hat{F} | \chi \rangle \\ \mathbf{S} &\equiv \langle \chi | \chi \rangle \end{aligned} \quad (2-8)$$

(La notación con "brackets" de Dirac indica una integración sobre las coordenadas  $\mathbf{r}_i$  de un electrón). Las soluciones HF se obtienen como  $m$  vectores columna  $\mathbf{c}_i$  de coeficientes referidos a un conjunto base dado  $\chi$ :

$$\begin{aligned} \mathbf{C} &= (\mathbf{c}_1, \dots, \mathbf{c}_m) \\ E_{ij} &= E_i \delta_{ij} \end{aligned} \quad (2-9)$$

La función de densidad monoeléctronica,  $\rho(\mathbf{r})$ , para la base elegida se expresa con la matriz:

$$\mathbf{R} = \sum_k^{oc} \mathbf{c}_k \mathbf{c}_k^\dagger \quad (2-10)$$

donde se suma sobre todos los orbitales  $k$  ocupados en el estado electrónico considerado. La matriz de Fock adopta la forma

$$F_{\mu\nu} = \langle \mu | \hat{h} | \nu \rangle + G(\mathbf{R})_{\mu\nu} \quad (2-11),$$

siendo la parte bielectrónica

$$G(\mathbf{R})_{\mu\nu} = \sum_{\rho} \sum_{\sigma} R_{\rho\sigma} [\langle \mu\nu | \rho\sigma \rangle - \langle \mu\rho | \nu\sigma \rangle] \quad (2-12).$$

La forma analítica de las funciones en la base, se elige de manera que las expresiones anteriores se puedan derivar e integrar analíticamente con relativa facilidad. También se busca que tales funciones mantengan las simetrías y las condiciones asintóticas asociadas a las propiedades angulares de los orbitales atómicos. Esta última característica es satisfecha por funciones similares a las soluciones analíticas del átomo de hidrógeno; cuya parte radial tiene una forma exponencial. A estas funciones se les llama orbitales de tipo Slater (STO)

y se les cataloga como a los orbitales atómicos del hidrógeno,  $1s, 2s, 2p_x, \dots$ , con la forma normalizada:

$$\begin{aligned}\phi_{1s} &= \left(\frac{\zeta_1^3}{\pi}\right)^{1/2} \exp(-\zeta_1 r) \\ \phi_{2s} &= \left(\frac{\zeta_2^5}{96\pi}\right)^{1/2} r \exp\left(-\frac{\zeta_2 r}{2}\right) \\ \phi_{2p_x} &= \left(\frac{\zeta_2^5}{32\pi}\right)^{1/2} x \exp\left(-\frac{\zeta_2 r}{2}\right)\end{aligned}\quad (2-13).$$

Las constantes  $\zeta_1$  y  $\zeta_2$  determinan el tamaño de los orbitales. Los STO's proporcionan representaciones razonables de los orbitales atómicos.

Una vez que se ha elegido una base, se puede proceder a evaluar las integrales que aparecen en la ec. (2-12):

$$\langle \mu\nu | \rho\sigma \rangle = \int \chi_\mu^*(\mathbf{r}_1) \chi_\nu(\mathbf{r}_1) \frac{1}{r_{12}} \chi_\rho^*(\mathbf{r}_2) \chi_\sigma(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (2-14).$$

La dificultad para evaluar estas integrales con los STO's, ha limitado su uso en cálculos prácticos. Una alternativa propuesta por Boys [6] son las funciones de tipo gaussiano (GTO):

$$g(\alpha, \mathbf{r}) = N_l(\alpha) x^i y^j z^k \exp(-\alpha r^2) \quad (2-15),$$

donde  $l$  indica el momento angular y  $\alpha$  es una constante que determina el tamaño (la extensión radial) de la función. La constante de normalización  $N$ , depende del valor de  $\alpha$  y del momento angular  $l$ . Los exponentes  $i, j$  y  $k$  son enteros y se eligen de modo que correspondan a la simetría angular del orbital que se quiera representar, bien con una sola función, o bien con una combinación de ellas. Aunque no son tan adecuadas como los STO's, especialmente por carecer de una cúspide en el origen y por su comportamiento asintótico de rápido decaimiento, los GTO's permiten efectuar las integrales (2-14) en forma analítica; por ello se les prefiere en los cálculos prácticos.

El problema ahora reside en que para lograr una buena representación de los orbitales atómicos, es necesario el uso de un mayor número de GTO's en la base, lo que trae como consecuencia que deban evaluarse una gran cantidad de integrales (del orden de millones para moléculas pequeñas, y de miles de millones para moléculas poliatómicas [4]). Formalmente, la cantidad de integrales es proporcional a  $m^4$  (un poco menos en la práctica), donde  $m$  es el número de funciones en la base. Entonces  $m$  debe mantenerse tan pequeño como sea posible. Una forma de lograrlo generalmente aceptada, es el uso de combinaciones lineales fijas de varios GTO's como funciones en la base para cálculos moleculares,

$$\chi_\mu(\mathbf{r}_1) = \sum_j c_{j\mu} g_j(\mathbf{r}_1) \quad (2-16).$$

Con estas combinaciones se busca reproducir más cercanamente la forma y el comportamiento asintótico de los STO's. Los coeficientes de los GTO's primitivos en estos CGTO

contraídos, se transfieren de cálculos en átomos. De hecho, actualmente se dedica mucho esfuerzo para (i) tratar de reducir el número de integrales que se almacenen para un desarrollo dado, y (ii) tratar de encontrar un desarrollo óptimo tan corto como sea posible para un propósito dado. La mayoría de las bases se derivan de cálculos atómicos, usando la idea química intuitiva de que en las moléculas aún es posible identificar átomos ligeramente distorsionados, de forma que la función de onda molecular se describe como una combinación lineal de orbitales atómicos (LCAO, de acuerdo a sus siglas en inglés).

## 2.2 Clasificación de bases para cálculos *ab initio*

Puesto que la selección de una base apropiada es crucial para que un cálculo *ab initio* resulte confiable, es conveniente mencionar la clasificación de las bases, así como la comparación de sus resultados con datos experimentales y con otros cálculos. Una base se denomina *mínima* (MB) cuando se emplea un solo CGTO para cada tipo de orbital atómico ocupado en el estado base del átomo respectivo. Cuando se usa más de una función para cada orbital atómico, la base puede llamarse *doble  $\zeta$*  (DZ), *triple  $\zeta$*  (TZ), y así sucesivamente. Esta nomenclatura se origina apartir del uso de los STO's antes mencionados. Una base DZ emplea el desarrollo de dos funciones con exponentes  $\zeta_1$  y  $\zeta_2$  distintos para describir, *e. g.*, un orbital 1s. Consecuentemente, las bases mínimas también se llaman  *$\zeta$  sencilla* (SZ). Ya que las capas internas prácticamente no intervienen en los enlaces químicos, algunas bases usan un solo CGTO para representar los orbitales en esas capas, pero dos para los orbitales de las capas de valencia. Estas bases se denominan *bases escindidas de valencia* (*split valence* (SV)), o *doble  $\zeta$  de valencia* (VDZ).

Dado que la solución de las ecuaciones HF se obtiene por el método variacional, cuanto más extensa sea la base, tanto mejor puede ser la aproximación a la solución exacta; aunque no es necesariamente cierto que siempre bases más extensas produzcan mejores resultados en otros observables que no sean la energía total del sistema (*vide infra*, sección 2.4). Una forma conveniente de aumentar la flexibilidad de la base, es la inclusión de funciones de polarización; éstas corresponden a orbitales atómicos con números cuánticos azimutales mayores que aquéllos correspondientes a los orbitales atómicos ocupados en los estados base de los átomos. Un caso típico es la base *doble  $\zeta$  más polarización* (DZP) que incluye funciones *d* en átomos de los periodos primero y segundo, *e. g.*, en C ó Si, y funciones *p* en H. Su composición, de acuerdo a la notación común, es {6, 4, 1/4, 2, 1/2, 1}, donde se da el número de funciones contraídas para cada orbital *s*, *p*, *d*, y las especificaciones para átomos de diferentes renglones de la tabla periódica, se hacen en orden decreciente, separadas por líneas inclinadas. Actualmente se ha generalizado el uso de las bases sugeridas por el grupo de Pople [3]: STO-3G (MB), 3-21G y 4-31G (SV), 6-31G\* y 6-31G\*\* (SV aumentados con funciones de polarización; el primer asterisco indica un conjunto de funciones *d* en todos los átomos, salvo el hidrógeno, en tanto que el segundo asterisco se refiere a un conjunto de funciones *p* en los hidrógenos), y 6-311G\*\* (SZ en el carozo, TZ en la valencia y funciones de polarización en todos los átomos). Las bases STO-3G\* y 3-21G\* tienen conjuntos de funciones *d* añadidas sólo a los elementos del segundo renglón en adelante. En las notaciones tales como 3-21+G ó 6-31++G\*\* el primer signo "+" se refiere a un conjunto adicional de funciones *s* y *p* difusas para todos los átomos, salvo el hidrógeno; en tanto

que el segundo "+" se refiere a una gaussiana *s* difusa en el hidrógeno. En tanto que las funciones de polarización se asignan a orbitales desocupados en el estado base, las funciones difusas corresponden a orbitales ocupados. En la base STO-3G cada STO de una base mínima se representa por tres GTO's primitivos, mientras que en los conjuntos *n*-k1G el CGTO de cada orbital del carozo consta de *n* primitivos y los orbitales de la capa de valencia se escinden en un CGTO que consta de *k* primitivos y un solo conjunto de GTO.

Se pueden encontrar descripciones más detalladas de diversas bases disponibles, con comentarios específicos sobre su selección para problemas particulares en varias revisiones, como p. ej. en las referencias [4, 7 - 10].

Es posible lograr una reducción significativa del número de integrales bielectrónicas, en particular para los elementos de los periodos más altos, cuando en el proceso en estudio puede considerarse a los electrones del carozo y reemplazarlos por un potencial efectivo del carozo, o *pseudopotencial* (e. g. [11]), o *potencial modelo*. En tanto estos pseudopotenciales se hayan deducido en una forma no empírica, y no contengan parámetros ajustables, los métodos generalmente siguen siendo llamados *ab initio*; aunque contengan una aproximación adicional que requiera ser probada. Cabe notar que los pseudopotenciales deben usarse con bases especialmente adaptadas a ellos; por eso algunos autores no los consideran dentro de los cálculos *ab initio* propiamente dicho. Se pueden encontrar detalles en distintas revisiones (e. g. [11]).

### 2.3 Aplicación a la determinación de estructuras moleculares

Uno de los principales objetivos de la química cuántica computacional es la determinación de los parámetros estructurales (longitudes y ángulos de enlace, ángulos diedros) usando el criterio de mínima energía, esto es, que se considera una geometría molecular *optimizada* con una base dada, cuando el valor de la energía, calculado con esa base, sea mínimo. Ya que el valor que se obtenga para la energía depende de la base empleada, distintas bases pueden dar diferentes estructuras optimizadas, por lo que asegurar que se ha llegado a una geometría óptima requiere de un estudio más detallado (véase el capítulo 4).

Se han desarrollado una serie de métodos [12, 13] para localizar mínimos en la hipersuperficie de energía, con respecto a las coordenadas nucleares de una molécula. Tales mínimos corresponden a las estructuras moleculares de equilibrio. Estos métodos requieren del cálculo repetido de las primeras derivadas de la energía  $E_R = E(X_1, \dots, X_{3N})$  con respecto a las coordenadas de *N* núcleos:

$$g_i(X_1, \dots, X_{3N}) = \partial E(X_1, \dots, X_{3N}) / \partial X_i \quad (2 - 17).$$

En un principio había que evaluar numéricamente las coordenadas  $g_i$  del gradiente, aplicando las diferencias finitas entre valores de energía, calculados para desplazamientos pequeños  $\Delta$  de las coordenadas nucleares, a partir de la estructura de referencia  $\mathbf{X}^c$ , y de la energía de esta estructura:

$$g_i \approx [E(X_i = X_i^c + \Delta) - E(X_i = X_i^c)] / \Delta \quad (2 - 18).$$

Actualmente ya se han deducido fórmulas analíticas para el cálculo de los gradientes, e incluso para derivadas de órdenes mayores de la energía (ver las revisiones en [12, 14, 15] y sus referencias). Las ventajas de los métodos analíticos son una mayor precisión numérica y una mayor eficiencia de cómputo. Por ejemplo, el tiempo requerido para la evaluación analítica de todas las componentes del gradiente, no es mayor que dos o tres veces el tiempo necesario para calcular la energía SCF, por cada geometría. Para algunas bases, como p. ej. 3-21G, existen programas muy eficientes que usan incluso menos tiempo para el cálculo de los gradientes que para el cálculo de la energía SCF [13, 16, 17, 18]. En cambio, la derivación numérica (ec. 2-18) requiere del cómputo de  $3N - 6$  energías SCF. Para una molécula de 10 átomos, esto significa un tiempo de cómputo de 8 a 24 veces mayor que el necesario para los métodos de derivación analítica.

Las derivadas analíticas también facilitan el cálculo de los espectros de vibración [19, 20]. Las constantes de fuerza armónica  $f_{ij}$  se definen como las segundas derivadas de la energía, con respecto a las coordenadas cartesianas en la posición de equilibrio  $\mathbf{X}^e$ :

$$f_{ij} = \left. \frac{\partial^2 E(X_1, \dots, X_{3N})}{\partial X_i \partial X_j} \right|_{\mathbf{X}=\mathbf{X}^e} \quad (2-19),$$

y pueden obtenerse, ya sea directamente de fórmulas analíticas para las segundas derivadas (disponibles en los programas comerciales GASUSSIAN-86 [16] y versiones siguientes, CADPAC [21] y HONDO 6.5 [22]), o bien como diferencias finitas de gradientes analíticos:

$$f_{ij} \approx [g_i(X_j = X_j^e + \Delta) - g_i(X_j = X_j^e)]/\Delta \quad (2-20).$$

Mientras que los cálculos con gradientes analíticos casi siempre son factibles, en tanto que los recursos de cómputo disponibles permitan efectuar el cálculo de energía, esto no es cierto para los cálculos con segunda derivada analítica. Los tiempos de cómputo para una sola corrida, pueden resultar prohibitivamente largos y se necesita espacio adicional en disco.

#### 2.4 Aplicación a la descripción de las reacciones químicas

Ya que se han determinado las estructuras óptimas (i. e. de mínima energía) de una serie de moléculas, los cálculos cuánticos proporcionan la energía de cada una de ellas. Entonces se hace posible una estimación del calor de una reacción química. En principio, cuando ya se ha optimizado la estructura de cada molécula participante en la reacción



se cuenta con los valores de las energías  $E(A)$ ,  $E(B)$ ,  $E(C)$  y  $E(D)$ . El calor de la reacción es simplemente la diferencia

$$\Delta E = E(C) + E(D) - E(A) - E(B) \quad (2-22).$$

En este procedimiento no se ha considerado ninguna interacción entre los reactantes y los productos, de forma que la situación corresponde a la de un gas infinitamente diluido.

Tampoco se han tomado en cuenta rotaciones, vibraciones ni traslaciones, por lo que se está manejando una reacción a temperatura  $T = 0$  K. Además, comúnmente tanto la optimización como el cálculo de las energías, se hace al nivel HF - SCF, i. e. sin incluir la contribución de la correlación electrónica. Todavía hay que mencionar un problema más: en tanto que la energía de cada molécula se ha calculado con un método variacional, que garantiza que el valor obtenido es cota superior de la energía HF exacta, y que al aumentar el tamaño de la base tanto más cerca se está de ella, el calor  $\Delta E$  de la reacción (ec. 2-21) se obtiene de una diferencia pequeña (típicamente del orden de milihartrees) entre números significativamente mayores (desde decenas hasta centenas de hartrees). Entonces el cálculo de  $\Delta E$  no cumple con los teoremas variacionales y, por lo tanto, no es necesariamente cierto que bases mayores produzcan resultados más confiables, en tanto no hayan rebasado cierto tamaño (típicamente el nivel DZP es el comienzo de la convergencia hacia el límite HF [4]). La Figura 2-1 muestra esquemáticamente esta situación.

FIGURA 2-1

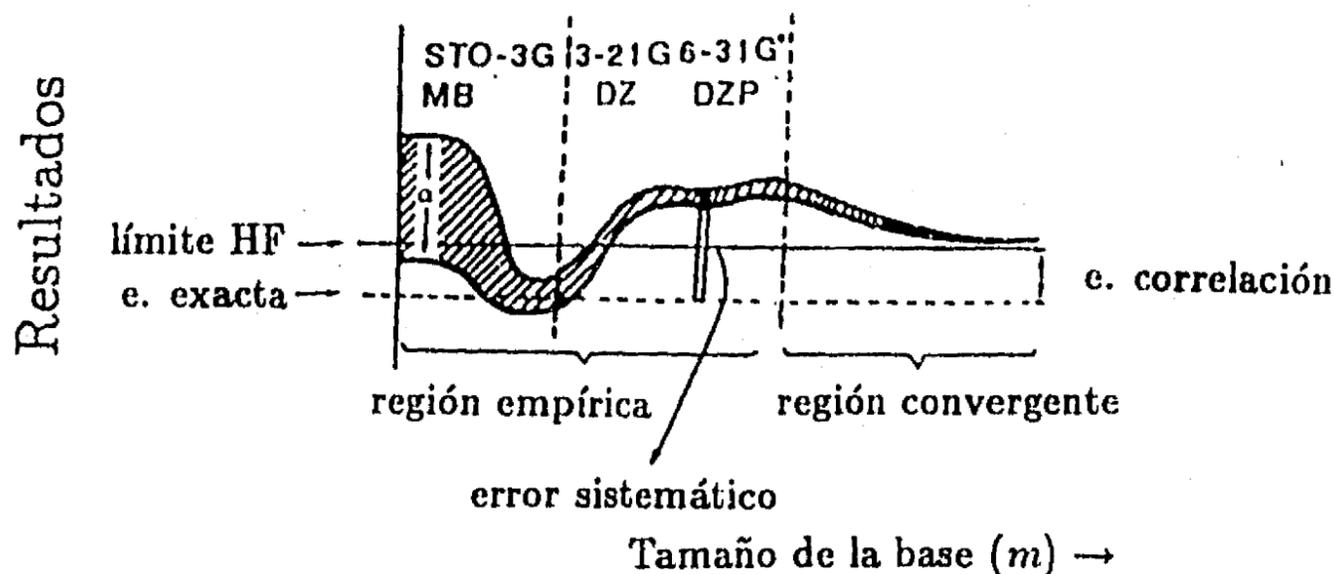


Figura 2-1. Aquí se muestra en forma esquemática cómo varía el error en los resultados de los cálculos *ab initio* a nivel HF - SCF, al aumentar el número de funciones en la base,  $m$ . (Hay que recordar que para cualquier base, la energía del sistema queda arriba del límite de Hartree - Fock).

Ante todos estos obstáculos, parece muy osado pretender que se pueden obtener valores confiables para calores de reacciones a partir de cálculos *ab initio*. Resulta sorprendente que existan, en efecto, resultados de estos cálculos que son comparables a los datos experimentales, aún cuando se hayan usado bases pequeñas [23]. Esto suele ocurrir gracias a una cancelación de errores que, en muchos casos, no tiene un comportamiento sistemático. Una racionalización de este efecto, es que aun cuando las bases usadas al nivel SCF describen cada molécula en forma incorrecta, el error es igual para todas, de modo que al calcular las diferencias ocurre la cancelación mencionada. Esto es cierto en el cálculo de energías de interacción intermolecular (e. g. [24]), y en este caso se habla de *bases bien balanceadas*.

Cuando se usan bases chicas entran en juego la habilidad y la experiencia (el *callo*) de la persona que realiza los cálculos, aunque existen algunas metodologías para calibrar bases, comparando sus resultados con los de bases mayores [24]. En este sentido, las bases que se han diseñado para cálculos con una sola molécula, p. ej. las que usan funciones difusas en orbitales ocupados, no pueden aplicarse confiablemente a estimar cantidades que impliquen diferencias entre distintas especies moleculares, ya que seguramente no estarán bien balanceadas y no habrá una cancelación sistemática de errores.

En lo que resta de este capítulo se hará una discusión de las formas de superar los problemas mencionados en los párrafos anteriores, comenzando con el referente al tamaño de la base usada.

Aquí no se hará un análisis general de la confiabilidad de distintas bases que se usan en los cálculos SCF, lo cual puede ser encontrado en diversas revisiones [4, 10, 25]. Sin embargo, cabe mencionar el rendimiento de algunas bases con respecto a la optimización de geometrías y al cálculo de calores de reacción.

Comúnmente las bases mínimas tienen asociadas incertidumbres grandes, en cuanto a sus predicciones de las longitudes y los ángulos de enlace, pero su comportamiento varía con el tipo de molécula estudiada. Así, mientras que para moléculas con átomos del primer renglón la base STO-3G da buenos resultados, éstos empeoran sensiblemente conforme se trata con elementos más pesados. Las bases MINI-1 de Huzinaga [26 - 29] producen longitudes de enlace que sistemáticamente están sobrestimadas entre 4.5% y 7.5% para enlaces A - H, y entre 5% y 10% para los demás enlaces. Sin embargo, el cálculo de los ángulos de enlace es muy bueno, e incluso hay alguna evidencia [26 - 31] de que sus predicciones sobre las frecuencias de vibración pueden competir con la base 3-21G. En general puede decirse que se obtienen geometrías y frecuencias de vibración correctas al usar bases DZP, pero también es común encontrar longitudes de enlace muy cortas y frecuencias de vibración muy altas al usar bases con funciones de polarización al nivel SCF.

En cuanto a la precisión en el cálculo de los calores de reacción, es conveniente hacer primero una clasificación de los distintos tipos de reacciones desde el punto de vista de la química cuántica computacional.

## **2.5 Clasificación de las reacciones químicas, con base en la energía de correlación**

### **2.5.1 La cancelación de la energía de correlación**

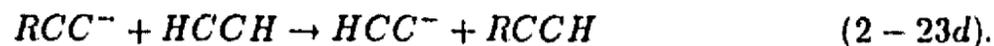
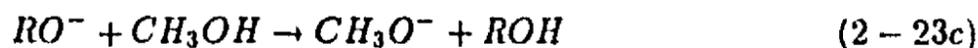
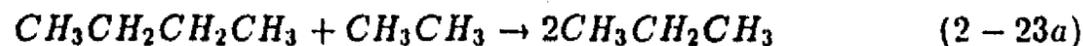
La clasificación que se presenta a continuación se basa en la relevancia que puede tener la energía de correlación en las diversas situaciones bajo estudio. Aunque se demuestra teóricamente que la aproximación HF - SCF es un tanto burda al no considerar la energía de correlación [23], el hecho de que en muchos cálculos prácticos de los calores de reacción ésta pueda soslayarse, se debe a que en la diferencia de la ec. (2-21) muchas veces los efectos de la correlación se cancelan. A esta cancelación Čársky la denomina la *conservación de la energía de correlación*, en referencia a que es una cantidad que permanece igual de ambos lados de una ecuación química. Las condiciones para que ocurra la cancelación, son las

siguientes [32]: (1) el número de pares electrónicos debe conservarse, y (2) para pares electrónicos que sean vecinos inmediatos también debe conservarse el arreglo espacial.

La primera condición la satisfacen automáticamente todas las reacciones en que participen sólo moléculas con capas cerradas. Snyder y Basch [33, 34] realizaron un estudio sistemático de este tipo de reacciones, y concluyeron que los calores de reacción obtenidos con métodos *ab initio* al nivel SCF, son más precisos que aquéllos que se encuentran usando relaciones semiempíricas de las energías de enlace para moléculas deformadas. Sin embargo, los mismos autores también concluyen que si se quiere llegar al nivel de precisión del experimento, se necesita una estimación semicuantitativa del cambio en la energía de correlación. Por supuesto, la conservación de la energía de correlación, depende del grado en que se cumple la segunda condición mencionada arriba. Con respecto a esta condición, y siempre considerando moléculas de capa cerrada, se ha hecho la siguiente clasificación [35]:

### 2.5.2 Reacciones homodésmicas

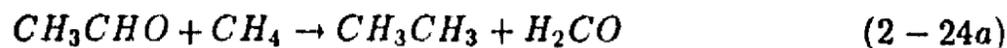
En las reacciones homodésmicas [36], se conserva el número de enlaces de un tipo particular (*e. g.* C[4]-C[4], C[4]-C[3], C[3]-C[3], C[3]=C[3], donde el número entre paréntesis indica la cantidad total de otros átomos ligados a cada carbono, de modo que C[4]-C[4] y C[3]=C[3] pueden representar enlaces CC en etano y en etileno, respectivamente). Además también se conservan cantidades iguales de cada tipo atómico (como C[4], C[3], etc.) con cero, uno dos y tres átomos de hidrógeno asociados en los reactantes y en los productos. Algunos ejemplos de reacciones homodésmicas son [36 - 38]:



En estas ecuaciones, R denota un alkilo. En las reacciones homodésmicas los elementos estructurales en los reactantes y en los productos son muy similares, de modo que puede esperarse un cambio muy pequeño en la energía de correlación. Así, cálculos con bases DZP dan errores de entre 1 y 2 kcal/mol con respecto a la medida experimental [36].

### 2.5.3 Reacciones isodésmicas

En las reacciones isodésmicas [39], se retiene el número de enlaces de cada tipo formal dado, pero con un cambio en las relaciones de uno con otro. En este sentido, las reacciones homodésmicas son una subclase de las isodésmicas. Los ejemplos típicos de este tipo de reacción son las de separación de enlace [38], como:



El grado de conservación de las relaciones espaciales depende de cada caso particular. El error intrínseco de la aproximación HF - SCF, en el cálculo de los calores de las reacciones isodésmicas, puede crecer a más de 5 kcal/mol, dependiendo de la base usada. La reacción de hidrólisis del pirofosfato neutro es de este tipo, por lo que en esta tesis se hace una discusión particular al respecto (secciones 4.4 y 4.5).

#### 2.5.4 Reacciones anisodésmicas

Desgraciadamente la mayor parte de las reacciones químicas pertenecen a esta última categoría en que la energía de correlación definitivamente no se conserva. A pesar de ello, una forma de tratar con estas reacciones, es la de combinar reacciones homodésmicas e isodésmicas en ciclos termoquímicos para describir estos procesos anisodésmicos [40, 41]. Un ejemplo de reacción anisodésmica con un cambio grande en la energía de correlación, es la dimerización  $2BH_3 \rightarrow B_2H_6$ , en que el cálculo SCF da una energía de 21 kcal/mol, en tanto que la inclusión de la correlación, lleva esta energía a 38 kcal/mol [42]. Aunque el número de pares electrónicos fue preservado, su arreglo espacial y el número de interacciones par - par no lo fueron.

Por otra parte, la reacción  $NH_3 + HCl \rightarrow NH_4Cl$  también puede tomarse como anisodésmica y, sin embargo, se satisfacen las dos condiciones para la conservación de la energía de correlación; de modo que se pudo predecir la existencia de  $NH_4Cl$  en la fase gaseosa usando cálculos SCF [43], la cual fue corroborada experimentalmente [44].

Algunas reacciones con moléculas de capa abierta también pueden ser tratadas al nivel SCF, pero requieren del uso de bases extendidas (al menos DZP [4]).

#### 2.6 Aproximación al límite de Hartree - Fock

En cada caso particular la convergencia al límite de Hartree - Fock depende de la base empleada. Para juzgar qué tan cerca de ese límite están los resultados, es conveniente hacer un estudio sistemático de esta convergencia, usando varias bases con un número creciente de funciones y comparando los resultados de unas bases con los de otras. Como lo representa la Figura 2-1, hay un cierto intervalo en que los resultados son un tanto caóticos, pero al rebasar el nivel DZP ya se llega a la región convergente.

Con el fin de hacer una comparación con los datos experimentales, sería adecuado contar con unas *barras de error* alrededor de los valores obtenidos. En general no se puede hacer una estimación cuantitativa del error; sino sólo asegurar que se ha llegado a un *mejor* cálculo. Sin embargo, se puede proponer una aproximación sistemática a un problema dado, con la cual sí se llegue a una estimación del error. En el caso del cálculo del calor de una reacción química, primero se hacen las optimizaciones de las estructuras de los reactantes y de los productos con distintas bases, con el fin de estudiar cuánto cambian los parámetros estructurales en función de la base usada (convergencia de geometrías). Desde aquí se comienzan los cálculos de la energía de la reacción de interés. Comúnmente al usar bases mayores que DZP, el resultado de esta cantidad ya no depende críticamente de las geometrías usadas, en tanto éstas sean *buenas*. Es decir: que se encuentren en cierto intervalo cercano a las geometrías óptimas. Entonces una tabla de

valores calculados para la energía de la reacción con distintas bases y geometrías, da una idea de la convergencia lograda; como es el caso en la Tabla II de la sección 4.5. Así puede hacerse una estimación del error introducido por usar una geometría no óptima y una base no saturada, al considerar las diferencias entre valores de columnas y renglones sucesivos (véase la sección 4.4).

## 2.7 La correlación electrónica

### 2.7.1 Los métodos de interacción de configuraciones CI

La precisión obtenida con el procedimiento arriba mencionado se refiere al límite HF. Pero como se explicó anteriormente, esta aproximación tiene el error intrínseco de no considerar los efectos de la correlación electrónica, que puede resultar muy importante en algunos casos.

Al usar una función de onda antisimétrica, aun la aproximación HF ya tiene incluido cierto grado de correlación, debido al principio de exclusión de Pauli: dos electrones con el mismo spin no ocupan el mismo orbital espacial. Pero existen otros efectos de correlación que se discutirán a continuación.

Los efectos de la correlación se incluyen con los métodos de interacción de configuraciones (CI) [45]. En ellos la función de onda del sistema se construye como una superposición del determinante HF,

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \dots & \psi_n(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \dots & \psi_n(\mathbf{r}_2) \\ \dots & \dots & \dots & \dots \\ \psi_1(\mathbf{r}_n) & \psi_2(\mathbf{r}_n) & \dots & \psi_n(\mathbf{r}_n) \end{vmatrix} \quad (2-2),$$

con todos aquéllos determinantes  $\psi_{ijk\dots}^{abc\dots}$  que se obtengan al sustituir un número creciente de orbitales moleculares ocupados  $i, j, k, \dots$  por orbitales desocupados (virtuales)  $a, b, c, \dots$ :

$$\psi = \psi^{HF} + \psi_S + \psi_D + \psi_T + \psi_Q + \dots \quad (2-25),$$

en que  $S, D, T, Q$  indican superposiciones de todos los determinantes de un tipo dado, i. e., con sustituciones  $S$  sencillas,  $D$  dobles,  $T$  triples,  $Q$  cuádruples, etc., como p. ej.:

$$\psi_D = \sum_{i \leq j}^{oc} \sum_{a \leq b}^{des} C_{ij}^{ab} \psi_{ij}^{ab} \quad (2-26),$$

donde

$$\psi_{ij}^{ab}(\mathbf{r}_1, \dots, \mathbf{r}_n) = |\psi_1(\mathbf{r}_1), \dots, \psi_a(\mathbf{r}_i), \dots, \psi_b(\mathbf{r}_j), \dots, \psi_n(\mathbf{r}_n)| \quad (2-27).$$

Aunque en principio siempre puede aproximarse la solución exacta de la ecuación de Schrödinger al tomar en cuenta *todas las sustituciones posibles*, este CI total (full CI)

no es factible para sistemas con más de una decena de electrones ni para bases de calidad mayor a la DZP, aun con el empleo de las supercomputadoras [46, 47].

Para la química cuántica computacional actual, es de la mayor relevancia el encontrar métodos eficaces y eficientes para tratar la correlación electrónica en una forma aproximada. Todos los métodos desarrollados parten de alguna idea de cómo acortar la serie en (2-25) y cómo confinarla a sólo cierto tipo de sustituciones. La inclusión de triples  $\psi_T$  está en la frontera actual, aunque existen algunas aplicaciones que incluyen cuádruples  $\psi_Q$ , afectando a las energías calculadas en algunos microhartrees [48].

Existe una gran variedad de métodos en uso, y aquí no se intenta hacer una presentación de todos ellos, por lo que la siguiente discusión se limitará a la mención de algunos, en especial el de perturbaciones de Møller - Plesset, que se utilizó en los cálculos de la entalpía de la hidrólisis del pirofosfato. Antes de ello, resulta útil distinguir entre los efectos dinámicos de correlación, y los *no dinámicos*, ya que requieren de ataques diferentes.

### 2.7.2 Efectos no dinámicos de correlación.

El ansatz a partir de un solo determinante HF, origina una representación cualitativamente incorrecta de la estructura electrónica. Debido a la cuasi - degeneración energética entre dos o varios determinantes, se necesita una superposición de los mismos para obtener una aproximación satisfactoria a orden cero. Un ejemplo de libro de texto sobre las limitaciones de la aproximación HF, lo constituye el límite de disociación que se obtiene del ansatz de un solo determinante para la fisión homolítica de los enlaces químicos [45], que resulta equivocado: considérese, p. ej., una molécula de  $H_2$  con los orbitales atómicos  $\chi_a$  y  $\chi_b$ , centrados en los núcleos  $H_a$  y  $H_b$ , respectivamente. En la aproximación HF de capa cerrada, ambos electrones se describen con el mismo orbital molecular  $\psi_1$  y difieren sólo en sus funciones de spin,  $\alpha$  y  $\beta$  (donde  $\sigma$  es la coordenada de spin):

$$\Psi_1(\mathbf{r}_1, \mathbf{r}_2) = |\psi_1(\mathbf{r}_1)\alpha(\sigma_1)\psi_1(\mathbf{r}_2)\beta(\sigma_2)| \quad (2-28).$$

Los coeficientes del desarrollo en orbitales son, por simetría, (cf. ec. 2-6)

$$\psi_1(\mathbf{r}) = c_1(\chi_a(\mathbf{r}) + \chi_b(\mathbf{r})) \quad (2-29).$$

La sustitución de (2-29) en (2-28) da por resultado (sin incluir las partes de spin ni las constantes de normalización):

$$\Psi_1(\mathbf{r}_1, \mathbf{r}_2) = \chi_a(\mathbf{r}_1)\chi_a(\mathbf{r}_2) + \chi_b(\mathbf{r}_1)\chi_b(\mathbf{r}_2) + \chi_a(\mathbf{r}_1)\chi_b(\mathbf{r}_2) + \chi_b(\mathbf{r}_1)\chi_a(\mathbf{r}_2) = \Phi^{ion} + \Phi^{cov} \quad (2-30).$$

O sea que la función de onda HF da iguales pesos a situaciones que pueden describirse como iónicas

$$\Phi^{ion} = \frac{H_a^- H_b^+ + H_a^+ H_b^-}{2} = \chi_a(\mathbf{r}_1)\chi_a(\mathbf{r}_2) + \chi_b(\mathbf{r}_1)\chi_b(\mathbf{r}_2) \quad (2-31)$$

y como covalentes

$$\Phi^{cov} = \frac{H_a H_b + H_b H_a}{2} = \chi_a(\mathbf{r}_1)\chi_b(\mathbf{r}_2) + \chi_b(\mathbf{r}_1)\chi_a(\mathbf{r}_2) \quad (2-32).$$

En tanto que esto parece ser una representación razonable de un enlace químico cerca de su distancia de equilibrio, falla en cambio para dos átomos de H separados por una distancia grande,  $d$ , los cuales quedan descritos por  $\Phi^{cov}$  a secas:

$$\dot{H}_a \dots d \dots \dot{H}_b$$

Ya que a separación infinita la energía de (2-31) es mayor que la de (2-32) (por la suma del potencial de ionización y de la afinidad electrónica), el ansatz de un determinante HF sobrestima la energía de la disociación, y entonces la curva de potencial termina en un estado incorrecto.

Hay otra solución al problema HF (cfr. cc. 2-7), que también se obtiene por simetría:

$$\psi_2(\mathbf{r}) = c_2(\chi_a(\mathbf{r}) - \chi_b(\mathbf{r})) \quad (2-33).$$

Este orbital de antienlace puede usarse para construir un determinante

$$\Psi_2(\mathbf{r}_1, \mathbf{r}_2) = |\psi_2(\mathbf{r}_1)\alpha(\sigma_1)\psi_2(\mathbf{r}_2)\beta(\sigma_2)| \quad (2-34),$$

que está doblemente sustituido con respecto a  $\psi_1$ . Al sustituir (2-33) en (2-34) se llega a

$$\Psi_2 = \Phi^{ion} - \Phi^{cov} \quad (2-35).$$

A largas distancias,  $\psi_2$  se degenera energéticamente con  $\psi_1$  y el estado base apropiado,  $\Phi^{cov}$ , se obtiene de

$$\Phi^{cov} = \Psi_1 - \Psi_2 \quad (2-36).$$

Para distancias arbitrarias, la combinación óptima de las contribuciones iónica y covalente, se pueden encontrar a partir del ansatz de dos configuraciones:

$$\Psi = C_1\Psi_1 + C_2\Psi_2 = C^{cov}\Phi^{cov} + C^{ion}\Phi^{ion} \quad (2-37).$$

En vez de buscar los coeficiente óptimos  $C_1$  y  $C_2$  en (2-37), de forma equivalente se puede efectuar el ansatz de dos configuraciones, conocido como método de valencia - enlace generalizado (*generalized valence - bond GVB*) [49]:

$$\Psi^{GVB}(\mathbf{r}_1, \mathbf{r}_2) = |\phi_1(\mathbf{r}_1)\alpha(\sigma_1)\phi_2(\mathbf{r}_2)\beta(\sigma_2)| + |\phi_2(\mathbf{r}_1)\alpha(\sigma_1)\phi_1(\mathbf{r}_2)\beta(\sigma_2)| \quad (2-38).$$

Y ahora se trata de obtener los coeficientes óptimos para los orbitales:

$$\phi_1(\mathbf{r}) = c_{1a}\chi_a(\mathbf{r}) + c_{1b}\chi_b(\mathbf{r}) \quad (2-39a)$$

$$\phi_2(\mathbf{r}) = c_{2a}\chi_a(\mathbf{r}) + c_{2b}\chi_b(\mathbf{r}) \quad (2-39b).$$

Entonces  $\Psi^{GVB}$  incluye a  $\Phi^{cov}$  y a  $\psi_1$  como casos especiales. A distancia infinita los orbitales óptimos son simplemente  $\phi_1 = \chi_a$  y  $\phi_2 = \chi_b$ , y  $\Psi^{GVB} \rightarrow \Phi^{cov}$ . Forzando la

igualdad  $\phi_1 = \phi_2$  se llega a que  $\Psi^{GVB} = \Psi_1$ . La aplicación del principio variacional a (2-38) y a (2-39) lleva a ecuaciones del tipo HF, y los problemas de cómputo son parecidos a los que se describieron en la sección 2.1.

Como lo muestra el ejemplo de  $H_2$ , los cálculos a nivel HF - SCF no permiten obtener superficies de potencial para procesos que involucren la ruptura o la formación de enlaces químicos. Para tales propósitos, los métodos como el GVB, y otros SCF multiconfiguracionales (MCSCF), son la elección adecuada. En tanto el interés resida sólo en una descripción cualitativamente correcta, se pueden usar bases pequeñas con polarización, como DZP, o incluso algunas escindidas de valencia sin polarización, como la 3-21G. Al combinarlas con técnicas de gradientes y métodos de búsqueda de puntos estacionarios en hipersuperficies, los cálculos MCSCF con bases pequeñas, han resultado eficaces en estudios cualitativos de superficies de potencial de disociación [50]. Aquí hay que distinguir las superficies de potencial de interacción intermolecular, en las que no se incluyen rupturas ni formaciones de enlaces químicos, y para las cuales el método SCF con un solo determinante, sí puede dar resultados razonables en una aproximación de orden cero; aunque en algunos casos la inclusión de la energía de correlación sea necesaria, como se explica en el siguiente párrafo.

### 2.7.3 Efectos dinámicos de correlación.

Para estos efectos, el determinante HF es una buena aproximación a orden cero. Este tipo de correlación produce correcciones pequeñas a los resultados HF de las geometrías moleculares, de las constantes de fuerza y de las propiedades moleculares; o a las energías de excitación y de ionización. Un efecto dinámico de correlación típico es la energía de dispersión, que es una contribución a las interacciones intermoleculares que no se obtiene del nivel HF. Aunque la energía de dispersión es muy pequeña, comparada con la energía total de los subsistemas interactuantes, es ella la responsable de la cohesión de van der Waals entre los átomos de los gases nobles y entre las moléculas no polares; y también contribuye significativamente a otros tipos de interacción intermolecular, como p. ej. los puentes de hidrógeno. Por lo tanto, no puede pretenderse que los cálculos HF produzcan energías de estabilización confiables para los gases nobles ni para los cristales, ya sean iónicos o moleculares.

Existen métodos económicos disponibles para tratar con los efectos dinámicos de correlación [51, 52], que se basan en que las contribuciones de los determinantes con sustituciones dobles, son las dominantes. Actualmente se cuenta con los métodos siguientes: la interacción de configuraciones (sencillas y) dobles, CI(S)D; los funcionales de pares acoplados y la aproximación por pares electrónicos acoplados, CPF y CEPA, respectivamente [53]; las aproximaciones de cúmulos acoplados, CCD y CCSD [54], y la teoría de perturbaciones de Møller - Plesset [55]. Esta última incluye sólo sustituciones dobles hasta los órdenes segundo y tercero (MP2, MP3), mientras que para el cuarto orden pueden incluirse sencillas, triples y cuádruples (SDTQ-MP4). Los métodos más avanzados parten de una función de onda multideterminante, construida con orbitales MCSCF; y consideran todos los determinantes que tengan sustituciones sencillas o dobles con respecto a todos los determinantes en la función de onda multideterminante. Esto lleva al esquema general MCSCF - MRCI (interacción de configuraciones multireferencial), que puede tratar corre-

lación tanto dinámica como no dinámica y es casi equivalente al CI total (full), con un costo computacional equivalente, sobre todo por la parte MCSCF. Existen, empero, algunos métodos aproximados que han ganado importancia práctica: GVB - CI [56], CASSCF - CI [57 - 59], MRD - CI [60, 61].

Es importante mencionar que la descripción de los efectos dinámicos de correlación, requiere bases mucho mayores que la simple aproximación SCF. De hecho, las funciones de polarización son indispensables. En contraste con los cálculos multiconfiguracionales con finalidades cualitativas, aquellos cálculos que pretenden incluir los efectos dinámicos de correlación, pero que emplean bases sin polarización, no tienen ningún sentido.

A pesar del costo computacional que representa considerar la correlación electrónica, los cálculos a este nivel se están volviendo rutinarios para la descripción de moléculas aisladas. La aproximación más simple al efecto dinámico de la correlación sobre la energía,  $E_{corr}^{(2)}$ , está dada por la teoría de perturbaciones de Møller - Plesset hasta el segundo orden (MP2) [55]:

$$E_{corr}^{(2)} = \sum_{i,j} \frac{\langle ia|jb \rangle [2\langle ia|jb \rangle - \langle ib|ja \rangle]}{E_i + E_j - E_a - E_b} \quad (2-40).$$

La expresión se refiere a un estado base de capa cerrada, en el que  $i$  y  $j$  son orbitales doblemente ocupados. El paso que mayor tiempo de cómputo consume es la transformación de las integrales sobre funciones en la base,  $\langle \mu\nu|\lambda\sigma \rangle$  (ec. 2-14), a integrales sobre los orbitales moleculares,  $\langle ia|jb \rangle$ :

$$\langle ia|jb \rangle = \sum_{\substack{\mu\nu \\ \lambda\sigma}} c_{i\mu} c_{a\nu} c_{j\lambda} c_{b\sigma} \langle \mu\nu|\lambda\sigma \rangle \quad (2-41).$$

Existen programas eficientes para el cálculo de la energía al nivel MP2, en un tiempo similar al cálculo de la energía SCF (e. g. GAUSSIAN 86 a 90 [16 - 18], CIPSI [62]), y aún para efectuar la optimización de las geometrías [16 - 18, 21, 63]. Con el uso de bases extensas y la aproximación al nivel MP2, se pueden obtener frecuencias armónicas de vibración muy precisas [62], así como momentos eléctricos multipolares y energías de interacción muy confiables. Incluso se pueden obtener porciones importantes de la energía de dispersión con el método MP2 [64 - 69].

En cuanto a los cálculos moleculares, el nivel MP2 está tomando el lugar que hasta hace poco tiempo tenía el nivel SCF como el método estándar para moléculas relativamente pequeñas ( $\sim 10$  átomos). A esto contribuye la ventaja de que si el método MP2 falla, se sabe cómo mejorar los resultados: llevando las sustituciones dobles a órdenes mayores, incluyendo otras sustituciones (S, T y Q), o cambiando a un tratamiento multireferencial. El costo computacional del método MP2 es todavía muy alto para moléculas poliatómicas, debido a la necesidad de usar bases extensas. Este costo también limita el uso del MP2 para el cálculo de superficies de interacción, debido al gran número ( $> 100$ ) de energías que deben obtenerse.

Aquí cabe mencionar que existe un tratamiento alternativo de los cálculos moleculares, que no se va por el camino HF - SCF  $\rightarrow$  CI; sino que utiliza el hecho de que la energía

de un sistema multielectrónico, es una funcional única de la densidad electrónica  $\rho(\mathbf{r})$  [70, 71]. Este es el tratamiento por los *funcionales de la densidad*, que no ha sido empleado en este trabajo y por lo tanto no se discutirá más. El lector interesado puede recurrir a una serie de revisiones sobre el tema, e. g., [72 - 74].

## 2.8 La energía de punto cero (ZPE) y los efectos térmicos.

En los párrafos anteriores se ha discutido la precisión de los cálculos cuánticos y se ha concluido que se pueden obtener resultados confiables al usar bases extensas y al incluir la correlación electrónica hasta el nivel MP2. En el caso de una reacción química, estos resultados se refieren a la diferencia de energías en la ec. (2-22),

$$\Delta E = E(C) + E(D) - E(A) - E(B) \quad (2-22).$$

Como se trata de una situación a  $T = 0$  K (*vide supra*, sección 2.4), esta energía de la reacción  $\Delta E$  debería ser comparable a una medida experimental de  $\Delta U_0$ , ya que se trata del cambio en la energía interna (intramolecular) al pasar del sistema de reactantes al sistema de productos. Pero aquí se ha considerado a los núcleos de cada molécula, fijos en el espacio. De manera que  $\Delta E$  no incluye las pequeñas vibraciones de los núcleos alrededor de sus posiciones de equilibrio. Estas vibraciones también deben ser tratadas cuánticamente y, por lo tanto, tienen un estado mínimo de energía, cuyo valor es lo que se conoce como la *energía del punto cero* (ZPE, de acuerdo a sus siglas en inglés) [75]. En una primera aproximación, estas vibraciones pueden ser consideradas armónicas, y entonces la ZPE del sistema se calcula con las constantes de fuerza armónica obtenidas de las segundas derivadas de la energía con respecto a las coordenadas cartesianas de los núcleos en las posiciones de equilibrio  $\mathbf{X}^e$ :

$$f_{ij} = \left. \frac{\partial^2 E(X_1, \dots, X_{3N})}{\partial X_i \partial X_j} \right|_{\mathbf{X}=\mathbf{X}^e} \quad (2-19),$$

y usando las expresiones para la energía y la frecuencia de un oscilador armónico,

$$\begin{aligned} E_{0ij} &= \frac{1}{2} h \omega_{ij}, \\ \omega_{ij} &= (f_{ij}/m)^{1/2} \end{aligned} \quad (2-42),$$

donde  $m$  es la masa de cada núcleo [75, 76].

Con la inclusión de la ZPE ya se tiene una energía de reacción comparable a  $\Delta U_0$ , pero en fisicoquímica suelen emplearse otro tipo de cantidades, con un significado termodinámico bien definido (véase el apéndice A). Estas cantidades son la entalpía  $\Delta H$ , la entropía  $\Delta S$  y la energía libre de Gibbs  $\Delta G$  de la reacción. El cambio de la energía interna  $\Delta U$  de un sistema, al pasar de un estado a otro, puede medirse como la cantidad de calor añadida (o desprendida) al mantener el sistema con un volumen constante y sin efectuar ningún trabajo,

$$\Delta U = q_v \quad (2-43).$$

Cuando el proceso es isobárico, este calor es una medida de la variación de la entalpía,

$$\Delta H = q_p \quad (2 - 44).$$

Cuando una reacción química ocurre a presión y volumen constantes, las variaciones en energía interna,  $\Delta U$ , y en entalpía,  $\Delta H$ , son iguales. Además, el cambio en entalpía  $\Delta H$  en una reacción química, puede visualizarse como el cambio en las energías de interacción entre los núcleos y los electrones al pasar de los reactantes a los productos (ver, e. g. [76]). Entonces la energía de reacción  $\Delta E$ , obtenida de los cálculos cuánticos, es el cambio de entalpía en la reacción.

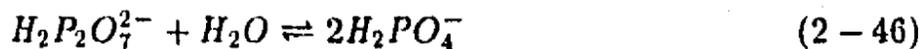
Por otra parte, en la bioquímica, comúnmente es más útil hablar de la energía libre de Gibbs  $\Delta G$  de una reacción (ver el capítulo 1), en la cual se toma en cuenta además la contribución del término entrópico en el sistema,  $T\Delta S$ . En cambio, como se mencionó en la sección 2.4, el cálculo *ab initio* representa un gas infinitamente diluido y a temperatura  $T = 0$  K. Una forma de considerar los efectos térmicos sobre las vibraciones y las rotaciones de las moléculas, consiste en el cálculo de las *funciones de partición molecular* (que se discuten con más detalle en el capítulo siguiente, por tratarse de conceptos mecánico-estadísticos). Así se toma en cuenta la contribución entrópica de cada una de las moléculas a una temperatura dada; pero no así los rearrreglos de unas moléculas con respecto a otras.

## 2.9 Cálculos cuánticos de la reacción de hidrólisis del pirofosfato

Casi una década después de la aparición del artículo de George *et al.* [77], Hayes *et al.* [78] realizaron una serie de cálculos *ab initio* de diversos ésteres y anhídridos de fosfato, tanto iónicos como neutros, en un intento por proporcionar valores confiables de los calores de las reacciones de hidrólisis en fase gaseosa. Dadas las capacidades de cómputo de la época, la optimización de geometrías que realizaron no es muy refinada. Además, utilizaron bases pequeñas (STO-3G y 4-31G), que son insuficientes para una buena descripción de los enlaces en el puente P - O - P. Finalmente, hacen una estimación de las energías de hidratación de los ácidos orto- y pirofosfórico con un modelo de campo promedio, y llegan a conclusiones similares a las de George *et al.* [77] con respecto a la importancia de la hidratación. Sin embargo, el cálculo de la energía  $\Delta E$  de la reacción



es de  $\Delta E = -4.5 \text{ kcal/mol}$  con una base 4-31G, que indica que la reacción es exotérmica en la fase gaseosa. Más aún, para la reacción



obtienen una energía de  $\Delta E = -70 \text{ kcal/mol}$ , que es diez veces mayor que el resultado experimental sobre la hidrólisis en solución acuosa. Aunque seguramente el valor es muy poco confiable, ya que requiere del cálculo de un anión divalente y las bases empleadas no lo describen correctamente; de cualquier forma es demasiado grande como para pasarlo por alto. Hayes *et al.* [78] presentan este resultado y, luego del cálculo de las energías de

hidratación, concluyen que el agua ayuda a estabilizar al anión divalente de pirofosfato, o sea que la energía de hidratación contribuye a hacer la reacción mucho menos energética que en la fase gaseosa, al contrario de lo propuesto en la hipótesis de George *et al.* [77]. Ya que no se cuenta ni con datos experimentales ni con otros cálculos *ab initio* de esta especie, la cuestión sobre su hidrólisis en fase gaseosa sigue abierta.

Existen otros dos trabajos sobre cálculos *ab initio* de la reacción de hidrólisis del pirofosfato. Uno de ellos, el de O'Keeffe *et al.* [79], consiste más bien en un estudio sobre la estructura molecular de varios fosfatos y su comparación con datos cristalográficos y con moléculas de silicatos. Ellos emplean bases grandes (6-31G\* y 6-31G\*\*) y comparan sus geometrías con las obtenidas con bases pequeñas (STO-3G\*), de donde concluyen que para moléculas con más de 10 átomos, las estructuras calculadas con bases chicas no son confiables, aunque pueden producir valores para ciertos parámetros (*e. g.* el ángulo del puente P - O - P) más cercanos a los datos cristalográficos que algunas bases medianas (véase la discusión sobre la optimización de las geometrías en la sección 2.3 y en el capítulo 4). Uno de los resultados más relevantes a los que llegan O'Keeffe *et al.* con respecto a la estructura del  $H_4P_2O_7$  es la correlación que existe entre las longitudes y el ángulo de enlace en el puente P - O - P. Esta correlación se obtiene a partir del cálculo de las constantes de fuerza  $K_r = \partial^2 E / \partial r^2$ ,  $K_\theta = \partial^2 E / \partial \theta^2$  y  $K_{r\theta} = \partial^2 E / \partial r \partial \theta$ , con la relación  $\delta r / \delta \theta = -K_{r\theta} / K_r$ .

Además de las estructuras moleculares, O'Keeffe *et al.* calculan la energía de la reacción (2-45), obteniendo un valor de  $\Delta E = -5.5 \text{ kcal/mol}$ . En principio, éste es un valor más confiable que el de Hayes *et al.* [78], ya que se obtiene de una base mayor, pero todavía está sujeto a una incertidumbre de  $\pm 5 \text{ kcal/mol}$  (ver las secciones 2.4 y 2.5), pues sólo se calculó al nivel SCF.

Finalmente, en el trabajo más reciente de Ewig y Van Wazer [80], se hace una optimización más completa de las estructuras de diversas especies moleculares (*i. e.* sin restricciones sobre los parámetros estructurales), empleando tres distintas bases, STO-3G\*, 4-31G\* y otra de calidad DZP. En el caso del  $H_4P_2O_7$  usan únicamente la base STO-3G\*, y concluyen que el ácido difosfórico es muy similar a un par de moléculas de ácido ortofosfórico unidas por el puente P - O - P. Con esta base pequeña también calculan la energía de la reacción (2-45) y llegan al valor de  $\Delta E = -0.49 \text{ kcal/mol}$ , del cual concluyen que la hidrólisis es termoneutra en la fase gaseosa.

Puede observarse que no hay un consenso claro entre los datos de los cálculos cuánticos. En esta tesis se retoma el problema con una aproximación más sistemática, tanto para la optimización de las geometrías, como para el cálculo de la energía de la reacción. En el siguiente capítulo se presenta una revisión sobre los cálculos cuánticos y su aplicación al estudio de las reacciones químicas. En el capítulo cuatro se hará una comparación de los resultados aquí presentados con los que se obtuvieron en esta tesis.

## 2.10 Referencias del capítulo 2

- [1].- L. Pauling y E. B. Wilson, *Introduction to Quantum Mechanics*. McGraw-Hill Kogakusha, Ltd., (1935).
- [2].- T. A. Albright, J. Burdett y M. H. Whangboo, *Orbital Interactions in Chemistry*. Wiley Interscience, N. Y. (1985).
- [3].- W. J. Hehre, L. Radom, P. V. R. Schleyer y J. A. Pople, *Ab Initio Molecular Orbital Theory*. Wiley Interscience, N. Y. (1986).
- [4].- J. Sauer, *Chem. Rev.* **89**, 199 (1989).
- [5].- C. C. J. Roothan, *Rev. Mod. Phys.* **23**, 69 (1951).
- [6].- S. F. Boys, *Proc. Roy. Soc. London* **A200**, 542 (1950).
- [7].- T. H. Dunning y P. J. Hay, en *Methods of Electronic Structure Theory*, compilado por H. F. Schaefer para *Modern Theoretical Chemistry*, vol. 3, p. 1, Plenum, N. Y. (1977).
- [8].- R. Ahlrichs y P. R. Taylor, *J. Chim. Phys. Phys. - Chim. Biol.* **78**, 315 (1981).
- [9].- S. Huzinaga, *Comp. Phys. Rep.* **2**, 279 (1985).
- [10].- E. R. Davidson y D. Feller, *Chem. Rev.* **86**, 681 (1986).
- [11].- J. C. Barthelat y P. Durand, en *Química Teórica*, S. Fraga, ed., Colección Nuevas Tendencias, vol. 4, p. 49, CSIC, Madrid (1987).
- [12].- P. Pulay, en *Applications of Electronic Structure Theory*, compilado por H. F. Schaefer, para *Modern Theoretical Chemistry*, vol. 3, p. 153, Plenum, N. Y. (1977).
- [13].- H. B. Schlegel, en *Ab Initio Methods in Quantum Chemistry - I*, p. 249, K. P. Lawley, ed., Wiley Interscience, N. Y. (1987).
- [14].- P. Pulay, en *Ab Initio Methods in Quantum Chemistry - II*, p. 241, K. P. Lawley, ed., Wiley Interscience, N. Y. (1987).
- [15].- J. F. Gaw y N. C. Handy, *Annu. Rep. C, Chem. Soc.* **81**, 291 (1984).
- [16].- M. J. Frisch, J. S. Binkley, H. B. Schlegel, K. Raghavachari, C. F. Melius, R. L. Martin, J. J. P. Stewart, F. W. Bobrowicz, C. M. Rohlfing, L. R. Kahn, D. J. Defrees, R. Seeger, R. A. Whiteside, D. J. Fox, E. M. Fleuder y J. A. Pople, *GAUSSIAN - 86*, Carnegie Mellon QCPU, Pittsburgh (1984).

[17].- M. J. Frisch, M. Head-Gordon, H. B. Schlegel, K. Raghavachari, J. S. Binkley, C. González, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, E. M. Fleuder, S. Topiol y J. A. Pople, *GAUSSIAN - 88*, Gaussian, Inc., Pittsburgh (1988).

[18].- M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. González, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol y J. A. Pople, *GAUSSIAN - 90, Revision I*, Gaussian, Inc., Pittsburgh (1990).

[19].- G. Fogarassi y P. Pulay, *Annu. Rev. Phys. Chem.* **35**, 191 (1984).

[20].- B. A. Hess, L. J. Schaad, P. Čársky y R. Zharadnik, *Chem. Rev.* **86**, 709 (1986).

[21].- R. D. Amos, *The Cambridge Analytical Derivatives Package CCP1/84/4*, Computational Science Group, SERC, Daresbury (1984).

[22].- M. Dupuis, J. D. Watts, H. O. Villar y G. J. B. Hurst, *HONDO*, versión 7.0, KGN - 169, IBM, Kingston (1988).

[23].- P. Čársky y M. Urban, *Ab Initio Calculations, Methods and Applications in Chemistry*, en *Lecture Notes in Chemistry* No. 16, G. Berthier, M. J. S. Dewar, H. Fischer, K. Fukui, H. Hartman, H. H. Jaffé, J. Jortner, W. Kutzelnigg, K. Ruedenberg, E. Scrocco y W. Zeil, eds., Springer - Verlag, Berlin (1980).

[24].- A. Leś e I. Ortega-Blake, *Int. J. Quantum Chem.* **6**, 189 (1979)

[25].- S. Iwata, en *Quantum Chemistry Literature Data Base*, K. Ohno y K. Morokuma, eds., Elsevier, Amsterdam (1982).

[26].- H. Tatekawi y S. Huzinaga, *J. Chem. Phys.* **71**, 4339 (1979).

[27].- H. Tatekawi y S. Huzinaga, *J. Comput. Chem.* **1**, 205 (1980).

[28].- Y. Sakai, H. Tatekawi y S. Huzinaga, *J. Comput. Chem.* **2**, 100 (1981).

[29].- Y. Sakai, H. Tatekawi y S. Huzinaga, *J. Comput. Chem.* **3**, 6 (1982).

[30].- P. Hobza y J. Sauer, *Theor. Chim. Acta* **65**, 279 (1984).

[31].- J. Sauer, *Z. Chem.* **25**, 254 (1985).

[32].- W. Kutzelnigg, *Topics Curr. Chem.* **41**, 31 (1973).

[33].- L. C. Snyder, *J. Chem. Phys.* **46**, 3602 (1967).

[34].- L. C. Snyder y H. Basch, *J. Amer. Chem. Soc.* **91**, 2189 (1969).

- [35].- P. George, M. Trachtman, A. M. Brett y C. W. Bock, *Intl. J. Quantum Chem.* **12**, 61 (1977).
- [36].- P. George, M. Trachtman, A. M. Brett y C. W. Bock, *J. Chem. Soc., Perkin Trans.* **2**, 1036 (1977).
- [37].- W. H. Hehre, R. T. McIver, Jr., J. A. Pople y P. v. R. Schleyer, *J. Amer. Chem. Soc.* **96**, 7162 (1974).
- [38].- L. Radom, *J. Chem. Soc., Chem. Comm.* 403 (1974).
- [39].- W. H. Hehre, D. Ditchfield, L. Radom y J. A. Pople, *J. Amer. Chem. Soc.* **92**, 4796 (1970).
- [40].- L. Radom, W. H. Hehre y J. A. Pople, *J. Chem. Soc., Inorg. Phys. Theor.* 2299 (1971).
- [41].- L. Radom, W. H. Hehre y J. A. Pople, *J. Amer. Chem. Soc.* **93**, 289 (1971).
- [42].- R. Ahlrichs, *Theor. Chim. Acta* **35**, 59 (1974).
- [43].- E. Clementi y J. N. Gayles, *J. Chem. Phys.* **47**, 3837 (1967).
- [44].- P. Goldfinger y G. Verhaegen, *J. Chem. Phys.* **50**, 1467 (1969).
- [45].- R. McWeeney y B. T. Sutcliffe, *Methods of Molecular Quantum Mechanics*, Academic, London (1969).
- [46].- N. C. Handy, *Faraday Symp. Chem. Soc.* **19**, 17 (1984).
- [47].- C. W. Bauschlicher, Jr. y P. R. Taylor, *J. Chem. Phys.* **85**, 2779 (1986).
- [48].- S. A. Kucharski y R. J. Bartlett, *Adv. Quantum Chem.* **18**, 281 (1986).
- [49].- F. W. Bobrowicz y W. A. Goddard III, en *Methods of Electronic Structure Theory*, compilado por H. F. Schaefer, para *Modern Theoretical Chemistry*, vol. 3, p. 79, Plenum, N. Y. (1977).
- [50].- F. Bernardi y M. A. Robb, en *Ab Initio Methods in Quantum Chemistry - I*, p. 155, K. P. Lawley, ed., Wiley Interscience, N. Y. (1987).
- [51].- K. P. Lawley, ed., *Ab Initio Methods in Quantum Chemistry - I*, Wiley Interscience, N. Y. (1987).
- [52].- K. P. Lawley, ed., *Ab Initio Methods in Quantum Chemistry - II*, Wiley Interscience, N. Y. (1987).
- [53].- R. Ahlrichs y P. Scharf, en *Ab Initio Methods in Quantum Chemistry - I*, p. 501, K. P. Lawley, ed., Wiley Interscience, N. Y. (1987).

- [54].- J. Noga y R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987).
- [55].- J. A. Pople, J. S. Binkley y R. Seeger, *Intl. J. Quantum Chem. Symp.* **S10**, 1 (1976).
- [56].- W. A. Goddard III y T. C. McGill, *J. Vac. Sci. Technol.* **16**, 1308 (1979).
- [57].- P. E. M. Siegbahn, J. Almlöf, A. Heiberg y B. O. Roos, *J. Chem. Phys.* **74**, 2384 (1981).
- [58].- B. O. Roos, P. R. Taylor y P. E. M. Siegbahn, *Chem. Phys.* **48**, 157 (1980).
- [59].- P. E. M. Siegbahn, *Intl. J. Quantum Chem.* **23**, 1869 (1983).
- [60].- R. J. Buenker y S. D. Peyerimhoff, *Theor. Chim. Acta* **35**, 33 (1974).
- [61].- R. J. Buenker, S. D. Peyerimhoff y W. Butscher, *Mol. Phys.* **35**, 771 (1978).
- [62].- B. Huron, J. P. Malrieu y P. Rancurel, *J. Chem. Phys.* **58**, 5745 (1973).
- [63].- N. C. Handy, J. F. Gaw y E. D. Simandiras, *J. Chem. Soc., Faraday Trans. 2*, **83**, 1577 (1987).
- [64].- M. D. Newton y R. D. Kestner, *Chem. Phys. Lett.* **94**, 198 (1983).
- [65].- M. M. Szczeniński y S. Scheiner, *J. Chem. Phys.* **84**, 6328 (1986).
- [66].- M. J. Frisch, J. E. Del Bene, J. S. Binkley y H. F. Schaefer III, *J. Chem. Phys.* **84**, 2279 (1986).
- [67].- J. Sauer, P. Hobza, P. Čársky y R. Zharadnik, *Chem. Phys. Lett.* **134**, 553 (1987).
- [68].- P. Hobza, B. Schneider, J. Sauer y P. Čársky, *Chem. Phys. Lett.* **134**, 423 (1987).
- [69].- J. Sauer, B. Kathan y R. Ahlrichs, *Chem. Phys.* **113**, 201 (1987).
- [70].- P. Hohenberg y W. Kohn, *Phys. Rev.* **136B**, 864 (1964).
- [71].- W. Kohn y L. J. Sham, *Phys. Rev.* **140A**, 1133 (1965).
- [72].- R. O. Jones, en la referencia [50], p. 413.
- [73].- D. R. Salahub, en la referencia [51], p. 447.
- [74].- B. I. Dunlap, en la referencia [51], p. 287.

[75].- P. W. Atkins, *Quanta, a Handbook of Concepts*, para *Oxford Chemistry Series*, P. W. Atkins, J. S. E. Holker y A. K. Holliday, eds., Clarendon, Oxford (1974).

[76].- P. W. Atkins, *Physical Chemistry*, Oxford University Press, Oxford (1978).

[77].- P. George, R. J. Witonsky, M. Trachtman, C. Wu, W. Dorwart, L. Richman, W. Richman, F. Shurayh y B. Lentz, *Biochim, Biophys. Acta* **223**, 1 (1970)

[78].- M. D. Hayes, L. G. Kenyon y P. A. Kollman, *J. Am. Chem. Soc.* **100**, 4331 (1978)

[79].- M. O'Keeffe, B. Domengès y G. V. Gibbs, *J. Phys. Chem.* **89**, 2304 (1985)

[80].- C. S. Ewig y J. R. Van Wazer, *J. Am. Chem. Soc.* **110**, 79 (1988)

CAPÍTULO 3

---

SIMULACIONES DE LOS  
LÍQUIDOS MOLECULARES

### 3.1 La mecánica estadística y la función de partición

Partiendo de primeros principios, con los cálculos cuánticos se logra la descripción molecular de un sistema. Esta aproximación es aplicable a sistemas con un número limitado de átomos, dadas las restricciones discutidas en el capítulo anterior. En cambio, cuando se desea comparar las predicciones de un modelo molecular con los resultados experimentales de sistemas macroscópicos, es necesario recurrir a la mecánica estadística de una colección de muchas partículas. La conexión con la termodinámica se hace a través de la *función de partición del ensemble* definido con la colección mencionada de partículas. El concepto de *ensemble* fue introducido por Gibbs (ver e. g. [1, 2]) y consiste en considerar una cantidad arbitrariamente grande de réplicas del sistema bajo estudio, suponer que cada una de ellas está sujeta a las mismas condiciones macroscópicas, y aceptar la hipótesis de que cada estado macroscópico instantáneo, está relacionado con el promedio de la magnitud de interés tomado sobre todas las réplicas. La *función de partición* describe la distribución (la partición) de las réplicas de un ensemble en el espacio  $\Gamma$  de microestados accesibles al sistema, y adopta diferentes formas funcionales, dependiendo de las condiciones del sistema. Estas condiciones definen los distintos tipos de ensemble: cuando se mantienen constantes la energía  $E$ , el volumen  $V$ , y el número de partículas  $N$ , el ensemble se denomina *microcanónico* ( $E, V, N$ ) y su función de partición está dada por

$$\Omega(E, V, N) \delta E = \Omega_o(E, V, N) \quad (3-1),$$

donde  $\Omega$  es el número de microestados accesibles al sistema en el intervalo  $[E, E + \delta E]$  y también es constante,  $\Omega_o$ . Esta función corresponde a la hipótesis estadística de que en un estado de equilibrio macroscópico, todos los microestados estacionarios con igual energía, tienen igual probabilidad *a priori*. El potencial termodinámico que se relaciona directamente con esta función es la entropía (véase [1,2] y apéndice A):

$$S(E, V, N) = k \ln \Omega_o(E, V, N) \quad (3-2),$$

en que  $k$  es la constante de Boltzmann.

Si en vez de la energía  $E$ , se mantiene constante la temperatura  $T$ , el ensemble se denomina *canónico* ( $T, V, N$ ), y su función de partición es:

$$Z(T, V, N) = \sum_i e^{-E_i(V, N)/kT} = \int_0^\infty e^{-E/kT} \Omega(E, V, N) dE \quad (3-3).$$

Cabe hacer notar que la relación entre las funciones de partición canónica,  $Z(T, V, N)$ , y microcanónica,  $\Omega(E, V, N) \delta E$ , refleja el hecho de que cuando la temperatura  $T$  se mantiene constante, cada réplica puede tener distinta energía  $E_i(V, N)$ . El potencial termodinámico que puede obtenerse del ensemble canónico, es la función de Helmholtz (véase [1,2] y apéndice A):

$$F(T, V, N) = -kT \ln Z(T, V, N) \quad (3-4).$$

Cuando en lugar del número de partículas  $N$  lo que se mantiene constante es el potencial químico  $\mu$ , se obtiene el ensemble *gran canónico*  $(T, V, \mu)$ , con la función de partición

$$\begin{aligned}\Xi(T, V, \mu) &= \sum_{N=0}^{\infty} \sum_j e^{-E_{Nj}/kT} e^{-\mu N/kT} \\ &= \sum_{N=0}^{\infty} Z(T, V, N) e^{-\mu N/kT}\end{aligned}\quad (3-5),$$

que se relaciona con el potencial denominado entalpía, en la siguiente forma:

$$PV = H - U = kT \ln \Xi(T, V, \mu) \quad (3-6),$$

donde  $P$  es la presión.

Otro tipo de ensemble gran canónico es el  $(T, P, N)$ , donde se mantienen constantes la temperatura  $T$ , la presión  $P$  y el número de partículas  $N$ . La función de partición correspondiente es

$$Y(T, P, N) = \int_0^{\infty} e^{-PV/kT} Z(T, V, N) dV \quad (3-7),$$

relacionada con el potencial de Gibbs,

$$G(T, P, N) = -kT \ln Y(T, P, N) \quad (3-8).$$

Así como los potenciales termodinámicos se obtienen unos de otros por medio de transformaciones de Legendre, las ecuaciones (3-1), (3-3), (3-5) y (3-7) muestran que unas funciones de partición pueden obtenerse a partir de las otras, por medio de transformaciones de Laplace. Entonces el problema fundamental de la mecánica estadística se reduce al cálculo de la función de partición de un sistema dado, a partir del modelo mecánico supuesto para su descripción.

### 3.2 La función de partición molecular

Cuando se consideran moléculas que no interactúan, como en el caso de un gas ideal, pero que tienen estructura interna, la energía del sistema se divide entre los modos de movimiento translacional, rotacional, vibracional y electrónico. Ya que las moléculas son independientes unas de otras, la función de partición del sistema es  $Z = z^n/n!$ , donde  $z$  es la función de partición *molecular* en un ensemble canónico,

$$z = \sum_i e^{-\epsilon_i/kT} \quad (3-9).$$

Si la energía total de cada molécula puede separarse en las partes arriba enumeradas.

$$\epsilon_i = \epsilon_i^t + \epsilon_i^r + \epsilon_i^v + \epsilon_i^e \quad (3-10).$$

entonces la función de partición molecular se reduce a un producto de las funciones para cada modo de movimiento:

$$z = \sum_i e^{-(\epsilon_i^t + \epsilon_i^r + \epsilon_i^v + \epsilon_i^e)/kT} = z^t z^r z^v z^e \quad (3-11).$$

Esta factorización no es exacta, pues los distintos modos de movimiento no son completamente independientes unos de otros, pero en general pueden considerarse en forma separada, como se hace *e. g.* en la aproximación de Born - Oppenheimer para cálculos moleculares. Entonces la función de partición clásica para la traslación es

$$z^t = (2\pi mkT/h^2)^{3/2} V \quad (3-12).$$

que se aplica a una molécula de masa  $m$  en un recipiente de volumen  $V$ , siendo  $h$  la constante de Planck. El término entre paréntesis corresponde a la energía cinética, y el volumen a la parte configuracional.

Para el cálculo de la función de partición rotacional  $z^r$ , se sustituyen los valores experimentales de los niveles de energía rotacional en la expresión para la función de partición y se suman las exponenciales numéricamente, sin olvidar que varios estados rotacionales pueden corresponder a una misma energía. Por ejemplo, en el caso de una molécula diatómica, los estados se indican con los números cuánticos  $J$  y  $M$ , pero la energía sólo depende de  $J$ ; ya que hay  $2J+1$  valores de  $M$  para una  $J$  dada, existen  $2J+1$  estados con la misma energía para cada valor de  $J$  y la función de partición rotacional resulta ser

$$z^r = \sum_{J,M} e^{-E_J/kT} = \sum_J (2J+1) e^{-E_J/kT} \quad (3-13).$$

En general la función de partición rotacional, depende del tipo de molécula que se trate. Una expresión genérica para cualquier molécula, habrá de incluir sus tres momentos de inercia:  $I_A$ ,  $I_B$  e  $I_C$ , y un número  $\sigma$  que indique la simetría ante distintas orientaciones (véase [3]):

$$z^r \sim (\pi^{1/2}/\sigma) \{(2I_A kT/h^2)(2I_B kT/h^2)(2I_C kT/h^2)\} \quad (3-14).$$

La contribución vibracional presenta características similares en cuanto a la complejidad, que depende del tipo de molécula bajo estudio: para una diatómica, la función de partición vibracional puede evaluarse al sustituir los niveles energéticos vibracionales experimentales en la expresión

$$z^v = \sum_j e^{-\epsilon_j^v/kT} \quad (3-15).$$

pero para una molécula poliatómica, cada modo normal de vibración tiene su propio conjunto independiente de niveles energéticos, de manera que

$$z^v = z^v(1) z^v(2) \dots \quad (3-16).$$

donde  $z^v(K)$  es la función de partición para el  $K$ -ésimo modo normal. Este es un método directo para calcular la función de partición, con la información obtenida por espectroscopía

vibracional. Una alternativa cuando las excitaciones vibracionales no son muy grandes, es la de aproximarlas por vibraciones armónicas simples. En ese caso las energías son  $\epsilon_j = (j + \frac{1}{2})h\omega$ ,  $j = 0, 1, 2, \dots$ , donde  $\omega = (k/\mu)^{1/2}$  (cfr. ec. 2-41), siendo  $k$  la constante de fuerza y  $\mu$  la masa reducida. En el nivel de punto cero (ZPE, ver la sección 2.8), los valores de la energía son  $\epsilon_j^{ZPE} = jh\omega$  y la aproximación armónica lleva a una expresión simple para la función de partición vibracional:

$$z^v = \sum_j e^{-j h \omega / k T} = \sum_j (e^{-h \omega / k T})^j \quad (3-17),$$

serie que converge a

$$z^v = \frac{1}{1 - e^{-h \omega / k T}} \quad (3-18).$$

Cuando la molécula tiene varios modos normales, la función de partición vibracional total, se forma multiplicando expresiones como la ec. (3-18), que tengan los valores apropiados a las frecuencias vibracionales de cada modo.

Finalmente, la contribución electrónica a la función de partición es muy simple: ya que las separaciones entre las energías electrónicas son en general muy grandes, las exponenciales  $\exp(-\epsilon^e/kT)$  son todas muy pequeñas, salvo por el estado base, en el cual  $\epsilon^e = 0$ . Por esto en la mayor parte de los casos

$$z^e = 1 \quad (3-19).$$

Una excepción relevante ocurre cuando los átomos tienen un estado base degenerado, en cuyo caso

$$z^e = g \quad (3-20),$$

donde  $g$  es la degeneración del estado base (i. e. el número de microestados con la misma energía).

La función de partición molecular es entonces la multiplicación expresada en la ec. (3-11).

### 3.3 La función de partición cinética y la configuracional

En el caso de un ensemble canónico clásico, cuyo hamiltoniano sea  $H(\mathbf{p}, \mathbf{x})$ , la función de partición puede escribirse en la forma [1]

$$Z = \frac{1}{N! h^{N_f}} \int \dots \int e^{-H(\mathbf{p}, \mathbf{x}) / k T} d\mathbf{p} d\mathbf{x} \quad (3-21),$$

ya que el número de microestados accesibles a un sistema con una energía en el intervalo  $[E, E + \delta E]$  es

$$\Omega(E, V, N) = \frac{1}{N! h^{N_f}} \int_{E \leq H \leq E + \delta E} \dots \int d\mathbf{p} d\mathbf{x} \quad (3-22).$$

Cuando la energía potencial depende sólo de la posición, la función de partición puede separarse en:

$$Z = \frac{1}{N!h^{N_f}} \int \dots \int e^{-K(\mathbf{p})/kT} d\mathbf{p} \int \dots \int e^{-U(\mathbf{x})/kT} d\mathbf{x} = \frac{1}{N!h^{N_f}} Z_c Q_N \quad (3-23),$$

donde  $K(\mathbf{p})$  es la energía cinética del sistema,  $U(\mathbf{x})$  es la energía potencial,  $N$  el número de moléculas,  $\mathbf{p}$  los momentos y  $\mathbf{x}$  las coordenadas. La parte cinética  $Z_c$  generalmente puede obtenerse de una manera sencilla al integrar sobre las  $N_f$  componentes de momento, como se hizo en la sección anterior para la función de partición molecular. En cambio la parte configuracional  $Q_N$  presenta mayores problemas, ya que hay que considerar la función de la energía potencial intermolecular total  $U_N(\mathbf{x}_1, \dots, \mathbf{x}_N)$ . Esta función puede expresarse como una suma de potenciales de muchos cuerpos

$$U_N = \sum_{i < j} V_2(\mathbf{x}_i, \mathbf{x}_j) + \sum_{i < j < k} V_3(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) + \dots \quad (3-24),$$

siempre que no haya campos externos ni se incluyan las condiciones a la frontera con el recipiente. Cuando se consideran deformaciones de las moléculas, también deben tomarse en cuenta términos de un solo cuerpo  $V_1(\mathbf{x}_i)$ .

Hasta fechas recientes los tratamientos mecánico estadísticos de los fluidos moleculares, habían usado sólo el primer término en el desarrollo (3-24), aproximación que se conoce como la suposición de aditividad por pares. Sin embargo, ya varios autores han mostrado la importancia de los términos de tres y cuatro cuerpos en la descripción de los fluidos moleculares, en especial del agua líquida [4 - 6] y en la actualidad se cuenta ya con una serie de potenciales para agua que incluyen efectos no aditivos (*e. g.* [7 - 9]). Además, ya también se comienzan a diseñar potenciales para moléculas deformables (*e. g.* [10 - 12]), pero aún hay una serie de problemas por superar [13].

Ahora bien, una vez que se cuenta con una función analítica de la energía potencial  $U_N(\mathbf{x}_1, \dots, \mathbf{x}_N)$ , el problema del cálculo de la función de partición configuracional, se reduce a un problema de integración en el espacio de configuraciones:

$$Q_N = \int \dots \int e^{-U_N/kT} d\mathbf{x} \quad (3-25).$$

Este no es un problema sencillo, dada la forma funcional de  $U_N(\mathbf{x}_1, \dots, \mathbf{x}_N)$ . De hecho, las aproximaciones analíticas tienen un alcance restringido a potenciales con muchas simetrías, e incluso diversos esquemas de integración numérica resultan imprácticos por la cantidad necesaria de operaciones [14]. Así, el ataque teórico analítico, se ha enfocado más bien a desarrollar esquemas generales aplicables a cualquier tipo de sistema mecánico estadístico [15]. La alternativa viable para el estudio de los fluidos moleculares la constituyen los métodos de simulación numérica.

### 3.4 Métodos de simulación numérica

Se han desarrollado dos esquemas para estudiar los fluidos moleculares: el método de Monte Carlo (MC) y el de dinámica molecular (MD). El primero (MC) es básicamente un método de integración numérica [16] que tiene una gran variedad de aplicaciones [17]. En la mecánica estadística generalmente se aplica el algoritmo de Metropolis [18], que hace uso de un ensemble canónico  $(T, V, N)$  y muestrea los estados *importantes* en el espacio de configuraciones. Como *importantes* se denominan aquellos estados que, dada su energía, contribuyen en mayor medida a la integral de la función de partición canónica (ec. 3-25), por el simple hecho de ser una mayoría. Tales estados se eligen, luego de un sorteo aleatorio, de acuerdo con una probabilidad que depende del cambio en energía  $\Delta U_N = U_N(\text{conf. } 2) - U_N(\text{conf. } 1)$  al pasar de una configuración a otra:

$$P(\text{conf. } 1 \rightarrow \text{conf. } 2) = \begin{cases} 1 & \text{si } \Delta U_N \leq 0 \\ e^{-\Delta U_N/kT} & \text{si } \Delta U_N > 0 \end{cases} \quad (3-26).$$

En el caso de las simulaciones MC, los valores de expectación de las variables termodinámicas, se estiman a partir de promedios sobre el ensemble,

$$\langle A \rangle \approx \frac{1}{M} \sum_{\nu=1}^M A(\mathbf{x}_\nu) \quad (3-27),$$

donde  $M$  es el número de configuraciones incluidas en la simulación, una vez que se ha llegado al equilibrio. Este último punto es más o menos delicado: tradicionalmente se ha considerado haber llegado al equilibrio cuando las energías de un número grande de configuraciones (algunos millones), se mantienen estables alrededor de un valor promedio, con una desviación de  $\pm 0.1\%$ . Algunos resultados obtenidos en nuestro laboratorio, muestran que en ciertos casos este criterio puede ser engañoso. Estos resultados y su discusión se presentan en el capítulo 5.

Por otra parte, el método de dinámica molecular [19, 20] consiste en obtener una solución de la evolución en el tiempo de un número finito de moléculas aisladas de su entorno; lo que corresponde a un ensemble microcanónico  $(E, V, N)$ , siguiendo la trayectoria en el espacio fase clásico de un solo punto representativo, restringido a moverse en una superficie de energía constante  $E$ .

El modo de obtener los valores de expectación termodinámicos, es la forma propuesta originalmente por Boltzmann, de manera que si  $A(\mathbf{x}, \mathbf{p})$  es alguna propiedad definida en el espacio fase, el promedio de ensemble  $\langle A \rangle$  es igual al promedio en un tiempo infinito (ésta es la hipótesis ergódica):

$$\langle A \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau A(\mathbf{x}(t), \mathbf{p}(t)) dt \quad (3-28).$$

En la práctica, las simulaciones MD corresponden a intervalos de tiempo de entre  $10^{-11}$  y  $10^{-10}$  segundos [14], por lo que sólo se aplican a procesos cuyos tiempos de relajación son

significativamente menores. En la sección 3.6 se discutirán algunos resultados recientes, relativos al tiempo que debe cubrir una simulación MD para dar resultados significativos.

### 3.5 La información que se obtiene de las simulaciones numéricas

Ya que en las simulaciones se cuenta con la posición de cada una de las moléculas del sistema bajo estudio, el tipo de información de que se dispone, corresponde a los parámetros estructurales: de las configuraciones muestreadas se obtienen directamente las funciones de correlación radial de las distintas especies atómicas involucradas (ver *e. g.* [9]), las orientaciones y las magnitudes de los momentos dipolares, su correlación espacial, etc. Además, también se conoce la energía de interacción de cada molécula con las otras, por lo que es posible contar, p. ej., el número de puentes de hidrógeno formados por los distintos tipos de moléculas del sistema (ver la revisión en [21]).

Con la función de correlación radial se calculan las funciones de estructura, que son comparables a los datos de difracción de rayos X y a los de dispersión de neutrones (*e. g.* [8, 9]), y con la correlación de dipolos es posible estimar la constante dieléctrica (*e. g.* [22]), aunque esto requiere de considerar a la celda central de la simulación, inmersa en un medio continuo con una constante dieléctrica dada.

De una simulación MD se obtienen, además, una serie de parámetros dependientes del tiempo con los cuales se calculan *e. g.* las constantes de difusión [23].

En principio, sería posible obtener cualquier cantidad termodinámica a partir de las simulaciones numéricas; sin embargo, esto presenta algunas dificultades, ya que el muestreo del espacio fase del sistema bajo estudio es incompleto: durante la simulación no se accede más que a una fracción pequeña de los microestados del sistema. Esto es particularmente crítico para las propiedades que dependen de las fluctuaciones. De cualquier forma, el valor de la energía que se obtiene de las simulaciones, es bastante confiable en un caso: MD, porque se impone la condición de energía constante (ensemble microcanónico), y en otro: MC, porque está diseñado para muestrear los microestados *importantes* para la energía (según el algoritmo de Metropolis).

Dado que en una simulación MC se emplea un ensemble canónico, se esperaría poder obtener la energía libre de Helmholtz  $F$  a partir de una simulación, usando la relación

$$F = -kT \ln Z \quad (3 - 29).$$

Esta relación puede expresarse como [24 - 26]:

$$F = kT \ln \langle e^{E/kT} \rangle \quad (3 - 30),$$

donde  $E = E(x_\nu)$  es la energía total de interacción del sistema en la configuración  $x_\nu$  y el promedio de ensemble puede estimarse aplicándole la ecuación (3-27):

$$\langle e^{E/kT} \rangle = \frac{1}{M} \sum_{\nu=1}^M e^{E(x_\nu)/kT} \quad (3 - 31).$$

Los resultados que se han obtenido con la evaluación de esta ecuación, son en general poco confiables debido al problema mencionado del muestreo incompleto del espacio fase, en especial en las regiones de energías altas. Para superar este problema se han empleado diversas técnicas, como la integración termodinámica [27], el método de perturbaciones [28, 29], y cálculos del potencial de la fuerza promedio [30]; pero todas ellas involucran una serie de simulaciones, cada una de las cuales es onerosa en términos de recursos computacionales (memoria y tiempo de CPU) [31].

Recientemente Jayaram y Beveridge [31] han propuesto un esquema para obtener la energía libre con los resultados de una sola simulación, en el cual se hace un cambio de variable, definiendo

$$E_r = E_{max} - E_{min} \quad (3-32),$$

que representa al intervalo de energías muestreado en la simulación, de modo que  $E_{max}$  y  $E_{min}$ , son las energías máxima y mínima que alcanza el sistema.  $E_{max}$  es una cantidad que no queda bien definida en el contexto de una simulación, ya que depende del número de configuraciones usado. Entonces  $E_r$  no puede evaluarse directamente de una simulación. Sin embargo, si expresamos a las energías  $E$  en unidades adimensionales, definiendo una variable  $\xi$  como

$$\xi = (E - E_{min})/E_r \quad (3-33),$$

que está en el intervalo  $0 \leq \xi \leq 1$ , entonces la energía configuracional total del sistema puede expresarse como

$$E = E_{min} + E_r \xi \quad (3-34)$$

y la ecuación (3-30) puede reescribirse como

$$F = E_{min} + kT \ln \langle e^{E_r \xi / kT} \rangle \quad (3-35).$$

La proposición de Jayaram y Beveridge [31] consiste en reemplazar en la ec. (3-23) el promedio de ensemble a la Boltzmann de la variable uniformemente distribuida  $E_r \xi$ , por el promedio de ensemble correspondiente a una distribución uniforme de una variable  $E_r^* \xi^*$  con distribución de Boltzmann:

$$\langle e^{E_r \xi / kT} \rangle = \langle e^{E_r^* \xi^* / kT} \rangle_u \quad (3-36),$$

lo que implica que  $e^{E_r^* \xi^* / kT}$  contiene información sobre la forma y el ancho de la distribución de energías. Una vez hecho esto, el promedio puede calcularse como una integral sobre el intervalo unitario

$$\begin{aligned} \langle e^{E_r^* \xi^* / kT} \rangle_u &= \int_0^1 e^{E_r^* \xi^* / kT} d\xi^* \\ &= \frac{e^{E_r^* / kT} - 1}{E_r^* / kT} \end{aligned} \quad (3-37)$$

y en tal caso la energía libre se expresa en la forma

$$F = E_{min} + kT \ln (e^{E_r^* / kT} - 1) - kT \ln (E_r^* / kT) \quad (3-38).$$

Ya que típicamente  $e^{E_r^*/kT} \gg 1$  en las simulaciones moleculares, se puede hacer la aproximación

$$F = E_{min} + E_r^* - kT \ln(E_r^*/kT) \quad (3-39).$$

Ahora el problema es determinar  $E_r^*$  a partir de la distribución de las energías. Al multiplicar ambos lados de la ecuación (3-33) por  $e^{-\xi}$  se obtiene

$$Ee^{-\xi} = E_{min}e^{-\xi} + E_r\xi e^{-\xi} \quad (3-40).$$

Si se define la transformación

$$\langle X \rangle_t = \frac{\int_0^1 X e^{-\xi} d\xi}{\int_0^1 e^{-\xi} d\xi} \quad (3-41)$$

y se aplica a ambos lados de (3-39) se llega a

$$\langle Ee^{-\xi} \rangle_t = \langle E_{min}e^{-\xi} + E_r\xi e^{-\xi} \rangle_t = E_{min}\langle e^{-\xi} \rangle_t + E_r\langle \xi e^{-\xi} \rangle_t \quad (3-42),$$

en que se despeja  $E_r$ :

$$E_r = \frac{\langle Ee^{-\xi} \rangle_t}{\langle \xi e^{-\xi} \rangle_t} - \frac{E_{min}\langle e^{-\xi} \rangle_t}{\langle \xi e^{-\xi} \rangle_t} \quad (3-43).$$

Ahora se busca expresar  $\langle Ee^{-\xi} \rangle_t$  en términos de la energía promedio  $\langle E \rangle$  obtenida en la simulación. La estrategia consiste en relacionar  $\langle Ee^{-\xi} \rangle_t$  con promedios de Boltzmann de la energía y reemplazar  $E_r$  por  $E_r^*$ . Se propone el ansatz

$$\langle Ee^{-\xi} \rangle_t = \langle E \rangle f(E, \xi) \quad (3-44),$$

donde  $f$  es una función de distribución desconocida. La propuesta es considerar a  $f$  como una función de distribución normal (gaussiana). Entonces es posible expresar a  $E_r^*$  como [31]

$$E_r^* = 2.5241\langle E \rangle - 2.9115E_{min} \quad (3-45),$$

en cuyo caso

$$F = \langle E \rangle - [C_1 E_{min} - C_2 \langle E \rangle + kT \ln \{((C_2 + 1)\langle E \rangle - (C_1 + 1)E_{min})/kT\}] \quad (3-46),$$

donde  $C_1 = 1.9115$  y  $C_2 = 1.5241$ . Tanto la energía promedio  $\langle E \rangle$  como la energía mínima  $E_{min}$  pueden obtenerse de una simulación numérica sin mayores complicaciones.

La aproximación descrita ha sido aplicada con buenos resultados a simulaciones de agua líquida, de soluciones acuosas de algunos iones monovalentes y de metano [31], como se verá en la siguiente sección.

### 3.6 La energía de hidratación

Una cantidad muy relevante para el caso de estudio de esta tesis, es la energía de hidratación; y más aún, la diferencia entre las energías de hidratación del pirofosfato y del ortofosfato. Nuevamente se presenta la situación de calcular una cantidad a partir de

una diferencia pequeña de números grandes, aunque los órdenes de magnitud no son tan dispares como en el caso de los cálculos cuánticos.

En principio, la energía de hidratación de una especie molecular, puede obtenerse de una simulación numérica con un procedimiento similar al descrito en la sección 2.4 para la estimación del calor de una reacción (ec. 2-21): primero se hace una simulación con un número determinado de moléculas de agua,  $N_w$ , y se obtiene la energía promedio del sistema,  $\langle E_w \rangle$ . Del sistema así equilibrado, se elimina una cantidad  $N_a$  de moléculas de agua, que se sustituye por la especie cuya hidratación se quiere estudiar. Entonces se efectúa la simulación del sistema con  $N_w - N_a$  moléculas de agua y una del soluto; así se obtiene un nuevo valor de la energía,  $\langle E_a \rangle$ , y la energía de hidratación se calcula como

$$\Delta E^{hid} = \langle E_a \rangle - (N_w - N_a) \frac{\langle E_w \rangle}{N_w} \quad (3-47),$$

donde  $\langle E_w \rangle / N_w$  es la energía de interacción por molécula, del solvente.

Cuando se cuenta con varias especies moleculares  $A, B, C, D, \dots$ , se repite el procedimiento arriba descrito para cada una de ellas, y luego es posible encontrar las diferencias requeridas. Por ejemplo, si se considera una reacción química



que ocurre en solución acuosa, la contribución de la hidratación a la energía de la reacción se obtiene de la diferencia

$$\Delta \Delta E^{hid} = \Delta E^{hid}(C) + \Delta E^{hid}(D) - \Delta E^{hid}(A) - \Delta E^{hid}(B) \quad (3-49).$$

Cuando en la ecuación (3-46) se utilizan directamente los promedios de las energías obtenidas de simulaciones MC, la cantidad  $\Delta E^{hid}$  resultante incluye únicamente las interacciones del soluto con el solvente, y correspondería a la entalpía de hidratación  $\Delta H^{hid}$ , al usar un ensemble  $(T, P, N)$ . Si se usa un ensemble  $(T, V, N)$ , se estará considerando la contribución de la energía potencial a la energía interna,  $\Delta U^{hid}$ . Sin embargo, cuando en la simulación con un ensemble  $(T, V, N)$ , se considera una sola molécula de soluto, es razonable suponer que el cambio en la presión, con respecto a un sistema de solvente puro, es muy pequeño, por lo que  $\Delta E^{hid}$  puede compararse a  $\Delta H^{hid}$ . Para la estimación de las energías libres de hidratación  $\Delta F^{hid}$  y  $\Delta G^{hid}$  pueden emplearse métodos como el descrito en la sección anterior [31] para  $\Delta F^{hid}$ , o métodos perturbativos para  $\Delta G^{hid}$  [27, 29, 30].

La aproximación de Jayaram y Beveridge [31] ha sido aplicada con buenos resultados a simulaciones de agua líquida con los modelos TIP4P [32], MCY [33] y SPC/E [34]. En el primer caso, el resultado es de  $\Delta F = -5.37 \pm 0.02 \text{ kcal/mol}$ , comparado con el valor experimental [35] de  $\Delta F = -5.74 \text{ kcal/mol}$ ; en el caso del MCY, la energía libre es  $\Delta F = -4.36 \text{ kcal/mol}$ , que se compara con el cálculo perturbativo de Mezei *et al.* [36] de  $\Delta F = -4.31 \text{ kcal/mol}$ . Finalmente, para el modelo SPC/E la energía libre es de  $\Delta F = -5.1 \text{ kcal/mol}$ , comparada a un cálculo anterior [35] de  $\Delta F = -5.5 \text{ kcal/mol}$ .

La aplicación a simulaciones de  $Li^+$ ,  $Na^+$  y  $Cl^-$  con el modelo TIP4P, da las siguientes energías libres de Helmholtz para la hidratación de los iones:  $-158.0 \pm 7.6 \text{ kcal/mol}$ ,  $-114.3 \pm 8.5 \text{ kcal/mol}$  y  $-87.2 \pm 9.0 \text{ kcal/mol}$ , respectivamente. Los resultados experimentales [37, 38] correspondientes a las energías libres de Gibbs de cada ión, son de  $-122.1 \text{ kcal/mol}$ ,  $-98.2 \text{ kcal/mol}$  y  $-77.0 \text{ kcal/mol}$ . También hay un resultado sobre la hidratación del metano, con un valor de  $\Delta F = 0.7 \pm 11.3 \text{ kcal/mol}$ , contra el valor experimental [37] de  $\Delta G = 2.0 \text{ kcal/mol}$ . Se encuentran diferencias importantes entre los datos teóricos y los experimentales, pero hay que recordar que las energías libres de Helmholtz,  $F$ , se toman a volumen constante, y las de Gibbs,  $G$ , a presión constante. Además, las diferencias pueden deberse bien a la aproximación de Jayaram y Beveridge, o bien al potencial empleado.

Dados los requerimientos computacionales, este tipo de cálculos apenas se está iniciando, y aún no queda claro a qué precisión pueden llegar. Sin embargo, un punto que se repite en especial en los métodos perturbativos, es la necesidad de hacer simulaciones que incluyan un gran número de configuraciones en MC, o bien un tiempo mayor de 100 ps en MD, tanto para llegar al equilibrio como para hacer una estadística razonable (ver *e. g.* [27] y resultados), por la necesidad de contar con sistemas de referencia bien conocidos. En todo caso, la cantidad más confiable es la energía promedio,  $\langle E \rangle$ , de modo que se pueden asegurar estimaciones de las entalpías de hidratación. Esta es, además, la magnitud más relevante en este trabajo, ya que existen datos experimentales con los cuales comparar los resultados de la simulación (ver las secciones 1.3, 1.4, 4.4 y 4.5); más aún, lo que aquí se busca es establecer la contribución de la interacción del soluto con el solvente a la energía de la reacción de hidrólisis del pirofosfato. Al comparar con el resultado experimental, se cuenta con un criterio para calibrar el método empleado. La ventaja de partir de primeros principios, es que se puede analizar la importancia de cada aproximación hecha y estudiar la posibilidad de refinarla, a diferencia de lo que ocurre con los modelos empíricos y los parámetros diseñados *ad hoc*.

En el siguiente capítulo se hace una descripción de los métodos desarrollados y aplicados al caso de interés, y se presentan los resultados de cálculos *ab initio* de la reacción de hidrólisis del pirofosfato, con los que se llega a un valor de la entalpía en fase gaseosa de  $-3.83 \text{ kcal/mol}$ . Además, se muestran resultados preliminares de simulaciones MC con los que se calculan las energías de hidratación del pirofosfato y del ortofosfato.

### 3.7 Referencias del capítulo 3

- [1].- J. Kestin y J. R. Dorfman, *A Course in Statistical Thermodynamics*, Academic Press, Nueva York (1971)
- [2].- R. Kubo, H. Ichimura, T. Usui y N. Hshitsume, *Statistical Mechanics*, North Holland, Amsterdam (1971)
- [3].- P. W. Atkins, *Physical Chemistry*, Oxford University Press, Oxford (1978)
- [4].- E. Clementi, W. Kolos, G. C. Lie y G. Ranghino, *Int. J. Quantum Chem.* **17**, 337 (1980)

- [5].- K. S. Hermansson, *J. Chem. Phys.* **89**, 2149 (1988)
- [6].- L. Gil-Adalid e I. Ortega-Blake, *J. Chem. Phys.* **94**, 3748 (1991)
- [7].- T. P. Lybrand y P. A. Kollman, *J. Chem. Phys.* **83**, 2923 (1985)
- [8].- U. Niesar, G. Corongiu, E. Clementi, G. R. Kneller y D. K. Battacharya, *J. Phys. Chem.* **94**, 7949 (1990)
- [9].- H. Saint-Martin e I. Ortega-Blake, *J. Chem. Phys.* **93**, 6448 (1990)
- [10].- E. Clementi y G. C. Lie, *Phys. Rev.* **A33**, 2679 (1986)
- [11].- A. Wallqvist, *Chem. Phys. Lett.* **165**, 437 (1990)
- [12].- Z. Slanina, *Chem. Phys. Lett.* **172**, 367 (1990)
- [13].- N. Pastor-Colón, *comunicación personal*
- [14].- D. W. Woods, en *Water, a Comprehensive Treatise*, **6**, F. Franks, ed. cap. 6, Plenum Press, Nueva York (1980)
- [15].- U. F. Edgal, *J. Chem. Phys.* **94**, 8179 (1991)
- [16].- I. M. Sóbol, *Método de Monte Carlo*, Mir, Moscú (1976)
- [17].- D. B. Hernández y L. J. Álvarez, *El Método de Monte Carlo*, Informe monográfico del depto. de Matemáticas, UAMI, México (1983)
- [18].- N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller y E. Teller. *J. Chem. Phys.* **21**, 1087 (1953)
- [19].- B. J. Alder y T. E. Wainwright, *J. Chem. Phys.* **27**, 1208 (1957)
- [20].- B. J. Alder y T. E. Wainwright, *J. Chem. Phys.* **31**, 459 (1959)
- [21].- H. Saint-Martin e I. Ortega-Blake, en *Biomolecules in Organic Solvents*, A. Gómez-Puyou, ed., cap. 3, CRC Press, Boca Ratón, FL (1991)
- [22].- M. Neumann, *J. Chem. Phys.* **82**, 5663 (1985)
- [23].- U. Niesar, G. Corongiu, M. J. Huang, M. Dupuis y E. Clementi, *Intl. J. Quantum Chem. Symp.* **23**, 421 (1989)
- [24].- D. L. Beveridge y F. M. DiCapua, *Ann. Rev. Biophys. Biophys. Chem.* **18**, 431 (1989)
- [25].- M. Mezei y D. L. Beveridge, en *Computer Simulations and Biomolecular Systems*, D. L. Beveridge y W. L. Jorgenson, eds., Annals of New York Academy of Sciences, 494. Nueva York (1986)

- [26].- J. P. Valleau y G. M. Torrie, en *Modern Theoretical Chemistry*, Vol. 6 B. J. Berne, ed., Plenum Press, Nueva York (1977)
- [27].- M. J. Mitchell y J. A. McCammon, *J. Compu. Chem.* 12, 271 (1991)
- [28].- T. P. Straatsma y J. A. McCammon, *Chem. Phys. Lett.* 167, 252 (1990)
- [29].- T. P. Straatsma y J. A. McCammon, *Chem. Phys. Lett.* 177, 433 (1991)
- [30].- M. Saito y H. Nakamura, *J. Compu. Chem.* 11, 76 (1990)
- [31].- B. Jayaram y D. L. Beveridge, *J. Phys. Chem.* 94, 7288 (1990)
- [32].- W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey y M. L. Klein, *J. Chem. Phys.* 79, 926 (1983)
- [33].- O. Matsuoka, E. Clementi y M. Yoshimine, *J. Chem. Phys.* 64, 1351 (1976)
- [34].- H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren y J. Hermans, en *Intermolecular Forces*, B. Pullman, ed., Reidel, Dordrecht, Holanda (1981)
- [35].- J. Hermans, A. Pathiasiril y A. Anderson, *J. Am. Chem. Soc.* 110, 5982 (1988)
- [36].- M. Mezei, S. Swaminathan y D. L. Beveridge, *J. Am. Chem. Soc.* 100, 3255 (1978)
- [37].- W. L. Jorgensen, J. F. Blake y J. K. Buckner, *Chem. Phys.* 129, 193 (1989)
- [38].- D. R. Rosseinsky, *Chem. Rev.* 65, 467 (1965)

LA REACCIÓN DE HIDRÓLISIS  
DEL PIROFOSFATO EN LA  
FASE GASEOSA

#### 4.1 Antecedentes

En el capítulo 1 se presentó una revisión sobre las teorías propuestas para explicar la alta reactividad de los anhídridos de fosfato en el ambiente celular. Las primeras consideraban únicamente efectos intramoleculares, en tanto que las últimas favorecen la idea de que la energía libre de las reacciones de hidrólisis, de los anhídridos de fosfato, depende de las diferencias en la hidratación de los reactantes y la de los productos. En esta tesis se presentan cálculos *ab initio* de la reacción de hidrólisis del pirofosfato neutro, con los que se muestra que los efectos intramoleculares no son despreciables.

En la sección 1.6 se discutieron las estimaciones que se han reportado, del calor de la reacción de hidrólisis del pirofosfato en la fase gaseosa. Estas estimaciones fueron obtenidas a partir de cálculos *ab initio* [1 - 3]. En esa sección se hizo notar la falta de un consenso en los resultados y la consecuente necesidad de un nuevo estudio, en el que se apliquen los criterios que se presentaron en el capítulo 2 para juzgar la confiabilidad de los cálculos.

En esta tesis se emplea una aproximación sistemática al problema, que permite demostrar que se ha alcanzado la convergencia al límite de Hartree - Fock, tanto en los datos estructurales como en los energéticos. También se incluyen los efectos de la energía de correlación y los de la energía del punto cero. Con todo esto se llega a un valor confiable de la entalpía de la reacción de hidrólisis



en la fase gaseosa.

A continuación se presentan los detalles del procedimiento usado en esta tesis.

#### 4.2 Descripción de los métodos empleados

Como se mencionó en la sección 2.3, una estimación de los valores que dependen de diferencias entre las energías calculadas *ab initio* para distintas especies (como las energías de interacción intermolecular o los calores de reacción), requiere que la geometría de cada molécula haya sido optimizada con la misma base.

Con el fin de comparar los parámetros estructurales calculados con distintas bases, y de estudiar la influencia de diferentes estructuras en el valor de la energía de la reacción (4-1), se realizaron varias optimizaciones de las geometrías de las moléculas  $H_2O$ ,  $H_3PO_4$  y  $H_4P_2O_7$ , usando el método semiempírico MNDO del paquete de cómputo MOPAC [4], y el método *ab initio* implementado en el programa GAUSSIAN-86 [5]. Los cálculos *ab initio* de energía para cada geometría optimizada, se efectuaron con el programa GAUSSIAN-88 [6], usando cinco bases distintas (ver la clasificación en la sección 2.2 y en [7]): (I) STO-3G\*, (II) 3-21G\* y (III) 6-31G\*, todas con funciones *d* de polarización sólo en los fósforos; (IV) una base 6-31G\* aumentada con funciones *d* de polarización en los oxígenos, y (V) otra que además tiene funciones *p* de polarización en los hidrógenos, 6-31G\*\*. Además se hicieron algunos cálculos con los pseudopotenciales de Barthelat [8] implementados en el programa MONSTERGAUSS [9], con una base mínima, más funciones *d* de polarización

en los fósforos (base VI), y una base de calidad DZP, con funciones *d* de polarización en los fósforos (base VII). La base (VI) tiene calidad similar a la base (I), y la base (VII) a la base (III).

Para verificar la convergencia en el calor de la reacción (4-1) obtenido al nivel SCF, se hicieron otros cálculos con dos bases mayores: (VIII) 6-311G(d,p), con funciones de polarización en todos los átomos, y (IX) 6-311G(2d,2p), con doble polarización en todos los átomos. En estos casos se usó el programa GAUSSIAN-90 [10].

#### 4.3 Estructuras de las moléculas de ortofosfato y de pirofosfato

(1)  $H_3PO_4$  (fig. 4-1). Se emplean en esta tesis cinco geometrías para esta molécula: la estructura *PI*, que se obtuvo con el método MNDO; la estructura obtenida por Ewig y Van Wazer [11] con una base como la (I), en una optimización completa, *PII*; la *PIII*, que fue generada con el método SCF, usando la base (II); la geometría *PIV* es la que obtuvieron O'Keeffe *et al.* [2] con una base equivalente a la (IV); por último, la estructura *PV* fue obtenida con la base (II), haciendo una optimización completa. En el apéndice B se dan algunos otros detalles de la geometría de esta molécula, y se presentan sus parámetros estructurales. La conclusión más relevante con respecto a esta especie molecular, es que la estructura queda bien determinada a nivel SCF y tiene una conformación tetraédrica (ver la figura 4-1).

(2)  $H_4P_2O_7$  (fig. 4-2). Para esta molécula también se cuenta con cinco grupos de estructuras: la *PPI* corresponde a la optimización con el método MNDO, sujeta a las restricciones de mantener una simetría  $C_2$  y los oxígenos en los fosforilos, en una posición opuesta a la de las otras geometrías optimizadas. Esto se hizo con el fin de estudiar la barrera de rotación alrededor del puente  $P - O - P$ . Hay dos estructuras *PPII*, una eclipsada y otra escalonada, que fueron obtenidas por Ewig y Van Wazer [3] con una base STO-3G\*; la escalonada fue optimizada casi sin ninguna restricción, y su energía es 2.3 kcal/mol menor que la de la geometría eclipsada, usando la base mínima. La estructura *PPIII* se calculó con la base (II) a partir de la geometría eclipsada *PPII*, permitiendo la variación de 16 grados de libertad, de los 33 que tiene la molécula. Se llegó a una geometría escalonada muy similar a la escalonada *PPII*. Hay también dos estructuras *PPIV*, una eclipsada y la otra escalonada, obtenidas con una base 6-31G\* por O'Keeffe *et al.* [2], y que tienen una diferencia de energías de 1.8 kcal/mol en favor de la escalonada. Ya que en la optimización de la estructura *PPIII* se variaron sólo la mitad de los parámetros, se procedió a una reoptimización, relajando todos los grados de libertad, y se llegó a la geometría *PPV* que no resultó muy distinta de la *PPIII*. Algunos otros detalles se describen en el apéndice B, en el cual se presentan los parámetros de todas las estructuras mencionadas. La característica más sobresaliente de la estructura del pirofosfato, es que se parece a un par de moléculas de  $H_3PO_4$  unidas por el puente  $P - O - P$ .

FIGURA 4-1

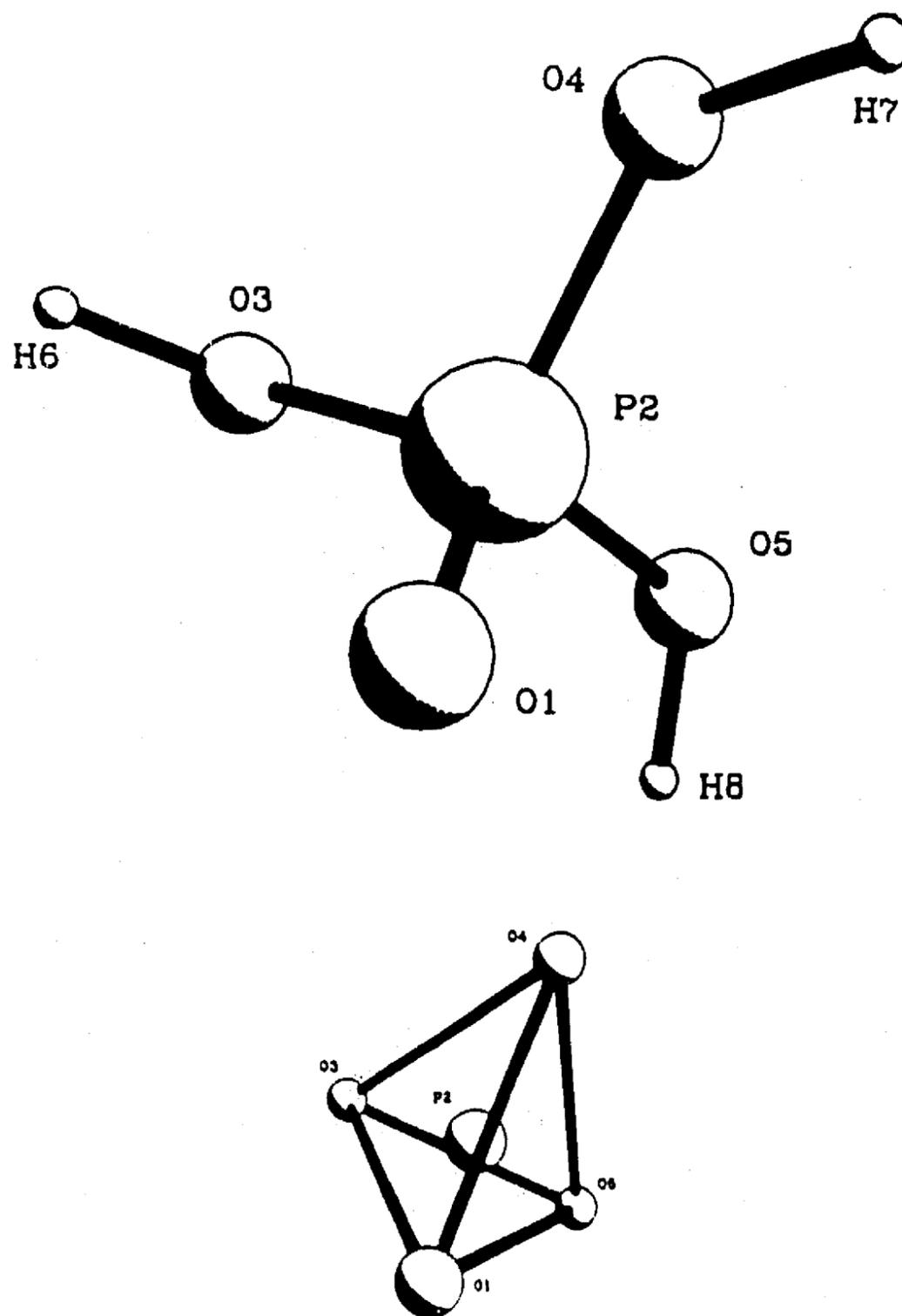


Figura 4-1. Estructura tetraédrica de la molécula del ortofosfato  $H_3PO_4$  (Geometría  $PIII$ ).

FIGURA 4-2

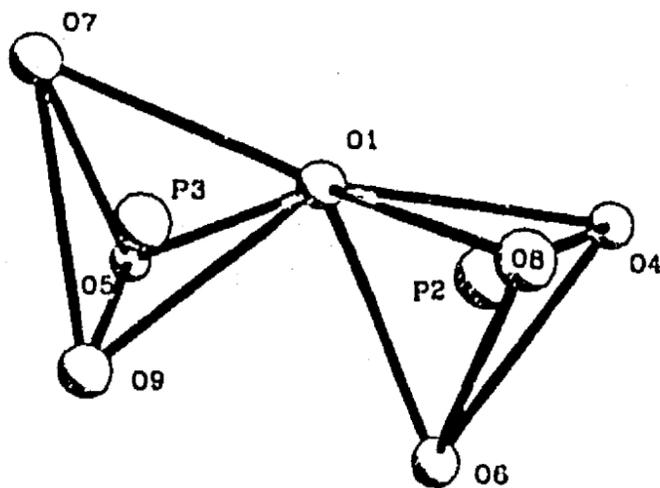
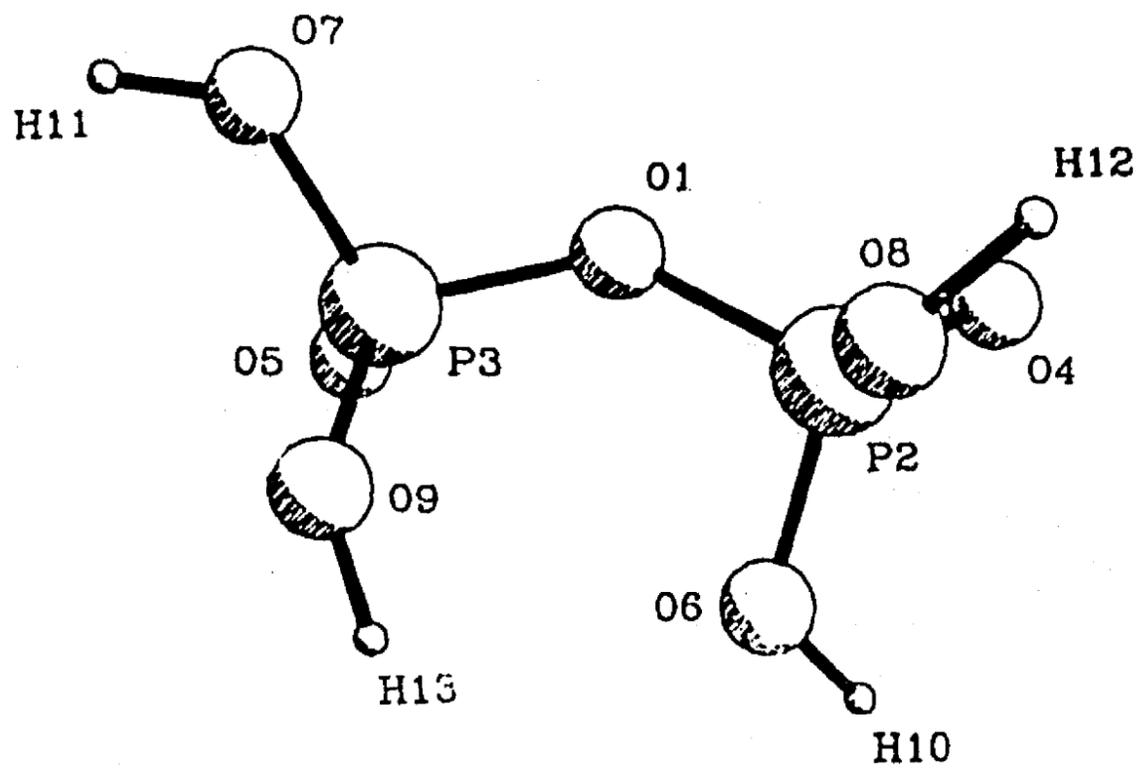


Figura 4-2. Estructura de la molécula del pirofosfato  $H_4P_2O_7$  (geometría *PPIII*). La molécula tiene gran similitud con un par de moléculas de ortofosfato, unidas por una rótula (el puente  $P - O - P$ ).

Puesto que al comparar las energías calculadas con la base (V), las geometrías *PIII* y *PPIII* resultan más estables que las geometrías *PV* y *PPV* (véase la Tabla AIII en la sección 4.5), todos los demás cálculos se hicieron con las primeras.

Ya que la flexibilidad del ángulo  $\phi$  del puente *P-O-P* parece ser uno de los grados de libertad más importantes de la molécula del  $H_4P_2O_7$ , se realizaron cálculos con la base (V), de estructuras en las que se modificó el ángulo  $\phi$  por  $-10^\circ$ ,  $-5^\circ$  y  $+5^\circ$ , con respecto al de la geometría *PPIII*,  $\phi_o = 138.96^\circ$ . Todos los demás parámetros se mantuvieron iguales. Se encontró que el ángulo  $\phi_o$  corresponde al mínimo de energía, al nivel SCF. A continuación se calcularon las energías de correlación al nivel MBPT(2), para las cuatro estructuras. Los resultados que se presentan en la Tabla 4-1 muestran que a este nivel, el ángulo  $\phi_o$  no es el óptimo, pero sí queda muy cerca de serlo.

TABLA 4-1

$\phi$	SCF	MBPT(2)	SUMA
$\phi_o - 10^\circ$	2.15	-1.11	1.04
$\phi_o - 5^\circ$	0.42	-0.46	-0.04
$\phi_o$	0.00	0.00	-0.00
$\phi_o + 5^\circ$	0.27	0.32	0.59

Tabla 4-1. Energías calculadas para la molécula del  $H_4P_2O_7$ , al variar el ángulo  $\phi$  del puente *P-O-P*, alrededor del ángulo  $\phi_o = 138.96^\circ$ , encontrado en la geometría *PPIII*. La primera columna muestra la variación del ángulo  $\phi$ . La segunda presenta las diferencias de energía al nivel SCF, con respecto al valor correspondiente a  $\phi_o$ ,  $E^{SCF} = -1208.037309$  Hartrees. En la tercera columna están las contribuciones de la correlación electrónica, al nivel MBPT(2) (valor de referencia:  $-1.475901$  Hartrees). La última columna muestra la energía relativa, con los dos efectos acumulados (SCF + MBPT(2)). Las energías están en kcal/mol.

Usando la base (V) se calculó la constante de fuerza de flexión,  $K_\theta = (1/r^2)(\partial^2 E/\partial\theta^2)$ , para el ángulo del puente *P-O-P*, tanto al nivel SCF ( $K_\theta = 36.6$  Nm $^{-1}$ ) como al nivel MBPT(2) ( $K_\theta = 30.5$  Nm $^{-1}$ ), llegando a resultados acordes con la observación de O'Keeffe *et al.* [2] de que existe una alta correlación entre las longitudes y los ángulos de enlace de dicho puente.

#### 4.4 La energía de la reacción de hidrólisis del pirofosfato

##### 4.4.1 Cálculos al nivel SCF

El punto principal de los cálculos realizados, es llegar a una estimación confiable de la energía de la reacción (4-1). Por tratarse de una reacción isodésmica (ver la sección 2.5), en principio la diferencia de energías SCF de los productos y los reactantes debería producir un valor con una incertidumbre de  $\pm 5$  kcal/mol. Para esto es necesario contar con estructuras óptimas y haber llegado a la región de convergencia al límite de Hartree-Fock,

discutido en la sección 2.4. Ninguno de los trabajos previos [1-3] demuestra haber llegado a esta región; en cambio, en la Tabla 4-2 se muestran los valores obtenidos usando distintas geometrías, y bases cada vez mayores. En esta Tabla puede observarse la convergencia al tomar las diferencias entre las energías de la reacción, calculadas con las geometrías *PIII* y *PPIII* (renglón *VIII*) al ir aumentando el tamaño de la base. Estas diferencias son de -2.49 kcal/mol, -1.11 kcal/mol y - 0.92 kcal/mol al pasar de la columna II a la III, de la III a la IV y de la IV a la V, respectivamente; notándose la disminución y, por ende, la convergencia.

Como se mencionó en la sección 4.2, también se hicieron cálculos del calor de la reacción (4-1) con las bases (VIII) y (IX), obteniendo los valores de -6.22 kcal/mol y de -9.01 kcal/mol, respectivamente. La diferencia entre el cálculo con la base (V) y el cálculo con la base (VIII) es de +1.16 kcal/mol, y al pasar de la (VIII) a la (IX) es de -2.79 kcal/mol (Tabla 4-3). Este comportamiento es discordante con la tendencia encontrada al usar las bases (II) a (V), y parece indicar que no se ha alcanzado la región convergente al límite HF, a pesar de que se están empleando bases de calidad TZP. En un principio se consideró que los valores obtenidos de -6.22 kcal/mol y -9.01 kcal/mol, representaban oscilaciones alrededor del valor en el límite; pero al revisar las construcciones de las bases (VIII) y (IX) se encontró que sus parámetros (los exponentes de las gaussianas y sus coeficientes de expansión), se habían elegido de manera que minimizaran las energías de los estados base atómicos, al realizar los cálculos al nivel MBPT(2), en vez de hacerlo al nivel SCF, como ocurre con los parámetros de los conjuntos (II) a (V) [12]. En ese caso, los cálculos del calor de la reacción (4-1), hechos con esas bases al nivel SCF, no son comparables a los valores presentados en las columnas (II) a (V) de la Tabla 4-3.

Se ha considerado relevante mencionar los resultados obtenidos con las bases (VIII) y (IX), porque representan un error que se puede cometer con relativa facilidad cuando se usan programas comerciales: al depositar la confianza en ellos, se pierde control sobre las condiciones del cálculo, lo cual da lugar a cifras que no tienen un significado claro. En el caso que aquí se estudia, la tendencia mostrada en la tabla 4-1 hasta la columna (V), permite concluir que es razonable usar el valor de  $\Delta E^{SCF} = -7.38 \text{ kcal/mol}$  como una estimación del calor de la reacción (4-1), calculada al nivel SCF. Una verificación ulterior podrá lograrse al emplear bases mayores, diseñadas para minimizar las energías atómicas calculadas al nivel SCF. En el Laboratorio de Biofísica ya se está trabajando en este sentido.

TABLA 4-2

I. Estimaciones del calor de la reacción de hidrólisis (4-1)

Geometría	Base						
	Todos los electrones			pseudo-potenciales			
	I	II	III	IV	V	VI	VII
I	2.36	-6.72		-11.86		-1.62	-3.09
II	4.22	-6.36		-11.72			-2.64
III	-2.63	-1.99		-4.51		8.35	-1.04
IV	-0.76	-1.63		-4.37			-0.60
V	-4.89	-2.84		-5.21		6.50	-0.91
VI	-3.03	-2.48		-5.08			-0.47
VII	0.71	-3.22	-6.21	-6.60	-7.67	1.97	-5.98
VIII	2.57	-2.86	-5.35	-6.46	-7.38		-5.54
IX		-0.92			-8.71		
X		-0.56			-8.43		

II. Las geometrías usadas para las diferentes moléculas son las siguientes:

Geometría	$H_4P_2O_7$	$H_3PO_4$	$H_2O$
I	PPI	PI	(a)
II	PPI	PI	(b)
III	PPII s	PII	(a)
IV	PPII s	PII	(b)
V	PPII e	PII	(a)
VI	PPII e	PII	(b)
VII	PPIII	PIII	(a)
VIII	PPIII	PIII	(b)
IX	PPV	PV	(a)
X	PPV	PV	(b)

Tabla 4-2. El calor de la reacción de hidrólisis (4-1), obtenido al nivel SCF usando diferentes geometrías, y calculado con distintas bases. Las estructuras usadas para los fosfatos, se describen en el texto. La letra *s* denota una estructura escalonada, y la *e*, una eclipsada. Las geometrías usadas para la molécula del agua son: (a) la experimental, y (b) la que se optimizó con la base correspondiente. Las energías están en kcal/mol.

TABLA 4-3

Base	II	III	IV	V
$\Delta E^{SCF}$	-2.86	-5.35	-6.46	-7.38
Diferencia con la columna anterior		-2.49	-1.11	-0.92

Base	VIII	IX
$\Delta E^{SCF}$	-6.22	-9.01
Diferencia con la columna anterior	+1.16	-2.79

Tabla 4-3. Aquí se muestra el calor de la reacción (4-1),  $\Delta E$ , calculado con las bases descritas en la sección 4.2. Las energías están en kcal/mol. En el renglón inferior se puede apreciar la convergencia al límite de Hartree - Fock, por la tendencia que siguen las diferencias en los valores obtenidos con cada base. Las cifras en las columnas VIII y IX no siguen esa tendencia, porque no son adecuadas para cálculos al nivel SCF (revisar la discusión en el texto).

#### 4.4.2 Inclusión de la energía de correlación y de la ZPE

Ahora bien, como se mencionó anteriormente, hay una incertidumbre de  $\pm 5$  kcal/mol debida a la omisión de la energía de la correlación electrónica. Ya que se ha mostrado que la base V da resultados en la región de convergencia al límite de Hartree-Fock, es razonable hacer cálculos de la energía de correlación (hay que recordar que no tiene sentido calcularla cuando se han usado bases pequeñas, sección 2.7). Esta contribución no había sido tomada en cuenta anteriormente, y al nivel MBPT(2) resulta ser de +2.25 kcal/mol, llevando la estimación de la energía de hidrólisis a

$$\Delta E = \Delta E^{SCF} + \Delta E^{MBPT(2)} = -5.14 \text{ kcal/mol} \quad (4-2).$$

Para poder comparar el valor obtenido del cálculo con la entalpía que se obtendría en un experimento, todavía falta agregar la energía de punto cero (ZPE, sección 2.8) y los efectos térmicos a temperatura ambiente. En este trabajo se usaron las frecuencias vibracionales armónicas, calculadas con la base (II) para la ZPE, y una aproximación a la función de partición molecular (sección 3.2) de cada especie, usando las frecuencias armónicas relacionadas con los grados de libertad de los reactantes y de los productos. Considerando estas contribuciones, se llega a una estimación de la entalpía de la reacción (4-1) en fase gaseosa de

$$\Delta H = -3.83 \text{ kcal/mol} \quad (4-3).$$

Este valor todavía puede estar sujeto a algunas incertidumbres pequeñas, causadas porque (a) aun la base más grande que se empleó no está saturada, (b) la energía de correlación

se calculó sólo hasta el segundo orden de la teoría de perturbaciones, y (c) las contribuciones ZPE y térmica fueron estimadas con aproximaciones armónicas. Dado el estudio sistemático sobre la convergencia de los resultados, es razonable proponer una incertidumbre total de  $\pm 1 \text{ kcal/mol}$ , de forma que la estimación de la entalpía de la reacción (4-1) en la fase gaseosa queda expresada como

$$\Delta H = -3.8 \pm 1.0 \text{ kcal/mol} \quad (4-4).$$

El valor asignado a la incertidumbre es común a cálculos al nivel  $MBPT(2)/6-31G^{**}$  sobre calores de reacciones isodésmicas [12, 13].

Las implicaciones de este resultado se discuten en la siguiente sección, y se replantearán en el capítulo 6.

#### 4.5 La entalpía de la reacción de hidrólisis del pirofosfato en la fase gaseosa

En vista del estudio sistemático hecho en esta tesis, queda demostrado que la entalpía de la reacción (4-1) es negativa en la fase gaseosa, con un valor estimado de  $\Delta H = -3.8 \pm 1.0 \text{ kcal/mol}$ , y por lo tanto la reacción es exotérmica. Si bien ésta no es la reacción de hidrólisis del verdadero sustrato de la enzima pirofosfatasa, sí se trata de una reacción de sustitución nucleofílica  $S_N2$  [14], sobre las cuales Dewar y Storch [15] han basado la hipótesis de *desolvatación*, que supone que son termoneutras en la fase gaseosa. El resultado aquí obtenido y el de Bash *et al.* [16] muestran que ocurre lo contrario y que por lo tanto la suposición de la termoneutralidad en la fase gaseosa debe ser descartada como una generalidad, en las explicaciones sobre la catálisis enzimática.

La otra suposición importante de la hipótesis de *desolvatación*, es la de considerar que la barrera de activación es debida exclusivamente al impedimento impuesto por el solvente para que los reactantes tengan colisiones favorables, a causa de cubrirlos con sendas capas de solvatación. Aunque la reacción sea exotérmica en la fase gaseosa, esto no demuestra la existencia de una barrera de activación en dicha fase, y queda todavía abierta la posibilidad de que la barrera se deba únicamente a los efectos del solvente. Warshel *et al.* [17] muestran que esta posibilidad presenta el problema de la gran cantidad de energía que se necesitaría para desolvatar especies cargadas. Esto, aunado a los resultados de Bash *et al.* [16], favorece la idea de que los sitios activos de las enzimas *anclan* diversos residuos de los sustratos y los acomodan de forma que al interactuar tengan mayor probabilidad de reaccionar, como lo propone la *teoría del estado de transición* [18], y por lo tanto su participación en la catálisis va más allá del simple hecho de remover el agua que rodea a los sustratos. Sin embargo, en un artículo muy reciente, Peng y Merz [19] presentan cálculos *ab initio* de la barrera de activación de la reacción



en la fase gaseosa, y simulaciones MD de la reacción en solución acuosa. Ellos muestran que en este caso, la barrera de activación en la solución acuosa se debe exclusivamente a la hidratación. De cualquier manera, la hipótesis de *desolvatación* no parece ser aplicable a todas las reacciones enzimáticas.

En cuanto a las teorías sobre la *alta energía* de los anhídridos de fosfato, en esta tesis se muestra que una parte importante es debida a los efectos intramoleculares; aunque no en la proporción considerada por Kalckar [20] o por Pullman [21]. De hecho, los conceptos usados por ellos, como el de la resonancia en oposición (véase la sección 1.4), han sido rebatidos por Ewig y Van Wazer [3], con base en que los cálculos *ab initio* no muestran indicios de que el ambiente electrónico en el oxígeno del puente  $P-O-P$ , sea muy distinto del de los demás oxígenos en el pirofosfato. En esta tesis se encontró la densidad de carga en cada oxígeno, usando la base  $V$  descrita en la sección 4.2. En el del puente  $P-O-P$  resultó ser de  $-0.79e$ , mientras que en los de doble ligadura fue de  $-0.70e$ , y en los de ligadura sencilla de  $-0.73e$ . Las diferencias relativas son del orden de un 10% solamente.

Aunque la conclusión inmediata es que los efectos intramoleculares son significativos, al comparar el valor estimado para la entalpía de la reacción (4-1) en la fase gaseosa,  $\Delta H_g = -3.8 \text{ kcal/mol}$ , con el valor medido para la misma reacción en solución acuosa,  $\Delta H_{aq} = -7.6 \text{ kcal/mol}$ , queda claro que la primera es apenas la mitad de la segunda, con lo que hay que concluir que el efecto de las interacciones de los solutos con el solvente puede ser tan relevante como las interacciones intramoleculares. Más aún, la energía libre medida por George *et al.* [22] para la reacción de hidrólisis (4-1) en solución acuosa, es de  $\Delta G = -9.5 \text{ kcal/mol}$ , por lo cual hay todavía  $1.9 \text{ kcal/mol}$  que corresponden al término entrópico  $T\Delta S$ . De hecho, De Meis propone que éste es el término que determina la  $\Delta G$  de la reacción en distintos ambientes, basándose en sus resultados sobre las hidrólisis de diversas especies de pirofosfato en mezclas de agua con solventes orgánicos [23 - 25].

Un problema en el tratamiento de este tipo de resultados experimentales, es la gran diversidad de reacciones de hidrólisis del "PP;" (ver las secciones 1.4 a 1.6 y [21]), con la consecuente carencia de un modelo molecular claro que permita explicarlos, lo que lleva al uso de conceptos empíricos como el de la actividad del agua, que no dan una imagen adecuada del fenómeno en términos moleculares; sino que sólo lo expresan en términos de otros parámetros ajustados a las desviaciones con respecto a los casos ideales. Además, los experimentos se realizan midiendo cantidades termodinámicamente significativas, y por lo tanto promediadas sobre poblaciones grandes de moléculas. Esta aproximación no puede llevar a una imagen molecular precisa [26], pero sí ha contribuido al desarrollo de nuevas ideas.

La alternativa que se propone en esta tesis, es la de partir de primeros principios con cálculos *ab initio*, y luego desarrollar simulaciones numéricas con potenciales analíticos basados en los cálculos, para llegar a una descripción molecular de una reacción particular. Este procedimiento también está siendo desarrollado por otros grupos, y ya ha producido algunos resultados relevantes (*e. g.* [16] y [19]).

Con los cálculos *ab initio* de esta tesis se llegó a una conclusión sobre la reacción (4-1) en la fase gaseosa, cuyas implicaciones se discutieron en esta sección y en la 4.5. Al comparar la entalpía de la fase gaseosa con la de la solución acuosa, se ha propuesto que la contribución de la hidratación es de la misma magnitud. Con el fin de someter a prueba esta aseveración, en esta tesis se realizaron simulaciones numéricas de soluciones acuosas de  $H_3PO_4$  y de  $H_4P_2O_7$ . Estas se describen en el siguiente capítulo.

#### 4.6 Referencias del capítulo 4

- [1].- M. D. Hayes, L. G. Kenyon y P. A. Kollman, *J. Am. Chem. Soc.* **100**, 4331 (1978)
- [2].- M. O'Keeffe, B. Domengès y G. V. Gibbs, *J. Phys. Chem.* **89**, 2304 (1985)
- [3].- C. S. Ewig y J. R. Van Wazer, *J. Am. Chem. Soc.* **110**, 79 (1988)
- [4].- J. J. P. Stewart, *QCPE Bull.* **5**, 133 (1985)
- [5].- M. J. Frisch, J. S. Binkley, H. B. Schlegel, K. Raghavachari, C. F. Melius, R. L. Martin, J. J. P. Stewart, F. W. Bobrowicz, C. M. Rohlfing, L. R. Kahn, D. J. Defrees, R. Seeger, R. A. Whiteside, D. J. Fox, E. M. Fleuder y J. A. Pople, *GAUSSIAN - 86*, Carnegie Mellon QCPU, Pittsburgh (1984)
- [6].- M. J. Frisch, M. Head-Gordon, H. B. Schlegel, K. Raghavachari, J. S. Binkley, C. González, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, E. M. Fleuder, S. Topiol y J. A. Pople, *GAUSSIAN - 88*, Gaussian, Inc., Pittsburgh (1988)
- [7].- S. Huzinaga, ed., *Gaussian Basis Sets for Molecular Calculations*, *Physical Sciences Data* **16**, Elsevier, Amsterdam (1984)
- [8].- J. C. Barthelat, Ph. Durand y A. Serafini, *Mol. Phys.* **33**, 159 (1977)
- [9].- M. Peterson y R. A. Poirier, *MONSTERGAUSS*, University of Toronto, Ontario (1987)
- [10].- M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. González, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol y J. A. Pople, *GAUSSIAN - 90, Revision I*, Gaussian, Inc., Pittsburgh (1990).
- [11].- C. S. Ewig y J. R. Van Wazer, *J. Am. Chem. Soc.* **107**, 1965 (1985)
- [12].- W. H. Hehre, L. Radom, P. v. R. Schleyer y J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley Interscience, Nueva York (1986)
- [13].- G. P. Ford y C. T. Smith, *J. Am. Chem. Soc.* **109**, 1325 (1986)
- [14].- J. R. Cox Jr. y O. B. Ramsay, *Chem. Rev.* **64**, 317 (1964)
- [15].- M. J. S. Dewar y D. M. Storch, *Proc. Natl. Acad. Sci. USA*, **82**, 2225 (1985)
- [16].- P. A. Bash, M. J. Field, R. C. Davenport, G. A. Petsko, D. Ringe y M. Karplus, *Biochemistry* **30**, 5826 (1991)

- [17].- A. Warshel, J. Åqvist y S. Creighton, *Proc. Natl. Acad. Sci. USA* **86**, 5820 (1989)
- [18].- J. Kraut, *Science* **242**, 533 (1988)
- [19].- Z. Peng y K. M. Merz, Jr., *J. Am. Chem. Soc.* (en prensa)
- [20].- H. M. Kalckar. *Chem. Rev.* **28**, 71 (1941)
- [21].- A. Pullman y B. Pullman, *Quantum Biochemistry*, Interscience, Nueva York (1963)
- [22].- P. George, R. J. Witonsky, M. Trachtman, C. Wu, W. Dorwart, L. Richman, W. Richman, F. Shurayh y B. Lentz, *Biochim, Biophys. Acta* **223**, 1 (1970)
- [23].- L. De Meis, *J. Biol. Chem.* **259**, 6090 (1984)
- [24].- L. De Meis, *Biochim, Biophys. Acta* **973**, 333 (1989)
- [25].- P. J. Romero y L. De Meis, *J. Biol. Chem.* **264**, 7896 (1989)
- [26].- L. R. Pratt, *Annu. Rev. Phys. Chem.* **36**, 433 (1985)

**LA REACCIÓN DE HIDRÓLISIS  
DEL PIROFOSFATO EN LA  
SOLUCIÓN ACUOSA**

## 5.1 La estimación de las energías de hidratación

### 5.1.1 Las aproximaciones termodinámicas

Las energías de hidratación del ortofosfato y del pirofosfato fueron estimadas por George *et al.* [1], usando cálculos de energías en redes cristalinas y medidas de calores de solución y de sublimación, dando un valor para el ortofosfato de  $\Delta E^{hid}(H_3PO_4) = -26 \text{ kcal/mol}$ , y para el pirofosfato de  $\Delta E^{hid}(H_4P_2O_7) = -34 \text{ kcal/mol}$ . De hecho, este último valor se obtiene a partir de considerar a la reacción como termoneutra en la fase gaseosa. En cambio, si se usa la entalpía aquí obtenida de  $\Delta H_g = -3.8 \text{ kcal/mol}$ , entonces la energía de hidratación del pirofosfato resulta ser de  $\Delta E^{hid}(H_4P_2O_7) = -38 \text{ kcal/mol}$ . Estas energías podrían resultar subestimadas y están sujetas a una incertidumbre de  $\sim \pm 10 \text{ kcal/mol}$  [2]. De cualquier manera, proporcionan un criterio cualitativo sobre las magnitudes de las energías de hidratación de los anhídridos de fosfato. Además, ésta es la única estimación reportada sobre la hidratación de las moléculas neutras. Los cálculos de Hayes *et al.* [3] a este respecto, se refieren únicamente a especies iónicas.

### 5.1.2 La aproximación por cálculos moleculares

Una forma alternativa de estudiar la hidratación de las moléculas, consiste en realizar cálculos *ab initio* de pares  $H_2O$  - molécula, en distintas configuraciones, con el fin de determinar las posiciones relativas más estables y de localizar las de energía mínima. Este procedimiento fue empleado por el grupo de Pullman para el estudio de la hidratación del anión de dimetilfosfato [4] y del anión  $H_2PO_4^-$  [5]. Los cálculos fueron hechos con bases STO-3G, dado el enorme esfuerzo computacional que representa el mapear la superficie de potencial. Como ellos mismos lo mencionan, la imagen obtenida a partir de este tipo de aproximación es sólo cualitativamente correcta, pero los valores numéricos no son muy confiables. En todo caso, encuentran que las configuraciones de energía mínima corresponden a las interacciones de los hidrógenos del agua con las cargas negativas en los oxígenos de los fosfatos, en posiciones que pueden considerarse como puentes de hidrógeno reforzados. Más aún, encuentran una estructura muy estable cuando la molécula de agua se orienta de tal modo que sus dos hidrógenos forman puentes con los dos oxígenos del fosfato. Las configuraciones elegidas en esta tesis para muestrear la superficie de potencial  $H_2O - H_3PO_4$ , están basadas en las obtenidas por Pullman [5] para el anión  $H_2PO_4^-$ .

### 5.1.3 La aproximación por simulaciones numéricas

El muestreo de la superficie de potencial con el método arriba descrito, es por supuesto incompleto, y depende en gran medida de la intuición química de la persona que realice los cálculos; pero además presenta el problema de omitir los efectos colectivos y los térmicos, que pueden ser significativos en soluciones acuosas [6, 7]. Aquí adquiere mayor relevancia la aproximación por las simulaciones numéricas. Éstas se están aplicando cada vez con mayor frecuencia y mejor eficiencia, a la descripción de sistemas relevantes en la bioquímica (c. g. [8, 9]), gracias al incremento en las capacidades de cómputo, que siguen mejorando continuamente.

En una simulación MC de la hidratación del anión de dimetilfosfato, Alagona *et al.*

[10] encuentran estructuras similares a las de Pullman [4]; pero no la del doble puente de hidrógeno. En el trabajo de Alagona, se emplearon los modelos TIPS2 [11] y TIP4P [12] para el agua líquida, y potenciales similares para la interacción con el dimetilfosfato, todos con la suposición de aditividad por pares [13]. La simulación se hace con 216 moléculas de agua en un ensemble NPT, a  $T = 298 K$ ,  $P = 1 atm$  y con condiciones de frontera periódicas. Se calcula la energía libre de Helmholtz de la hidratación, por medio de una integral que depende de la diferencia en energía interna  $\Delta U$  entre el sistema con el ión, y el sistema de referencia (agua pura). Para evaluar esta integral no es conveniente usar el muestreo convencional de Metropolis; sino que se emplea el *muestreo de sombrilla (umbrella sampling)*, que está diseñado para favorecer las configuraciones que contribuyan en mayor medida a la integral mencionada [14].

La energía de hidratación que calculan para el anión, es de entre  $-95 kcal/mol$  y  $-65 kcal/mol$ , dependiendo de la conformación del dimetilfosfato. El análisis de las estructuras de hidratación se efectúa por medio de las funciones de distribución radial (véase la sección 3.5 y [15]), y de este análisis Alagona *et al.* concluyen que el agua se agrupa alrededor del dimetilfosfato en cuatro formas distintas: (1) las que se ligan fuertemente a los oxígenos  $O^{\delta-}$ , cada uno de los cuales tiene un número de coordinación de 3; (2) las que se ligan por puentes de hidrógeno al oxígeno del éster; (3) las aguas ligadas entre ellas por el efecto hidrofóbico del grupo  $CH_3$ , y (4) las aguas sueltas (*bulk*), que se comportan igual que el agua en simulaciones del líquido puro. Tanto Pullman [4] como Alagona [10] concluyen que el efecto del soluto en la estructura del solvente, alcanza un máximo de dos capas de hidratación. Las interacciones entre las moléculas de agua del tipo (1), son más débiles que en el caso del agua pura, ya que forman menos puentes de hidrógeno. Las interacciones entre las del tipo (2) son muy similares a las del tipo (4), es decir, parecidas a las del agua pura, y finalmente las interacciones entre las moléculas del tipo (3) son más intensas que las otras, debido a la estructura ordenada impuesta por el efecto hidrofóbico del grupo  $CH_3$ .

## 5.2 Descripción del método desarrollado en esta tesis

En tanto que los cálculos *ab initio* producen resultados correspondientes a la fase gaseosa, las simulaciones numéricas permiten estudiar el estado líquido y las interacciones entre los solutos y los solventes. Para esto es necesario contar con potenciales analíticos que permitan evaluar confiablemente tales interacciones. Estos potenciales son funciones analíticas, cuyos parámetros se ajustan para reproducir algunos datos experimentales, o bien para reproducir superficies de energías de interacción calculadas *ab initio*. La primera aproximación ha sido empleada por diversos autores [16, 17], aplicándola principalmente al caso del agua. Ya que en esta aproximación hay que hacer muchas consideraciones empíricas, no se tiene información sobre las posibles fuentes de error al encontrar una situación donde los resultados de las simulaciones no coincidan con los del experimento [18]. En cambio, al partir de cálculos cuánticos es relativamente sencillo detectar las limitaciones del modelo y, en algunos casos, incluso superarlas. Este ha sido el caso de los potenciales desarrollados para la descripción del agua líquida, que, con la inclusión de los efectos no aditivos de tres y de cuatro cuerpos, han dado resultados sobre la estructura

mucho mejores que los modelos empíricos (ver *e. g.* [19-21] y Apéndice C).

En esta tesis además, se propone que es posible llegar a una descripción a nivel molecular de diversos sistemas bioquímicos, partiendo de primeros principios. Por ello se usan potenciales ajustados a superficies de energías de interacción calculadas *ab initio*. El procedimiento consiste en realizar varias decenas de cálculos *ab initio* de la interacción de un par de moléculas en distintas configuraciones, que son elegidas buscando las posiciones en las que la energía de interacción tenga mínimos, y explorando las regiones cercanas. También se aprovechan las simetrías de las moléculas, de manera que un solo cálculo corresponda a varias configuraciones. Este muestreo no corresponde necesariamente a las posiciones relativas de las moléculas que se encontrarán en la fase líquida (ver *e. g.* [22]), pero constituye una primera aproximación razonable.

Además de una superficie de interacciones por pares, para el caso de las soluciones acuosas es necesario contar con otra de contribuciones no aditivas a las interacciones de al menos tres cuerpos [7, 22, 23]. Una vez que se cuenta con estas superficies, se ajustan a ellas los parámetros de una forma funcional que representa a un modelo molecular. En el caso presente se usa el modelo MCHO [21], que ha dado los mejores resultados estructurales que se han reportado sobre el agua líquida. Este modelo fue desarrollado como parte del trabajo presente, y se anexa en el apéndice C. El potencial MCHO también ha sido aplicado al estudio de la tautomerización del uracilo [24], y al de las interacciones de la urea con el agua [25]. De hecho, los parámetros correspondientes a la molécula de agua se mantienen fijos, ajustando únicamente los de las otras especies moleculares.

La forma funcional del potencial MCHO entre dos átomos es la siguiente:

$$V_{ij} = A \exp(-n_1 R_{IJ}) + B \exp(-n_2 R_{IJ}) + C \frac{\delta_i \delta_j}{r_{ij}} + a_i r_{ij}^2 + a_j r_{ij}^2 + k r_{ij}^{-n_3} \quad (5-1).$$

La descripción del modelo, y de sus parámetros, se presenta en el apéndice C.

Al tener los potenciales analíticos, se comienzan las primeras simulaciones que ya van a producir información sobre el comportamiento de la solución. Sin embargo, se ha encontrado conveniente revisar las configuraciones obtenidas en la simulación, con el fin de refinar la superficie de energías, haciendo cálculos *ab initio* de los pares y los tríos de moléculas que más se repitan durante la simulación [22]; con lo cual también se refina el potencial analítico, al reajustar los parámetros a la superficie aumentada.

En esta tesis se ajustaron los parámetros del potencial MCHO a las siguientes superficies de interacción: (1)  $H_3PO_4 - H_2O$ , (2)  $H_3PO_4 - 2H_2O$ , (3)  $H_4P_2O_7 - H_2O$  y (4)  $H_4P_2O_7 - 2H_2O$ . Con los potenciales obtenidos se realizaron simulaciones de sistemas de agua con ortofosfato y de agua con pirofosfato. De estas simulaciones se obtuvieron resultados preliminares sobre las energías de hidratación de una y otra especie, y con ellas se estimó la contribución de la hidratación a la entalpía de la reacción (4-1).

### 5.3 La interacción del ortofosfato con el agua

Las energías de todas las interacciones se calcularon con la base (V), descrita en la sección 4.1, y sólo al nivel SCF, debido al costo computacional que representa obtener una superficie de potencial. En este caso se calcularon las energías del dímero  $H_3PO_4 - H_2O$  en 96 posiciones relativas distintas, y las no aditvidades de trímeros  $H_3PO_4 - 2H_2O$ , en 23 geometrías.

Las interacciones de los pares  $H_3PO_4 - H_2O$  con energías más bajas, corresponden a configuraciones con puentes de hidrógeno (HB) entre los oxígenos de las aguas y los hidrógenos del ortofosfato, y el valor mínimo es de -9.0 kcal/mol al nivel SCF. Al agregar la contribución de la energía de correlación el valor es de -10.5 kcal/mol. Esto da una idea del error producido por la omisión de la correlación, que podría incluirse para el ortofosfato, pero que resulta prohibitivamente costosa para el pirofosfato, por lo que se decidió mantener el mismo nivel de aproximación para ambas especies.

En la fig. 5-1 se muestran algunas configuraciones empleadas para el cálculo de la superficie, y se indica la de mínima energía. En la Tabla 5-1 se presentan las energías de interacción calculadas para los dímeros de ortofosfato con agua, las no aditvidades para los trímeros  $H_3PO_4 - 2H_2O$ , y los valores que da el potencial ajustado. El ajuste se hizo como se indica en el apéndice C (referencia [21]), usando un programa desarrollado en el laboratorio de Biofísica del IFUNAM [26]. Las coordenadas de las moléculas de  $H_3PO_4$  y de  $H_2O$  que corresponden a las energías de interacción en la Tabla 5-1 están reportadas en el apéndice B.

En la Tabla 5-2 se presentan los valores de los parámetros ajustados a la interacción agua - agua [27] y los de la interacción ortofosfato - agua.

La fig. 5-2 muestra la calidad del ajuste a las energías de los 96 pares  $H_3PO_4 - H_2O$  ( $m = 0.9988$ ,  $b = -0.0019$ ,  $\rho = 0.9995$ ). El error cuadrático medio del ajuste fue de  $e_{rms} = 0.0782$  kcal/mol. En la fig. 5-3 se puede observar el ajuste a las no aditvidades de los tríos  $H_3PO_4 - 2H_2O$  ( $m = 0.8621$ ,  $b = -0.0377$ ,  $\rho = 0.9649$ ), con un error de  $e_{rms} = 0.0802$  kcal/mol. Cuanto mejor es el ajuste logrado tanto más cerca de tener pendiente  $m = 1$  y ordenada al origen  $b = 0$  está la recta en las gráficas. Puede concluirse que el ajuste alcanzado es muy bueno.

FIGURA 5-1

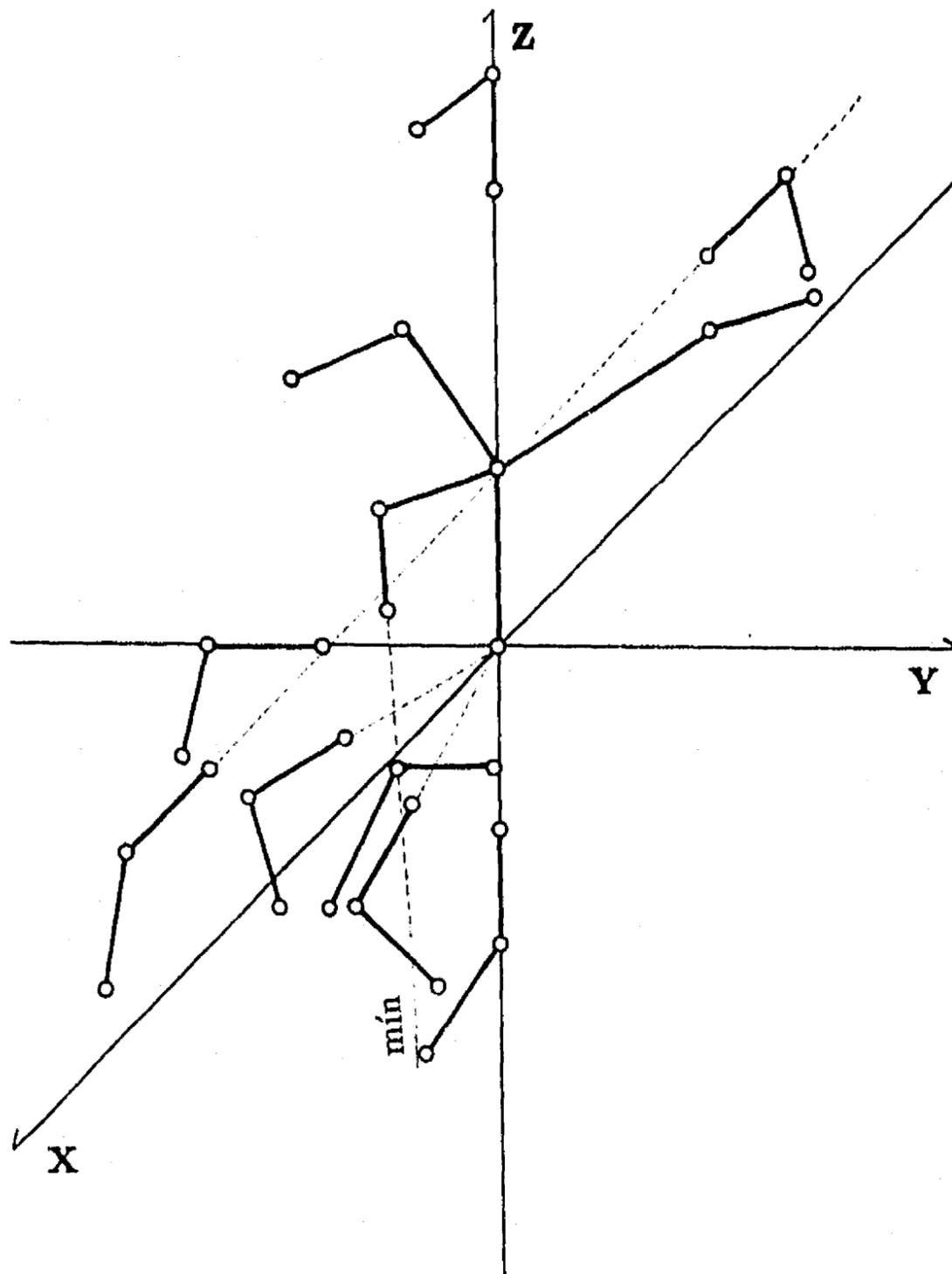


Figura 5-1. Configuraciones usadas para calcular la superficie de interacción del ortofosfato con el agua. A lo largo de la línea indicada con mín se encuentra el mínimo de energía, y corresponde a la configuración 59.

TABLA 5-1

Energías de interacción por pares (kcal/mol)									
conf.	SCF	MCHO	$e_a$	$e_r$	conf.	SCF	MCHO	$e_a$	$e_r$
1	0.983773	0.947349	0.04	0.04	2	-5.266400	-5.364270	0.10	-0.02
3	-6.018010	-6.042220	0.02	-0.00	4	-5.410090	-5.408640	-0.00	0.00
5	-4.693510	-4.686410	-0.01	0.00	6	-3.045790	-3.053970	0.01	-0.00
7	-2.306430	-2.342910	0.04	-0.02	8	-1.921470	-1.969360	0.05	-0.02
9	-1.270560	-1.296110	0.03	-0.02	10	-0.899469	-0.871635	-0.03	0.03
11	0.977766	0.938330	0.04	0.04	12	-5.271050	-5.371030	0.10	-0.02
13	-6.021830	-6.047780	0.03	-0.00	14	6.755260	6.673860	0.08	0.01
15	2.063470	2.059750	0.00	0.00	16	-1.860760	-1.883250	0.02	-0.01
17	-1.655940	-1.675900	0.02	-0.01	18	-1.743200	-1.773390	0.03	-0.02
19	-1.439970	-1.493220	0.05	-0.04	20	6.783630	6.808360	-0.02	-0.00
21	-1.228940	-1.294310	0.07	-0.05	22	2.103470	2.155100	-0.05	-0.02
23	-0.814273	-0.863525	0.05	-0.06	24	-1.628050	-1.638350	0.01	-0.01
25	-0.650175	-0.667917	0.02	-0.03	26	2.570990	2.559460	0.01	0.00
27	-1.002030	-1.149350	0.15	-0.15	28	0.091230	0.105727	-0.01	-0.16
29	-5.612610	-5.523390	-0.09	0.02	30	-6.816400	-6.768920	-0.05	0.01
31	-5.950540	-5.944740	-0.01	0.00	32	-6.756190	-6.767200	0.01	-0.00
33	-7.230340	-7.143290	-0.09	0.01	34	-6.572770	-6.513640	-0.06	0.01
35	-5.992530	-6.081300	0.09	-0.01	36	-6.244720	-6.145900	-0.10	0.02
37	-5.845820	-5.763250	-0.08	0.01	38	-4.913370	-5.003250	0.09	-0.02
39	-5.341340	-5.227420	-0.11	0.02	40	-5.076180	-4.983210	-0.09	0.02
41	-3.258890	-3.338170	0.08	-0.02	42	-4.083970	-3.952810	-0.13	0.03
43	-3.968550	-3.873870	-0.09	0.02	44	-2.310710	-2.396320	0.09	-0.04
45	-3.370980	-3.246980	-0.12	0.04	46	-3.329100	-3.250470	-0.08	0.02
47	5.897100	6.021800	-0.12	-0.02	48	2.300880	2.166570	0.13	0.06
49	1.294340	1.248860	0.05	0.04	50	0.865845	0.900843	-0.03	-0.04
51	2.626300	2.691600	-0.07	-0.02	52	0.010899	0.002607	0.01	0.76
53	-0.015071	-0.067999	0.05	-3.51	54	0.205821	0.136663	0.07	0.34
55	0.264590	0.249970	0.01	0.06	56	0.224966	0.281894	-0.06	-0.25
57	0.171090	0.275804	-0.10	-0.61	58	-5.781950	-5.778190	-0.00	0.00
59	-8.989020	-8.998490	0.01	-0.00	60	4.150040	4.143510	0.01	0.00
61	6.104180	6.008540	0.10	0.02	62	-8.481750	-8.421400	-0.06	0.01
63	1.373410	1.403340	-0.03	-0.02	64	2.618810	2.533660	0.09	0.03
65	-6.748180	-6.721540	-0.03	0.00	66	0.824638	0.943456	-0.12	-0.14
67	1.753460	1.750070	0.00	0.00	68	-5.593370	-5.592040	-0.00	0.00
69	0.695227	0.828040	-0.13	-0.19	70	1.274040	1.324780	-0.05	-0.04
71	-4.151250	-4.167660	0.02	-0.00	72	0.772838	0.854604	-0.08	-0.11
73	1.178760	1.220050	-0.04	-0.04	74	-3.382260	-3.413850	0.03	-0.01
75	0.871074	0.885418	-0.01	-0.02	76	1.142940	1.157330	-0.01	-0.01
77	-2.466600	-2.535990	0.07	-0.03	78	0.967309	0.902664	0.06	0.07
79	1.100680	1.080010	0.02	0.02	80	-1.997490	-2.087170	0.09	-0.04
81	0.968374	0.882230	0.09	0.09	82	1.023990	0.999384	0.02	0.02

TABLA 5-1 (continuación)

Energías de interacción por pares (kcal/mol)									
conf.	SCF	MCHO	$\epsilon_a$	$\epsilon_r$	conf.	SCF	MCHO	$\epsilon_a$	$\epsilon_r$
83	0.927210	0.853631	0.07	0.08	84	0.954566	0.942843	0.01	0.01
85	-0.445351	-0.509536	0.06	-0.14	86	-6.358740	-6.321090	-0.04	0.01
87	-6.878070	-6.791700	-0.09	0.01	88	-6.045340	-5.933530	-0.11	0.02
89	-5.228960	-5.099670	-0.13	0.02	90	8.221480	8.520980	-0.30	-0.04
91	-0.194694	-0.251405	0.06	-0.29	92	-1.822460	-1.959930	0.14	-0.08
93	-2.203390	-2.373250	0.17	-0.08	94	-2.001920	-2.141830	0.14	-0.07
95	-1.571970	-1.643990	0.07	-0.05	96	-0.680002	-0.583870	-0.10	0.14

Contribuciones no aditivas de tres cuerpos (kcal/mol)									
conf.	SCF	MCHO	$\epsilon_a$	$\epsilon_r$	conf.	SCF	MCHO	$\epsilon_a$	$\epsilon_r$
1	-0.458052	-0.327195	-0.13	0.29	2	0.391662	0.219504	0.17	0.44
3	0.260101	0.143831	0.12	0.45	4	0.210529	0.112573	0.10	0.47
5	0.161315	0.081326	0.08	0.50	6	0.137255	0.066905	0.07	0.51
7	0.107328	0.051084	0.06	0.52	8	-1.046030	-1.107220	0.06	-0.06
9	-0.484229	-0.529289	0.05	-0.09	10	-0.363782	-0.383030	0.02	-0.05
11	-0.252384	-0.249717	-0.00	0.01	12	-0.201114	-0.193201	-0.01	0.04
13	-0.141043	-0.135808	-0.01	0.04	14	-0.620127	-0.557748	-0.06	0.10
15	-0.423657	-0.349784	-0.07	0.17	16	-0.340726	-0.265718	-0.08	0.22
17	-0.253987	-0.183400	-0.07	0.28	18	-0.211085	-0.146739	-0.06	0.30
19	-0.131114	-0.090333	-0.04	0.31	20	0.439260	0.299959	0.14	0.32
21	0.252900	0.182159	0.07	0.28	22	0.213725	0.143212	0.07	0.33
23	0.170298	0.124167	0.05	0.27					

Tabla 5-1. Energías de interacción del ortofosfato con el agua (en kcal/mol), calculadas *ab initio* (columnas 2 y 7), y con el modelo MCHO (columnas 3 y 8). También se presentan los errores absoluto,  $\epsilon_a$ , (columnas 4 y 9) y relativo,  $\epsilon_r$ , (columnas 5 y 10), con respecto al valor *ab initio*. Las configuraciones correspondientes (columnas 1 y 6), se reportan en el apéndice B. El valor mínimo de energía ocurre en la configuración 59.

TABLA 5-2

I. Parámetros de pares de átomos						
<b>a. Interacción intermolecular agua - agua</b>						
	$O_w - O_w$	$O_w - H_w$	$H_w - H_w$			
A	$10^9$	999.9427	-1422.459			
B	$-10^9$	-749.6233	$2.18874 \times 10^{15}$			
$n_1$	2.934434	1.040209	2.317225			
$n_2$	2.941385	0.9693737	10.61030			
C	1.0	1.0	1.0			
k	2.558486	2615.392	-0.012511677			
$n_3$	5.645374	5.150225	5.763937			
<b>b. Interacción intramolecular del agua</b>						
		$O_w - H_w$	$H_w - H_w$			
	C	25.37466	1.064440			
	k	13962.38	0.08711197			
	$n_3$	1.011549	1.026989			
<b>c. Interacción intermolecular agua - ortofosfato</b>						
	$O_w - P_o$	$O_w - O_o$	$O_w - H_o$	$H_w - P_o$	$H_w - O_o$	$H_w - H_o$
A	1002.271	385.3814	1054.462	7222.559	3931.750	444.0482
B	-1.389745	-110.3132	-55.40633	-1.779791	-13.27000	-230.3417
$n_1$	1.141635	1.344110	1.556246	1.769200	1.867020	1.512459
$n_2$	0.3458269	0.8303445	1.004466	0.2757602	1.578108	1.204592
C	1.0	1.0	1.0	1.0	1.0	1.0
k	0.6001643	0.4224089	24067.94	0.3895013	0.5639487	0.2686426
$n_3$	1.811529	2.756796	6.693404	0.7203396	1.531071	4.129292
<b>d. Interacción intramolecular del ortofosfato</b>						
	$P_o - O_o$	$P_o - H_o$	$O_o - O_o$	$O_o - H_o$	$H_o - H_o$	
C	1.042091	0.4479815	0.2494013	1.387665	0.8228134	
k	118.7550	1.308975	1.055457	89.52290	0.6910009	
$n_3$	1.454920	0.7771192	0.9548834	4.368043	0.6762414	
<b>II. Parámetros atómicos (cargas y constantes de fuerza restitutiva)</b>						
	$O_w$	$H_w$	$P_o$	$O_o$	$H_o$	
a	60.09528	172.5012	639.6005	423.6041	2448.076	
$\delta$	-1.3349146	0.6674573	2.167303	-1.253958	0.9520311	

Tabla 5-2. Parámetros del potencial MCHO para las interacciones agua - agua y ortofosfato - agua. Las cargas están en u. a., y los demás parámetros dan energías en kcal/mol cuando las distancias están en u. a. Los coeficientes de las interacciones electrostáticas deben multiplicarse por 627.5. La pertenencia de un átomo a la molécula del agua o a la del ortofosfato, se indica con los subíndices  $w$  y  $o$ , respectivamente.

FIGURA 5-2

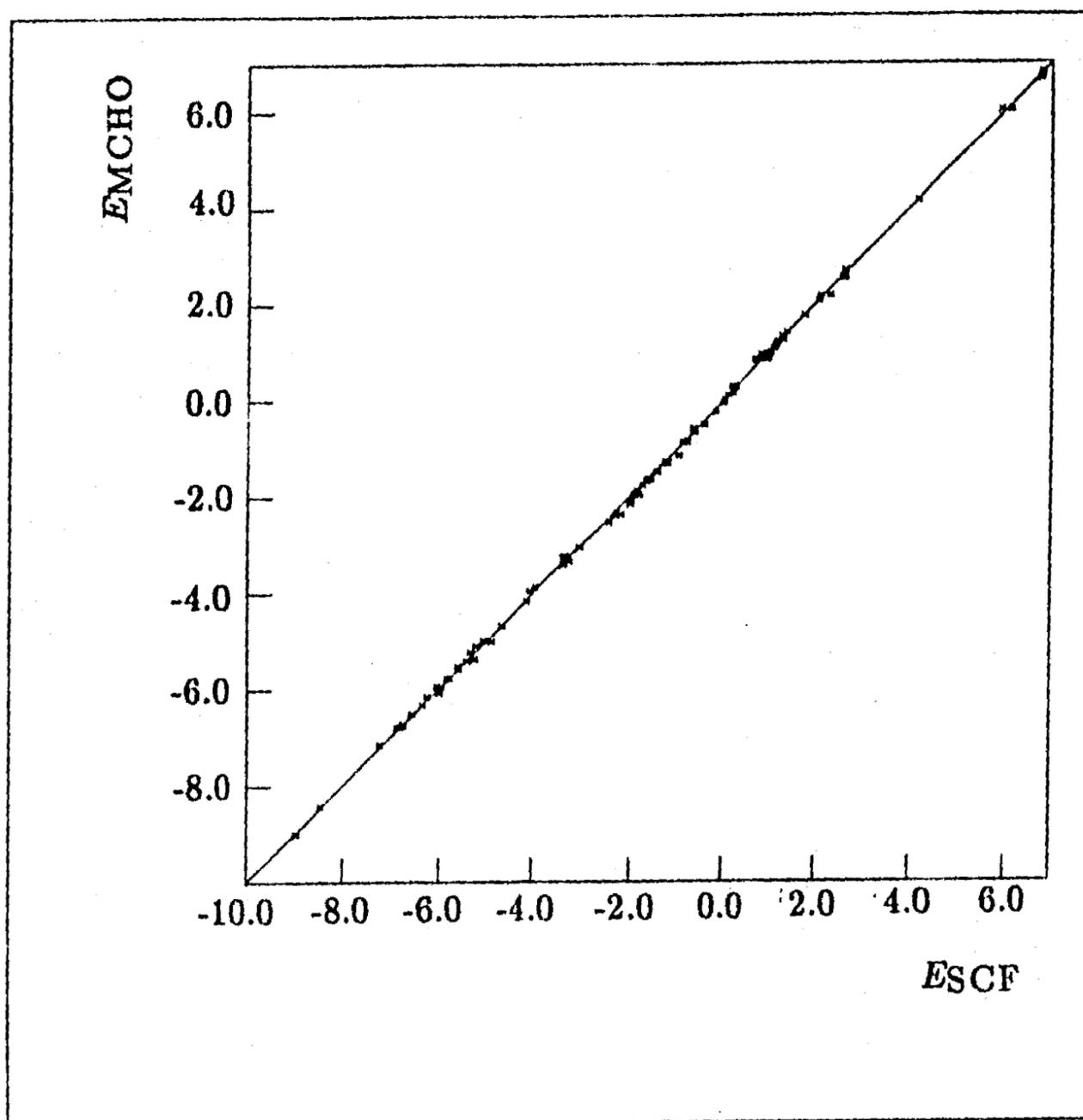


Figura 5-2. Comparación de los valores de las energías por pares  $H_3PO_4 - H_2O$  calculados *ab initio* (abscisas) y por el modelo MCHO (ordenadas). Las unidades son kcal/mol.

FIGURA 5-3

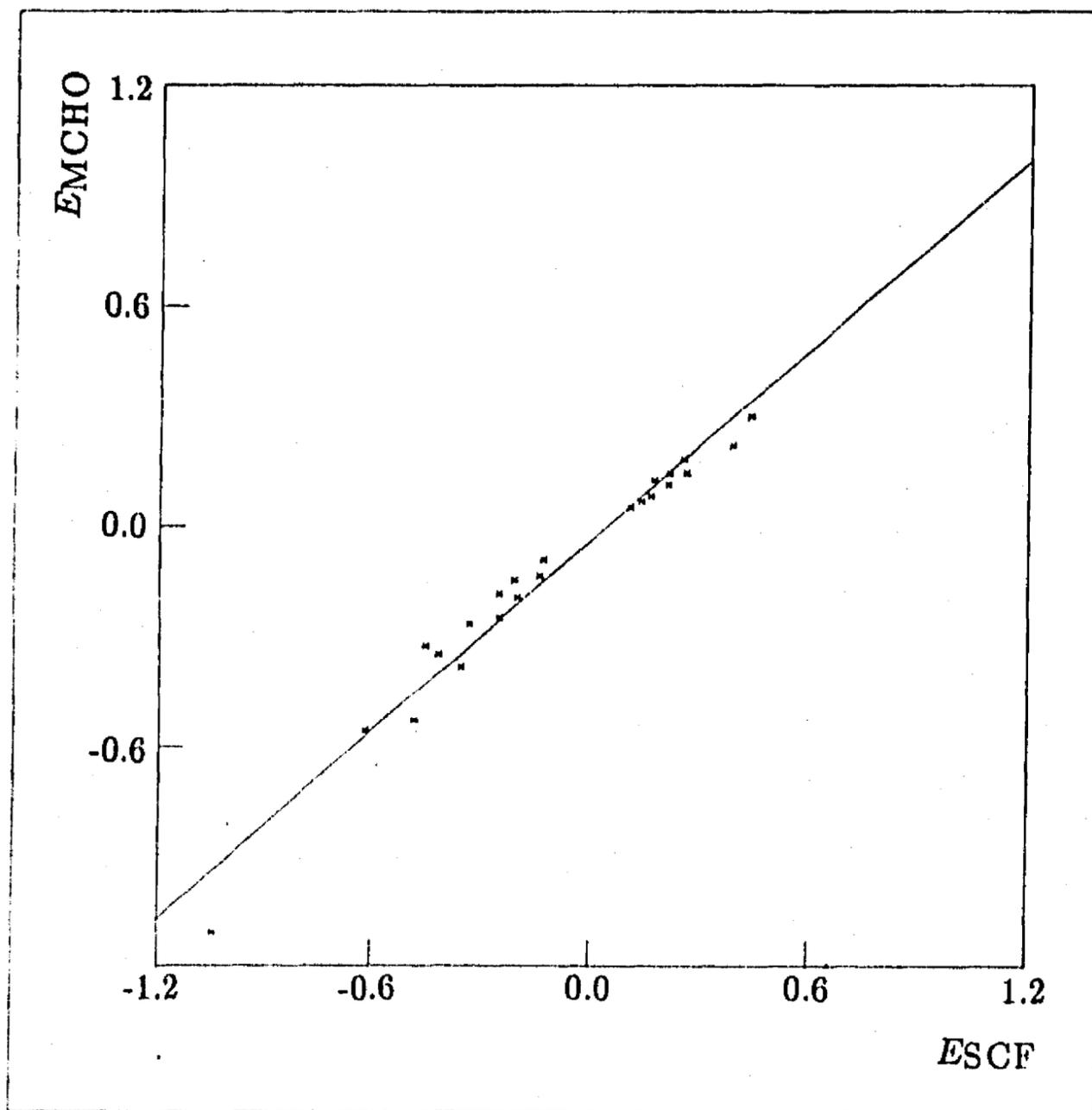


Figura 5-3. Comparación de los valores de las contribuciones no aditivas a las energías de interacción de trímeros  $H_3PO_4 - 2H_2O$  calculados *ab initio* (abscisas) y por el modelo MCHO (ordenadas). Las unidades son kcal/mol.

Con el potencial así obtenido se efectuó una simulación MC de ortofosfato en agua líquida a  $T = 298 K$ , con una molécula de ortofosfato fija en el centro de una celda computacional cúbica de 41.1029 u.a. por lado, que también se utiliza para simulaciones de agua pura. En este último caso se emplean 343 moléculas, con lo cual se mantiene la densidad experimental de  $\rho = 0.997 g/cm^3$ . En cambio, al incluir la molécula de  $H_3PO_4$ , es necesario eliminar de la celda algunas moléculas de agua. El criterio usado para esta eliminación se basa en las distancias interatómicas más cortas para las que se calcularon las energías, en la superficie a la cual se ajustó el potencial MCHO. Las distancias mínimas de acercamiento entre los átomos del ortofosfato y del agua se muestran en la Tabla 5-3.

TABLA 5-3

Atomo	$O_w$	$H_w$
$O_w$	4.5	
$H_w$	2.0	2.0
$P_o$	3.5	3.5
$O_o$	2.0	3.5
$H_o$	2.0	2.5

Tabla 5-3. Distancias mínimas de acercamiento entre los átomos de  $H_3PO_4$  y de  $H_2O$  (en u. a.). A distancias menores el potencial tiene valores de más de 10 kcal/mol. Los subíndices  $w$  y  $o$  señalan a cuál molécula, agua u ortofosfato, pertenece cada átomo.

A partir de varias configuraciones ya equilibradas de agua líquida, se eliminaron consistentemente tres moléculas de agua, cada vez que se aplicó el método descrito arriba. Se hizo la simulación con las 340 moléculas restantes y la de ortofosfato. Para obtener la energía de hidratación, se efectuaron algunos millones de configuraciones, con el fin de llegar a un equilibrio. A continuación se hizo el promedio de las energías del sistema, y se comparó con el de una simulación MC de agua pura, usando la ec. (3-47),

$$\Delta E^{hid}(H_3PO_4) = \langle E_o \rangle - \frac{N_w - N_o}{N_w} \langle E_w \rangle \quad (3-47).$$

Comúnmente se emplean alrededor de  $2 \times 10^6$  configuraciones para equilibrar un sistema, y luego se registran los datos para hacer la estadística por otras  $4 \times 10^6$  configuraciones [10]. En este caso se procedió de igual manera, y se obtuvieron valores para la energía de hidratación del ortofosfato de  $\sim -100 kcal/mol$ , con una dispersión muy grande. Ya que ésta parecía una energía de hidratación exagerada para una especie neutra, se decidió examinar el sistema de agua pura y se hizo una simulación más larga, con la finalidad de verificar la  $\langle E_w \rangle$  que debería restarse en la ec. (3-47). Se encontraron resultados que ameritan una discusión aparte, la cual se presenta en la sección 5.5.

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

#### 5.4 La interacción del pirofosfato con el agua

Con el mismo procedimiento que para el caso del ortofosfato, se calcularon las energías de 100 pares  $H_4P_2O_7 - H_2O$ , en las configuraciones esbozadas en la fig. 5-4, y las no aditivities de 21 tríos  $H_4P_2O_7 - 2H_2O$ .

Las interacciones de los pares con energías más bajas, corresponden a configuraciones con puentes de hidrógeno (HB) entre los oxígenos del agua y los hidrógenos del pirofosfato, y el valor mínimo de los calculados es de -10.9 kcal/mol. En la Tabla 5-4 se presentan los valores calculados de las energías y las no aditivities de la interacción del pirofosfato con el agua, así como los valores producidos por el modelo. Las coordenadas de las moléculas de  $H_4P_2O_7$  y de  $H_2O$  que corresponden a las energías de interacción en la Tabla 5-4 están reportadas en el apéndice B.

FIGURA 5-4

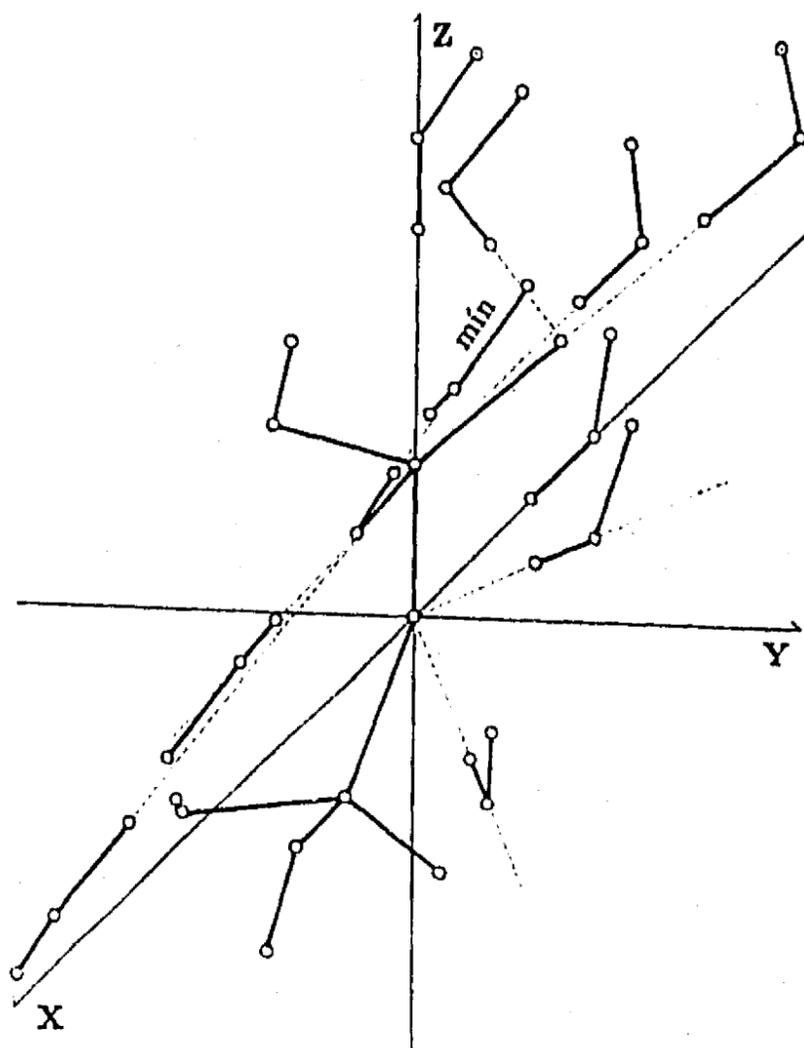


Figura 5-4. Configuraciones usadas para calcular la superficie de interacción del pirofosfato con el agua. A lo largo de la línea indicada con *min* se encuentra el mínimo de energía, que corresponde a la configuración 37.

TABLA 5-4

Energías de interacción por pares (kcal/mol)									
conf.	SCF	MCHO	$e_a$	$e_r$	conf.	SCF	MCHO	$e_a$	$e_r$
1	-0.240691	-0.166744	-0.07	0.31	2	-1.445010	-1.490030	0.05	-0.03
3	-1.571860	-1.644880	0.07	-0.05	4	-1.452910	-1.509420	0.06	-0.04
5	-1.087220	-1.090040	0.00	-0.00	6	1.854140	1.959680	-0.11	-0.06
7	-2.277020	-2.922590	0.65	-0.28	8	-5.419510	-5.425790	0.01	-0.00
9	-5.299450	-5.309520	0.01	-0.00	10	-5.014220	-5.028490	0.01	-0.00
11	-4.858270	-4.872750	0.01	-0.00	12	-4.534060	-4.546880	0.01	-0.00
13	-3.032800	-3.033820	0.00	-0.00	14	-2.555660	-2.563340	0.01	-0.00
15	1.967590	1.982190	-0.01	-0.01	16	-4.661690	-4.662140	0.00	-0.00
17	-5.586310	-5.573770	-0.01	0.00	18	-5.419600	-5.425820	0.01	-0.00
19	-4.538570	-4.546880	0.01	-0.00	20	-3.591030	-3.588310	-0.00	0.00
21	-3.037210	-3.033780	-0.00	0.00	22	0.615113	0.573356	0.04	0.07
23	-2.976030	-2.760200	-0.22	0.07	24	-4.141560	-3.981990	-0.16	0.04
25	-4.110460	-4.069890	-0.04	0.01	26	-3.460380	-3.461130	0.00	-0.00
27	-2.995980	-2.994060	-0.00	0.00	28	5.041130	4.919860	0.12	0.02
29	1.277960	1.233870	0.04	0.03	30	-1.044860	-1.068730	0.02	-0.02
31	-1.576880	-1.582400	0.01	-0.00	32	-1.742370	-1.719950	-0.02	0.01
33	-1.660130	-1.627400	-0.03	0.02	34	-1.512080	-1.480890	-0.03	0.02
35	4.730430	4.529670	0.20	0.04	36	-8.491640	-8.545490	0.05	-0.01
37	-10.91420	-10.87190	-0.04	0.00	38	-10.06570	-10.18350	0.12	-0.01
39	-6.638090	-6.879570	0.24	-0.04	40	-4.106600	-4.236490	0.13	-0.03
41	-3.368570	-3.483830	0.12	-0.03	42	4.941400	5.018070	-0.08	-0.02
43	2.094070	2.169160	-0.08	-0.04	44	1.663960	1.717760	-0.05	-0.03
45	1.509130	1.551020	-0.04	-0.03	46	1.394490	1.425390	-0.03	-0.02
47	7.767790	7.746320	0.02	0.00	48	5.975250	6.017670	-0.04	-0.01
49	4.137340	4.096460	0.04	0.01	50	2.943170	2.838730	0.10	0.04
51	0.786677	0.496071	0.29	0.37	52	-0.773219	-0.846837	0.07	-0.10
53	-1.013000	-1.004110	-0.01	0.01	54	-0.988144	-0.936852	-0.05	0.05
55	6.248080	6.248630	-0.00	-0.00	56	-2.069660	-1.911250	-0.16	0.08
57	-3.621180	-3.522050	-0.10	0.03	58	-3.220740	-3.213250	-0.01	0.00
59	-2.356020	-2.354610	-0.00	0.00	60	-1.831100	-1.828520	-0.00	0.00
61	-1.384640	-1.380620	-0.00	0.00	62	0.917473	0.689715	0.23	0.25
63	-5.370320	-5.308780	-0.06	0.01	64	-6.068870	-5.970440	-0.10	0.02
65	-4.722300	-4.598440	-0.12	0.03	66	-3.687610	-3.536430	-0.15	0.04
67	-3.089980	-2.937550	-0.15	0.05	68	-2.571330	-2.436480	-0.13	0.05
69	2.064370	2.258970	-0.19	-0.09	70	1.970730	1.986980	-0.02	-0.01
71	-1.324870	-1.337100	0.01	-0.01	72	-3.409620	-3.425740	0.02	-0.00
73	-4.660650	-4.659500	-0.00	0.00	74	-5.325250	-5.310670	-0.01	0.00
75	-5.585940	-5.571990	-0.01	0.00	76	-5.583270	-5.578680	-0.00	0.00
77	-5.419610	-5.424650	0.01	-0.00	78	-5.162740	-5.174450	0.01	-0.00
79	-4.858330	-4.871930	0.01	-0.00	80	-4.208350	-4.216350	0.01	-0.00
81	-3.890310	-3.894740	0.00	-0.00	82	-3.586740	-3.588050	0.00	-0.00

TABLA 5-4 (continuación)

	SCF	MCHO	$e_a$	$e_r$		SCF	MCHO	$e_a$	$e_r$
83	-3.300430	-3.300450	0.00	-0.00	84	-3.033050	-3.033700	0.00	-0.00
85	3.146790	2.902750	0.24	0.08	86	-0.189888	-0.435297	0.25	-1.29
87	-2.316170	-2.541870	0.23	-0.10	88	-3.608210	-3.796260	0.19	-0.05
89	-4.313840	-4.469790	0.16	-0.04	90	-4.613880	-4.754560	0.14	-0.03
91	-4.650200	-4.785320	0.14	-0.03	92	-4.523830	-4.655600	0.13	-0.03
93	-4.302830	-4.429680	0.13	-0.03	94	-4.032740	-4.151230	0.12	-0.03
95	-3.741570	-3.849270	0.11	-0.03	96	-3.447000	-3.542760	0.10	-0.03
97	-3.158800	-3.243620	0.08	-0.03	98	-2.883870	-2.958980	0.08	-0.03
99	-2.624880	-2.692780	0.07	-0.03	100	-2.383580	-2.446770	0.06	-0.03
Contribuciones no aditivas de tres cuerpos (kcal/mol)									
	SCF	MCHO	$e_a$	$e_r$		SCF	MCHO	$e_a$	$e_r$
1	-0.703946	-0.644848	-0.06	0.08	2	-0.515624	-0.416031	-0.10	0.19
3	-0.427984	-0.332804	-0.10	0.22	4	-0.274204	-0.216990	-0.06	0.21
5	-0.177759	-0.156897	-0.02	0.12	6	-0.154719	-0.140033	-0.01	0.09
7	0.283019	0.123263	0.16	0.56	8	0.141287	0.033869	0.11	0.76
9	-0.029798	-0.024314	-0.01	0.18	10	-0.037239	-0.019130	-0.02	0.49
11	-0.040796	-0.013584	-0.03	0.67	12	-0.170071	-0.047151	-0.12	0.72
13	-0.211091	-0.073614	-0.14	0.65	14	-0.248860	-0.084646	-0.16	0.66
15	0.242329	0.253685	-0.01	-0.05	16	0.309674	0.315343	-0.01	-0.02
17	0.308620	0.330575	-0.02	-0.07	18	0.024636	0.024351	0.00	0.01
19	0.681904	1.056150	-0.37	-0.55	20	-0.003203	0.059904	-0.06	19.70
21	-0.743446	-0.629748	-0.11	0.15					

Tabla 5-4. Energías de interacción del pirofosfato con el agua (en kcal/mol), calculadas *ab initio* (columnas 2 y 7), y con el modelo MCHO (columnas 3 y 8). También se presentan los errores absoluto,  $e_a$  (columnas 4 y 9) y relativo,  $e_r$  (columnas 5 y 10), con respecto al valor *ab initio*. Las configuraciones correspondientes (columnas 1 y 6), se reportan en el apéndice B. El valor mínimo de energía ocurre en la configuración 37.

También se usó el potencial MCHO y los valores de los parámetros se encuentran en la Tabla 5-5. La reproducción de los valores de energía calculados *ab initio*, puede apreciarse en la fig. 5-5, que corresponde al ajuste a las interacciones por pares ( $m = 0.9993$ ,  $b = -0.0282$ ,  $\rho = 0.9988$ ), con un error de  $e_{rms} = 0.1179$  kcal/mol; y en la fig. 5-6, que muestra el ajuste a las no aditividad de tres cuerpos ( $m = 0.9871$ ,  $b = 0.0534$ ,  $\rho = 0.9151$ ), con un error de  $e_{rms} = 0.1161$  kcal/mol. Al igual que en el caso del ortofosfato, el ajuste es muy bueno.

TABLA 5-5

I. Parámetros de pares de átomos						
a. Interacción intermolecular agua - pirofosfato						
	$O_w - P_p$	$O_w - O_p$	$O_w - H_p$	$H_w - P_p$	$H_w - O_p$	$H_w - H_p$
A	2116.217	12677.78	2279.771	17498.84	2644.206	407.4675
B	-0.8795200	-439.4043	-67.77884	-2.431389	-24.25462	-227.6637
$n_1$	1.349991	1.590736	1.689204	2.103948	1.809027	1.693778
$n_2$	0.6020350	0.9945048	1.016057	0.6465470	1.437973	1.604606
C	1.0	1.0	1.0	1.0	1.0	1.0
k	0.4211711	0.4704457	23901.84	0.2275091	601.9921	0.1895745
$n_3$	1.626258	3.053850	10.12865	0.6047708	6.966639	4.950058
b. Interacción intramolecular del pirofosfato						
	$P_p - P_p$	$P_p - O_p$	$P_p - H_p$	$O_p - O_p$	$O_p - H_p$	$H_p - H_p$
C	0.2088792	0.6101027	0.5038368	0.1757131	0.9061877	0.8305884
k	0.7265451	566.5412	1.445587	0.8952897	103.8502	0.5443260
$n_3$	0.7049666	2.383085	0.8528040	0.7297747	2.301782	0.3447974
II. Parámetros atómicos (cargas y constantes de fuerza restitutiva)						
	$P_p$	$O_p$	$H_p$			
a	121.1618	124.9758	469.3221			
$\delta$	2.138048	-1.066348	0.7930878			

Tabla 5-5. Parámetros del potencial MCHO para las interacciones agua - agua y pirofosfato - agua. Las cargas están en u. a., y los demás parámetros dan energías en kcal/mol cuando las distancias están en u. a. Los coeficientes de las interacciones electrostáticas deben multiplicarse por 627.5. Los subíndices  $w$  y  $p$  señalan a cuál molécula, agua o pirofosfato, pertenece cada átomo.

FIGURA 5-5

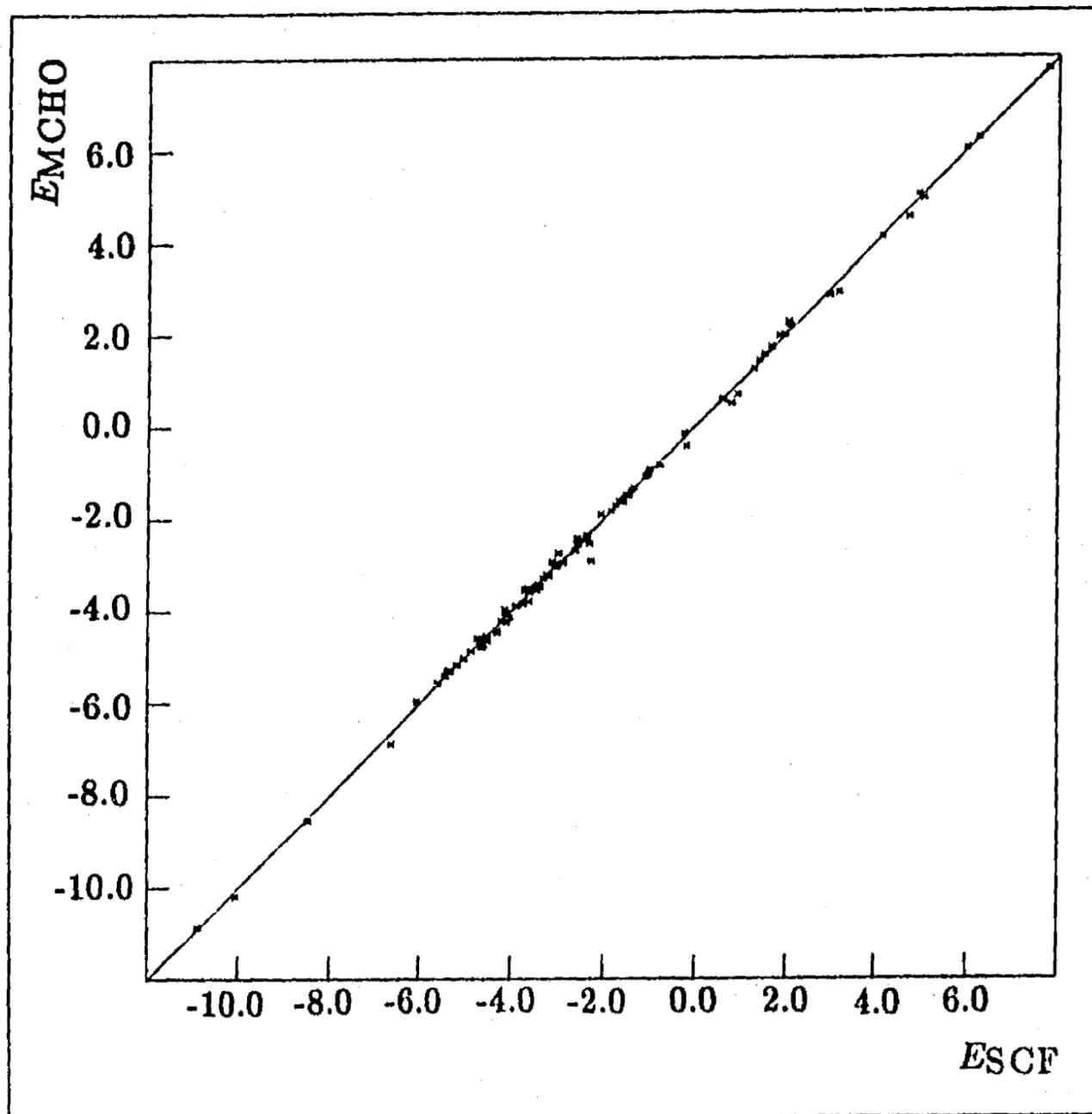


Figura 5-5. Comparación de los valores de las energías por pares  $H_4P_2O_7 - H_2O$  calculados *ab initio* (abscisas) y por el modelo MCHO (ordenadas). Las unidades son kcal/mol.

FIGURA 5-6

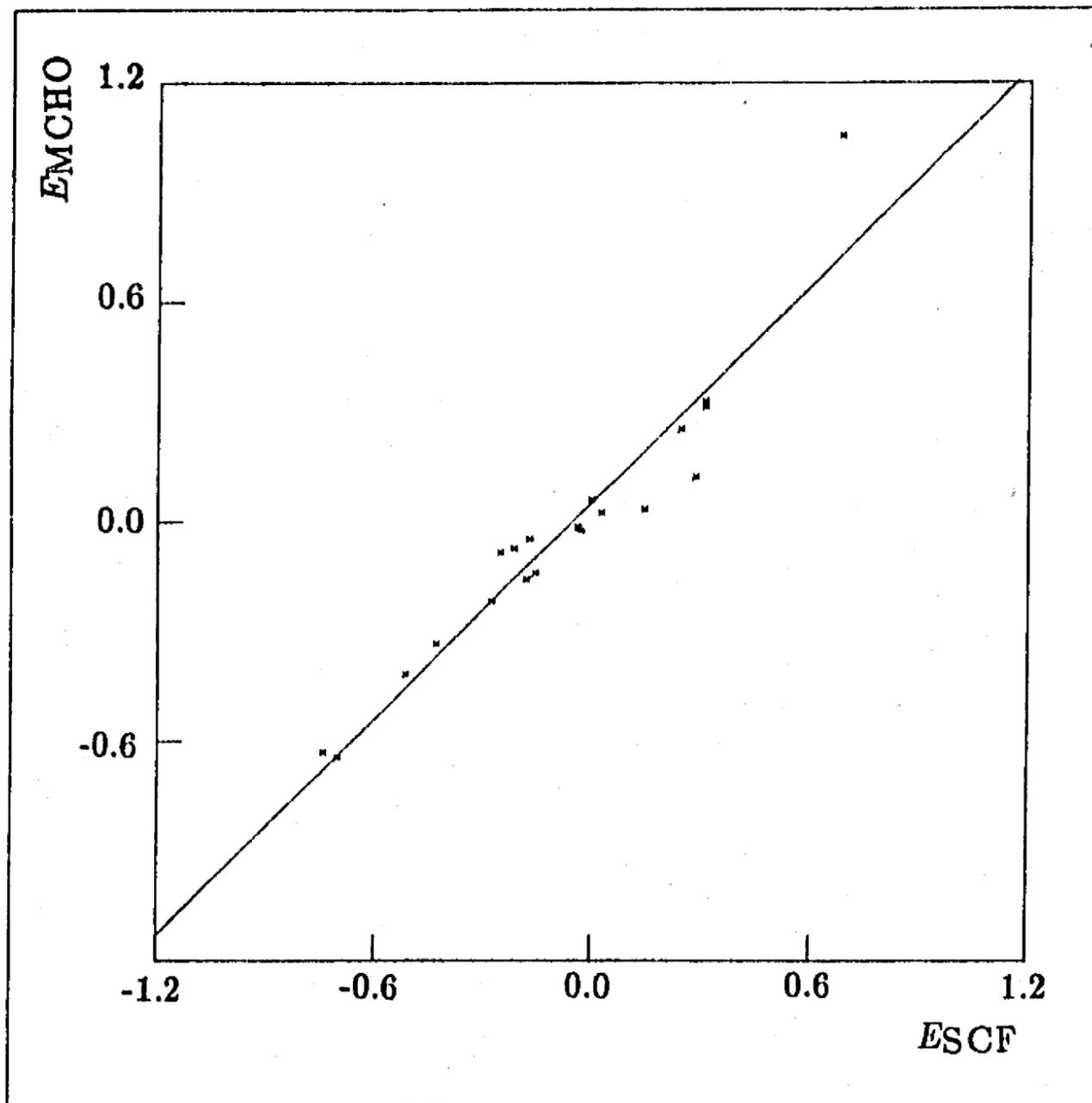


Figura 5-6. Comparación de los valores de las contribuciones no aditivas a las energías de interacción de trímeros  $H_4P_2O_7 - 2H_2O$  calculados *ab initio* (abscisas) y por el modelo MCHO (ordenadas). Las unidades son kcal/mol.

Para la simulación se usó la misma celda computacional que en los casos anteriores, con el mismo procedimiento para eliminar moléculas de agua, usando los valores para distancias mínimas mostrados en la Tabla 5-6. En este caso se eliminaron 4 moléculas de agua y la simulación se hizo con las 339 restantes y la de pirofosfato. Otra vez se obtuvo el promedio de energías después de  $\sim 4 \times 10^6$  configuraciones y se llegó a una energía de hidratación de  $\sim -100 \text{ kcal/mol}$ , con una dispersión muy grande.

TABLA 5-6

Atomo	$O_w$	$H_w$
$O_w$	4.5	
$H_w$	2.0	2.0
$P_p$	3.5	3.5
$O_p$	2.0	3.5
$H_p$	2.0	2.5

Tabla 5-6. Distancias mínimas de acercamiento entre los átomos del pirofosfato y los del agua, en u. a. Los subíndices  $w$  y  $p$  señalan a cuál molécula pertenece cada átomo.

En vista de que los resultados de las simulaciones con el ortofosfato y con el pirofosfato presentaron problemas, se decidió revisar los que se habían obtenido con el agua pura, con el fin de buscar algún posible error. De hecho, ésta es una de las ventajas del método que se aplicó en el trabajo presente: el conocimiento preciso de las aproximaciones y de los procedimientos usados, permite realizar la búsqueda de las causas de los posibles errores. Ya que se usaron las mismas condiciones en los cálculos y las simulaciones de todos los sistemas, es relativamente fácil detectar el problema y corregirlo, como se muestra en las siguientes secciones.

### 5.5 La descripción del agua líquida con el potencial MCHO

En este trabajo se pretende llegar a estimaciones confiables de las energías de hidratación del ortofosfato y del pirofosfato, a partir de simulaciones MC de soluciones acuosas, cuyos potenciales de interacción intermolecular fueron ajustados para reproducir superficies de energía calculadas *ab initio*. La confiabilidad de los valores propuestos depende de varios factores, algunos de los cuales ya fueron mencionados.

#### 5.5.1 La convergencia de los resultados de las simulaciones

En el caso de las simulaciones, un criterio que ha sido usado comúnmente (*e. g.* [18]) para juzgar la convergencia, consiste en revisar las variaciones con respecto a la energía promedio, obtenida luego de algunos millones de configuraciones. Si estas variaciones son del orden del 0.1%, se considera haber llegado al equilibrio. Un criterio adicional lo proporcionan las funciones de distribución radial,  $g(r)$ , al compararlas con datos experimentales de difracción de rayos X o de dispersión de neutrones, cuando éstos existen.

Hasta fechas recientes, en los estudios sobre líquidos moleculares, se han empleado simulaciones con un número de configuraciones menor a  $10^7$ , en buena medida debido al costo computacional que representa generarlas. Sin embargo, ya existen varios trabajos [28 - 30] que indican la necesidad de simulaciones más largas.

Con el potencial MCHO se obtuvieron valores de la energía y de las funciones de estructura del agua, que dan la mejor reproducción reportada de los datos experimentales [21] (Apéndice C). Estos resultados corresponden a una simulación con apenas  $1.5 \times 10^6$  configuraciones.

Por otra parte, en el laboratorio de Biofísica, se intentó agregar al potencial MCHO la posibilidad de incluir en la evaluación de la energía, los efectos de las rotaciones y de las vibraciones de la molécula de agua. Esto hizo necesario calcular las contribuciones no aditivas de un mayor número de trimeros y de tetrameros de moléculas de agua, y modificar los parámetros para ajustarse a la nueva superficie y tratar de hacerlos compatibles con la parte vibracional. La reparametrización mantuvo una misma calidad para el ajuste a las interacciones por pares, pero mejoró en el ajuste a las no aditividadades [27]. Los valores que se presentan en la Tabla 5-2 corresponden a esta reparametrización, y los resultados de una simulación de  $\sim 10^6$  configuraciones, fueron similares a los reportados en la ref. [21]. En la fig. 5-7 se muestran los ajustes a (1) las interacciones por pares, (2) los efectos no aditivos de tres cuerpos, y (3) los efectos no aditivos de cuatro cuerpos.

Ante los problemas en los resultados sobre las energías de hidratación de los fosfatos, se decidió hacer una simulación con varios millones de configuraciones, para verificar la convergencia de los resultados para el agua líquida. Entonces se encontró que las configuraciones utilizadas en la ref. [21] no estaban bien convergidas.

Para los fines de esta tesis, los datos más relevantes son la energía y las funciones de correlación radial, producidas por el potencial MCHO, con la parametrización reportada arriba; así que las dos siguientes secciones se dedican a la discusión de estas cantidades.

FIGURA 5-7

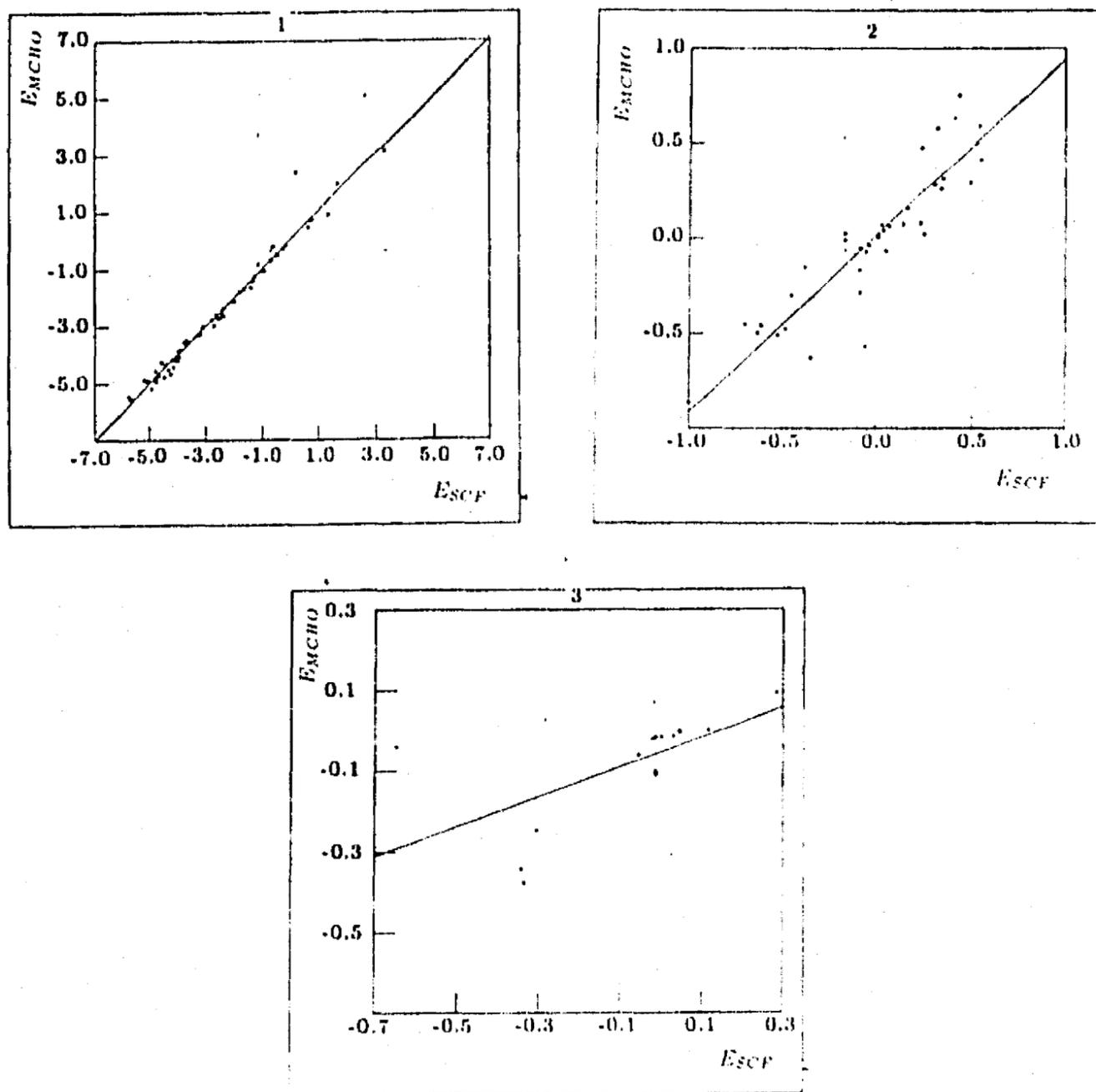


Figura 5-7. En estas gráficas se muestra el ajuste del potencial MCHO a los cálculos *ab initio* de (1) las interacciones por pares, (2) los efectos no aditivos de tres cuerpos, y (3) los de cuatro cuerpos. Se usaron los parámetros de la Tabla 5-2. Las unidades son kcal/mol, en los tres casos.

### 5.5.2 La energía intermolecular del agua líquida

El valor *experimental* de la energía intermolecular del agua, al que se refieren los trabajos sobre el tema, es de  $E_i = -9.92 \text{ kcal/mol}$ , y se basa en medidas reportadas en un libro de 1940 [31], que es difícil de conseguir. La determinación de este valor no proviene de una medida directa, sino que se deduce a partir del calor de vaporización,  $\Delta H_{vap} = -E_i + P\{V(g) - V(l)\}$ . En esta relación,  $P$  es la presión, y  $V(g)$  y  $V(l)$  son los volúmenes del gas y del líquido, respectivamente. El término  $P\Delta V$  se considera igual a  $RT$  (ecuación del gas ideal, con  $R$  la constante de los gases, y  $T$ , la temperatura). Entonces, el valor que se cita para la energía  $E_i$  del agua a  $25^\circ\text{C}$ , es de  $E_i = -10.51 + 0.59 = -9.92 \text{ kcal/mol}$ . La consideración hecha requiere de dos suposiciones: (1) que el vapor de agua puede considerarse como un gas ideal, y (2) que la suma de las energías de traslación y de vibración en el vapor es la misma que en el líquido [11]. En tanto que la primera suposición ha sido justificada [11], la segunda no ha quedado bien establecida. De hecho, en un análisis sobre las correcciones cuánticas a las propiedades termodinámicas del agua líquida, que incluye la consideración de osciladores cuánticos en los términos vibracionales, Berens *et al.* [32] obtienen resultados que son consistentes con una energía intermolecular *experimental* de entre  $-10 \text{ kcal/mol}$  y  $-10.5 \text{ kcal/mol}$ . Además, los valores de la energía que resultan de simulaciones de agua líquida con diversos potenciales, también se encuentran en este intervalo [12, 16, 17].

Por otra parte, la energía calculada con el potencial por pares MCY [33, 34] es de  $E = -8.6 \text{ kcal/mol}$ ; en cambio, cuando se usa una extensión que incluye los efectos no aditivos de tres y cuatro cuerpos [20], el resultado es de  $E = -10.65 \text{ kcal/mol}$ .

En el caso del potencial MCHO, el valor reportado [21] es de  $E = -9.9 \text{ kcal/mol}$ , en concordancia con la energía *experimental*, sin corrección cuántica. Esta cifra se obtuvo de una simulación con  $1.5 \times 10^6$  configuraciones. Sin embargo, después de  $10^7$  configuraciones, se alcanzó un valor de  $E = -11.14 \text{ kcal/mol}$ , que se mantuvo estable por  $5 \times 10^7$  configuraciones más. Aunque está un poco sobrestimada, esta cifra va más de acuerdo con los resultados discutidos en los párrafos anteriores. De hecho, es razonable que el potencial MCHO no reproduzca la energía *experimental*, ya que está ajustado a una superficie de interacciones por pares que tiene únicamente 66 puntos, y pueden ser una muestra muy restringida e insuficiente. Además, en las simulaciones se han usado condiciones de frontera toroidales [35], con una distancia de corte de  $20.5 \text{ u.a.}$  para las interacciones intermoleculares de corto alcance, y la consideración de imagen mínima [36], para las de largo alcance. Este tipo de simulaciones ha sido criticado por diversos autores [36 - 40], en el sentido de que no produce energías correctas. En el laboratorio de Biofísica del IFUNAM, ya se están haciendo estudios para superar los problemas mencionados.

Con el fin de obtener resultados consistentes, en todas las simulaciones de esta tesis se utilizaron las mismas aproximaciones (i. e., moléculas rígidas y condiciones de frontera toroidales). En todo caso, queda claro que para poder discriminar si el potencial que se use es adecuado, se requiere de simulaciones largas (de más de  $10^7$  configuraciones) para llegar a equilibrar los sistemas bajo estudio.

### 5.5.3 La estructura del agua líquida a temperatura ambiente ( $T = 298\text{ K}$ )

Como se mencionó en las secciones 3.5 y 5.5.1, de las simulaciones numéricas, se pueden obtener los datos estructurales de un líquido. La manera más común de presentarlos es por medio de las funciones de distribución (o de correlación) radial,  $g(r)$ . Estas funciones consisten en la evaluación de la densidad de un tipo de átomo  $B$ , en cascarones esféricos centrados en un átomo  $A$ , y en la comparación de la densidad de cada cascarón con la densidad de los átomos del tipo  $B$  en todo el sistema (fig. 5-8). Así se consigue una representación gráfica de la distribución de los átomos del tipo  $B$  alrededor de los del tipo  $A$ .

Así como se hace con la energía, las predicciones sobre la estructura que resultan de las simulaciones, se comparan con datos experimentales. La manera más directa para medir las propiedades estructurales, es por medio de los efectos de difracción en la radiación dispersada en forma coherente, por una muestra del sistema bajo estudio. Para lograr una resolución en escala atómica, se requiere usar una longitud de onda de  $1\text{ \AA}$ , que puede obtenerse con el uso de rayos X, de neutrones lentos, o de electrones de baja energía.

Los patrones de difracción producidos por las estructuras periódicas de las redes cristalinas, constan de máximos de Bragg bien definidos, debidos a las interferencias constructivas ocasionadas por las reflexiones en los distintos planos de la red. En cambio, para los sistemas desordenados el patrón de difracción, tiene una forma más oscilatoria; pero aún contiene la información estructural relevante. Este patrón de difracción es la función de forma  $H(s)$ , en que la variable independiente es la magnitud del vector de dispersión  $s$ , y se define como  $s = \frac{4\pi}{\lambda} \sin\theta$ , donde  $\lambda$  es la longitud de onda de la radiación, y  $2\theta$  es el ángulo de dispersión [41, 42]. La función de forma  $H(s)$  contiene la información estructural de todos los átomos en el sistema. La conversión de esta función  $H(s)$  en una función de correlación radial compuesta,  $\bar{g}(r)$ , requiere del uso de técnicas de transformaciones de Fourier.

La molécula de agua consta de dos tipos de átomos,  $O$  y  $H$ , así que existen tres funciones de correlación radial por pares,  $g_{OO}(r)$ ,  $g_{OH}(r)$  y  $g_{HH}(r)$ , que definen las relaciones espaciales de unos átomos con respecto a los otros. La función compuesta  $\bar{g}(r)$  que se obtenga de cualquier experimento de difracción, será una combinación ponderada de esas tres funciones parciales. Si se efectúan tres mediciones independientes, ya sea con distintos tipos de radiación, o usando la dispersión de neutrones de muestras con distintas proporciones isotópicas  $H/D$ , en principio se puede determinar separadamente las tres funciones. En la práctica, la interpretación de los resultados de estos experimentos es muy difícil, porque se necesita que las amplitudes de la dispersión atómica tengan valores suficientemente distintos como para discriminarlos. De manera que los datos reportados por diferentes grupos han estado sujetos a una gran controversia: hasta fechas relativamente recientes, la función  $g_{OO}(r)$  experimental contra la cual se comparaban las que se obtenían de las simulaciones, era la reportada por Narten y Levy [42], que se basa en un experimento de difracción de rayos X. De hecho, en este artículo no se presenta una función  $g_{OO}(r)$ , sino una función de correlación entre los centros dispersores de rayos X, de las moléculas de agua,  $g_{MM}(r)$ . En la obtención de esta función se supone que las moléculas de agua

tienen una densidad electrónica lo suficientemente esférica, que permite omitir las desviaciones de la esfericidad. También se considera que los centros dispersores están situados prácticamente en los átomos de oxígeno.

FIGURA 5-8

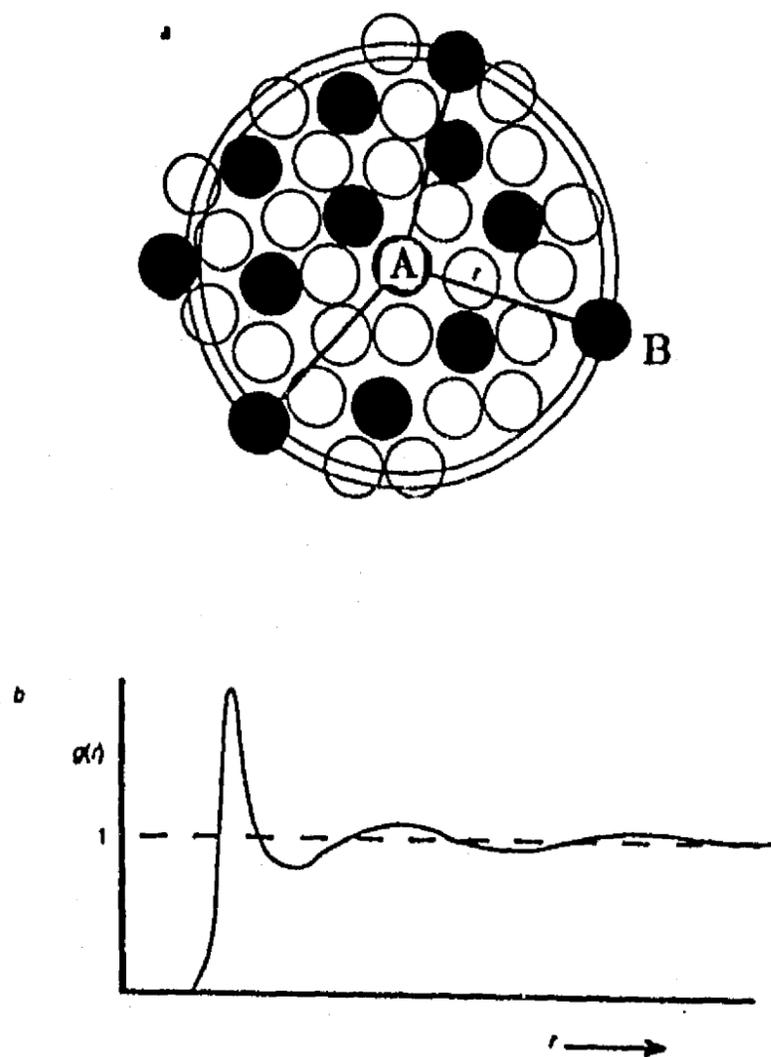


Figura 5-8. Representación de la función de correlación radial: (a) éste es un diagrama esquemático de una configuración particular de un sistema con átomos de dos tipos, *A* y *B*. Para calcular la  $g(r)$  mostrada en la gráfica (b), se cuenta el número de átomos del tipo *B* que hay en cada cascarón esférico centrado en un átomo del tipo *A*,  $N_B(r_0)$ , y se divide entre el volumen del cascarón. La densidad así obtenida se compara con la densidad de átomos del tipo *B* en todo el sistema,  $\rho_B$ , y ése es el valor de  $g_{AB}(r_0) = \frac{N_B(r_0)}{4\pi\rho_B r_0^2 dr}$ .

Posteriormente al experimento de Narten y Levy [42], se hicieron algunas medidas de la dispersión de neutrones por sistemas con agua pesada [43 - 46], con la finalidad de determinar la función  $g_{HH}(r)$ . Los datos reportados en cada trabajo fueron tan diferentes (cfr. [46]), que ameritaban una revisión sobre el tipo de análisis hecho a las funciones de estructura, que son las mismas en todos los trabajos. Esta revisión la hizo Soper [46], y concluyó que se había realizado una interpretación incorrecta de los efectos dinámicos sobre la sección transversal diferencial de los átomos de hidrógeno. En los trabajos anteriores, dichos efectos fueron representados por medio de polinomios ajustados a los datos de la dispersión, lo cual puede causar una mala representación de la forma de la autodispersión (*self-scattering*) en el intervalo  $1 \text{ \AA}^{-1} \leq s \leq 5 \text{ \AA}^{-1}$ . En cambio, Soper [46] utilizó un método empírico que no necesita de ajustes arbitrarios. Con este método, aplicado a datos de difracción de neutrones, Soper y Phillips [47] lograron determinar un nuevo conjunto de funciones  $g_{OO}(r)$ ,  $g_{OH}(r)$  y  $g_{HH}(r)$ .

Poco después, Lie [48] diseñó una verificación para la consistencia de las funciones de correlación, que se basa en que la información para obtener la constante de compresibilidad isotérmica,  $\kappa$ , está contenida en esas funciones [49]:

$$\rho k T \kappa = 1 + \int_0^{\infty} 4\pi r^2 \rho [g(r) - 1] dr \quad (5-1).$$

En esta ecuación  $\rho$  es la densidad de moléculas,  $k$  la constante de Boltzmann, y  $T$  la temperatura. Si se considera una distancia  $R_{min}$ , tal que para distancias mayores, el líquido puede considerarse como un continuo de densidad uniforme (i. e. donde  $g(r) = 1$ ), la ec. (5-1) puede escribirse como:

$$\int_0^{R_{min}} 4\pi r^2 \rho g(r) dr = \frac{4}{3} \pi R_{min}^3 \rho - 1 + \rho k T \kappa \quad (5-2).$$

Esta última ecuación es invariante ante la transformación  $R_{min} \rightarrow R_{min} + r$ , con  $r > 0$ . Lie considera como consistente a un grupo de funciones de correlación, si cumplen la igualdad (5-2), cuando en ella se usa el valor experimental [50]  $\kappa = 41 \times 10^{-6} \text{ atm}^{-1}$ . En su artículo, Lie [48] compara las desviaciones que con respecto a esta igualdad, resultan de las funciones de Narten [42 - 44] y de las funciones de Soper [47]. La conclusión es que estas últimas son las más consistentes.

Es interesante notar en el trabajo de Soper [46] la referencia a los resultados de las simulaciones: una motivación para revisar las funciones de correlación radial *experimentales* fue que ningún modelo reproducía todas las características de la  $g_{MM}(r)$  de Narten [42]; sino que todos presentaban un primer máximo más alto, y una transición muy lisa hacia el segundo; a diferencia de la  $g_{MM}(r)$ , que tiene un punto de inflexión en  $r \approx 4 \text{ \AA}$  (ver fig. 5-9). En la simulación de  $1.5 \times 10^6$  configuraciones, el potencial MCHO sí produjo una  $g_{OO}(r)$  similar a la  $g_{MM}(r)$  (fig. 5-9).

FIGURA 5-9

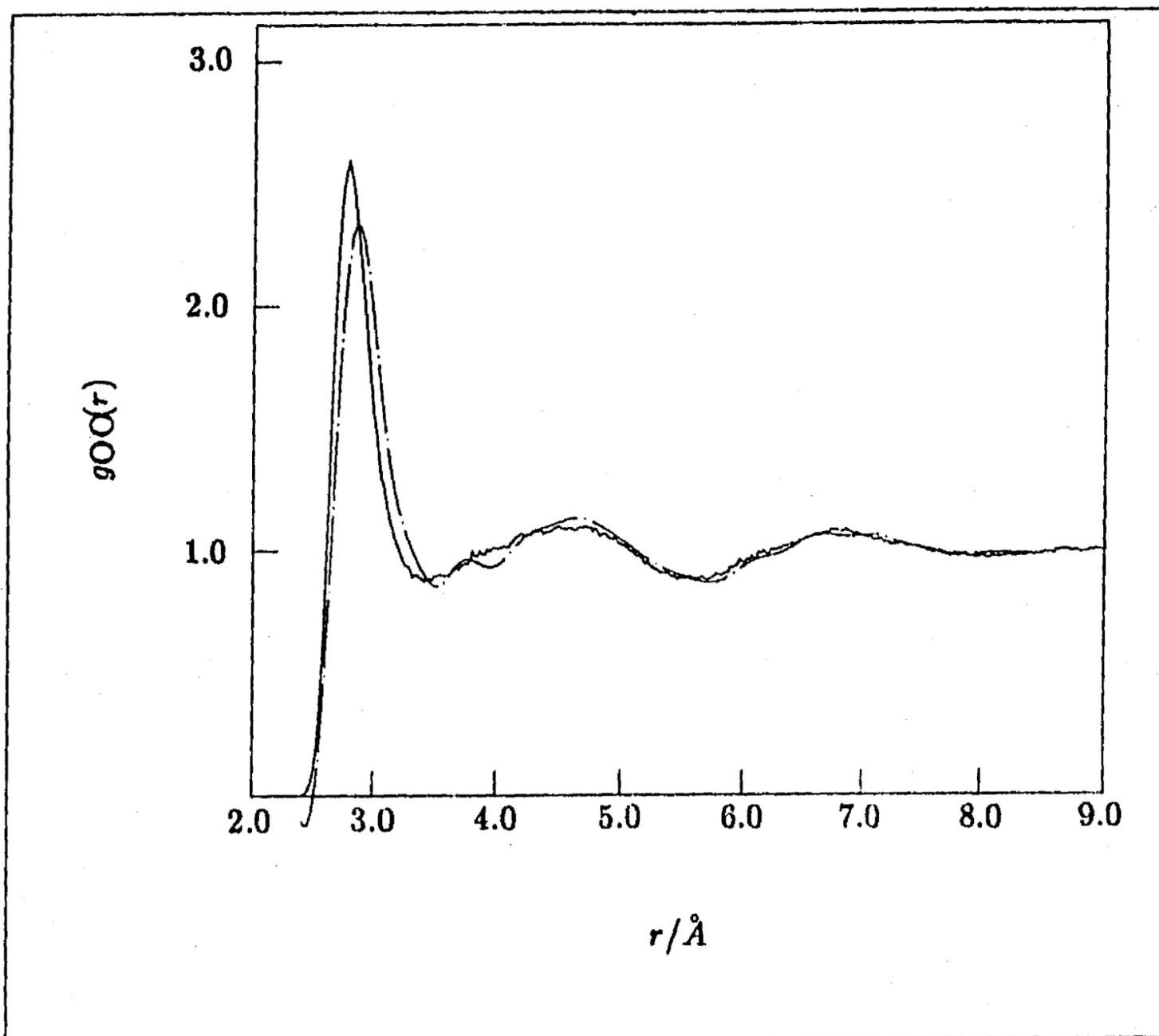


Figura 5-9. Función de correlación radial  $g_{OO}(r)$  reportada para el modelo MCHO en la ref. [21] (línea continua), comparada con la  $g_{MM}(r)$  de Narten y Levy [42] (línea punteada).

Un cambio dramático que se obtiene en la simulación más larga, está en las funciones de correlación radial: se encuentra la mejor reproducción que se ha reportado de las funciones de Soper y Phillips [47] (fig. 5-10).

FIGURA 5-10

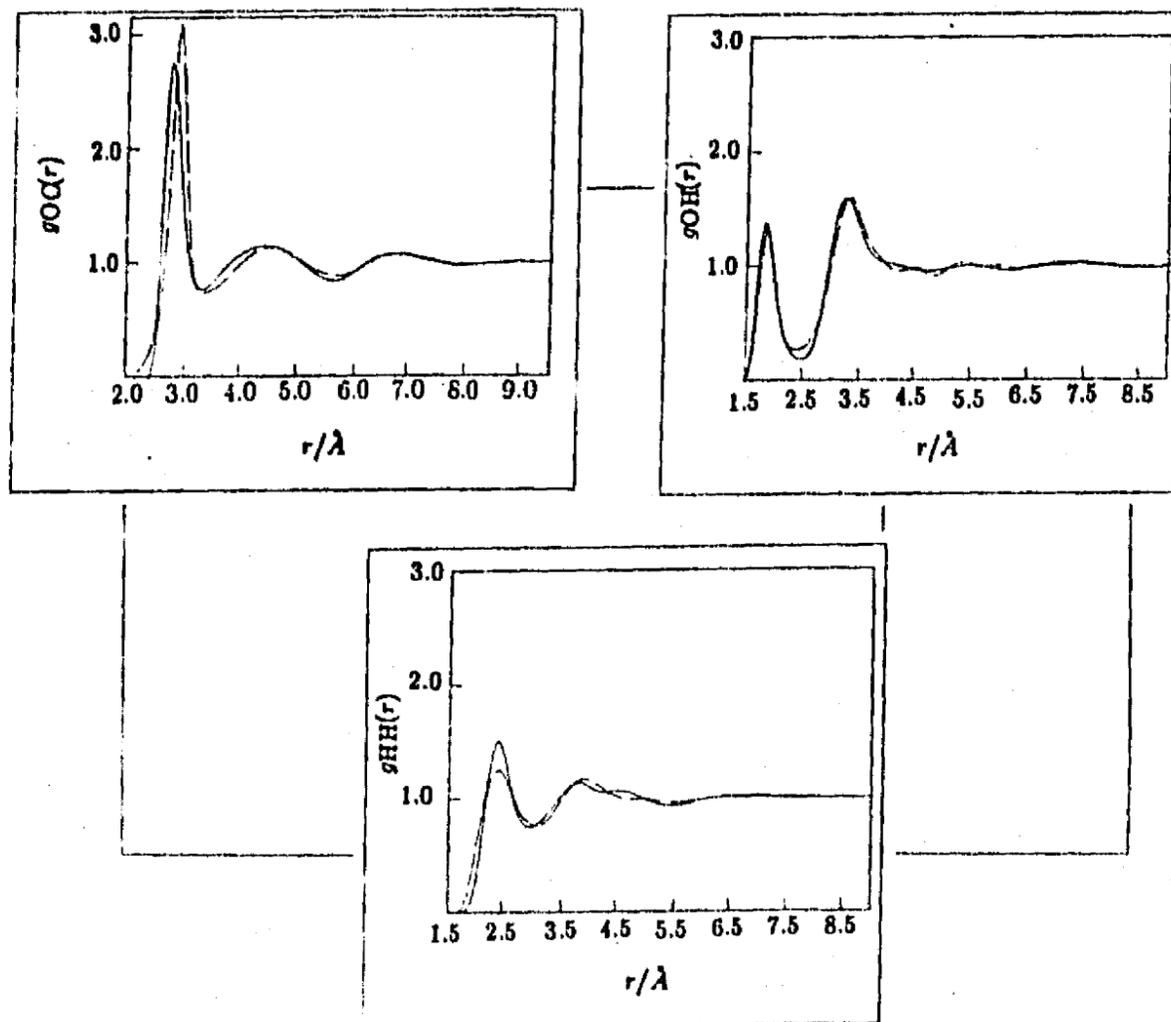


Figura 5-10. Comparación de las funciones de correlación radial producidas en una simulación de  $5 \times 10^7$  configuraciones con el modelo MCHO (línea continua), con respecto a las curvas experimentales de Soper y Phillips [47] (línea punteada). (a)  $g_{OO}(r)$ , (b)  $g_{OH}(r)$  y (c)  $g_{HH}(r)$ .

Por supuesto, las funciones de estructura de dispersión de rayos X,  $H_X(s)$ , y de neutrones,  $H_N(s)$ , son muy parecidas a las experimentales (fig. 5-11). De hecho, éste es el primer caso reportado en el que se reproducen los dos máximos de  $H_X(s)$  en  $s = 2.08 \text{ \AA}^{-1}$  y en  $s = 2.88 \text{ \AA}^{-1}$ , con las alturas correctas.

FIGURA 5-11

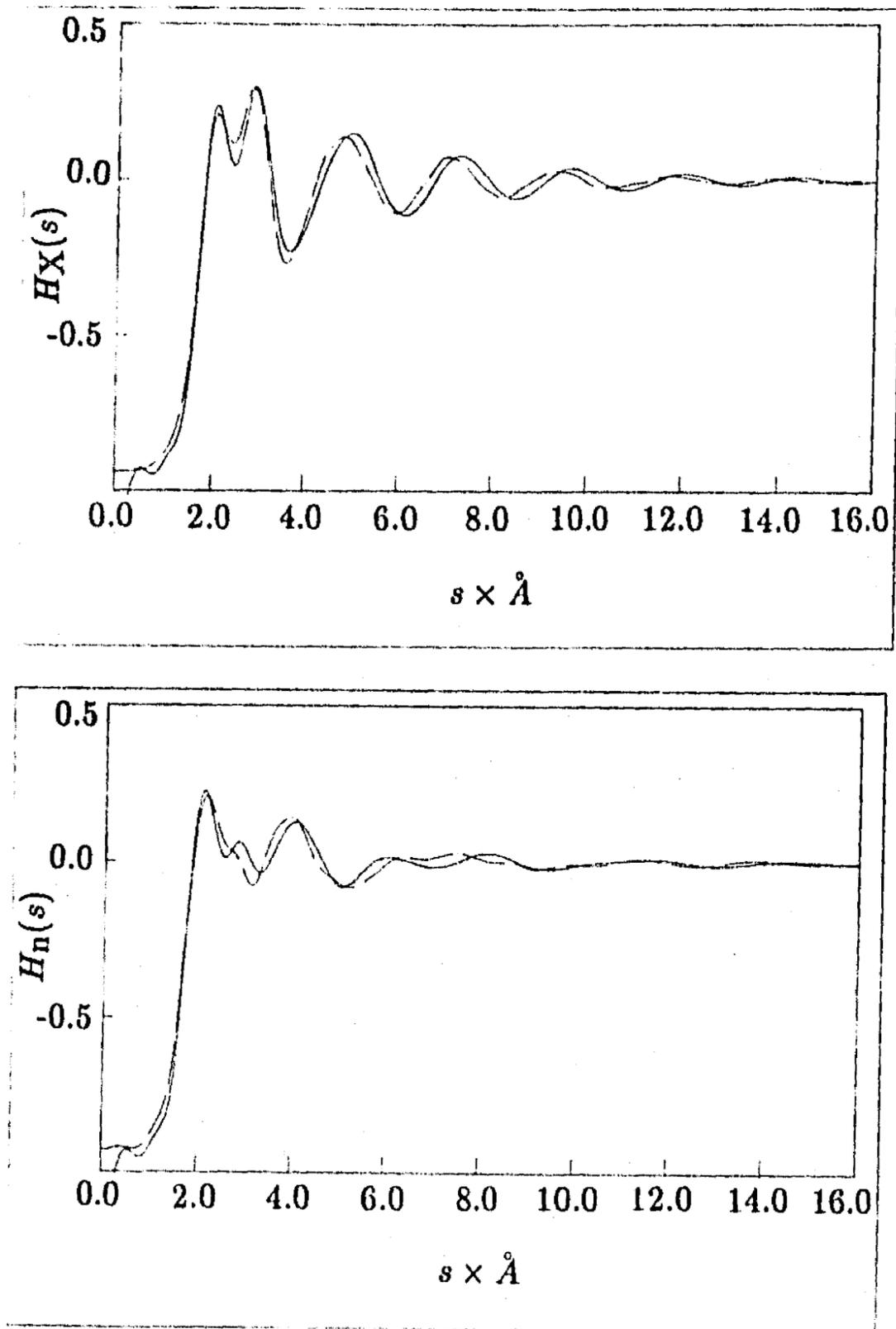


Figura 5-11. Comparación de las funciones de estructura producidas en una simulación de  $5 \times 10^7$  configuraciones con el modelo MCHO (línea continua), con respecto a las curvas experimentales [42 - 47] (línea punteada). (a)  $H_X(s)$ , correspondiente a la difracción de rayos X, y (b)  $H_N(s)$ , a la dispersión de neutrones.

Los datos aquí presentados demuestran que el potencial MCHO es el que mejor representa la estructura del agua líquida a  $T = 298 K$ . El corrimiento del primer máximo en  $g_{OO}(r)$  hacia distancias más cortas puede provenir de una exageración de los efectos no aditivos de tres y cuatro cuerpos, o a la falta de flexibilidad [21]. Esta característica es común a varios modelos (cfr. [12] y [51]), algunos de los cuales no son polarizables. El mismo corrimiento puede observarse en modelos con moléculas flexibles [52], así que la relajación de grados de libertad no hace la corrección deseada. Actualmente en el laboratorio de Biofísica del IFUNAM, se están desarrollando estudios sobre los efectos no aditivos, tendientes a explicar el efecto mencionado [27] y a refinar el modelo para corregirlo. También se han comenzado las simulaciones para verificar la capacidad que tiene el modelo MCHO, de reproducir la estructura del agua, a diferentes temperaturas, y en diferentes fases.

Ya que se ha demostrado que el potencial MCHO es confiable para la descripción del agua líquida, y que la obtención de datos estadísticamente significativos requiere de simulaciones muy largas, es posible proponer la estimación de las energías de hidratación de los fosfatos, a partir de simulaciones numéricas, usando el potencial MCHO, con los parámetros de las Tablas 5-2 y 5-5.

### 5.6 Las energías de hidratación del ortofosfato y del pirofosfato

Ya que se verificó la convergencia de los resultados para el agua pura con el modelo MCHO, se procedió a efectuar las estimaciones de las energías de hidratación del  $H_3PO_4$  y del  $H_4P_2O_7$ . A partir de una configuración ya equilibrada de 343 moléculas de agua se realizó el proceso de eliminación descrito en las secciones 5.3 y 5.4, para hacer simulaciones de una molécula de ortofosfato con 340 de agua, y de una de pirofosfato con 339 de agua. En ambos casos, se analizó la energía del sistema durante  $2 \times 10^7$  configuraciones, al cabo de las cuales se mantenía estable para el  $H_3PO_4$ , en tanto que la del  $H_4P_2O_7$  seguía teniendo variaciones importantes.

Para el procesamiento estadístico de las energías de los sistemas, se usaron otras  $5 \times 10^7$  configuraciones, registrando la energía cada  $10^5$  pasos. Cada corrida total se dividió en cinco grupos de  $10^7$  configuraciones. Para cada grupo se obtuvo la energía promedio de cada sistema (Tabla 5-7) y con estos valores se hicieron los promedios de las corridas, reportando como incertidumbre la desviación estándar en cada caso. En el último renglón de la Tabla 5-7 se presentan estos resultados.

Las energías de hidratación que se presentan en la Tabla 5-7 se calcularon con la ec. (3-47). La energía del agua ya fue discutida en la sección 5.4. La energía de hidratación de  $H_3PO_4$ ,  $\Delta E_{hid} = -29.58 kcal/mol$ , es muy similar a la que reportan George *et al.* [1], de  $\Delta E_{hid} = -26 kcal/mol$ , y los datos de la Tabla 5-7 muestran que se ha logrado una convergencia razonable. En el caso del  $H_4P_2O_7$ , se ha llegado a una energía de hidratación de  $\Delta E_{hid} = -40.76 kcal/mol$ , que también está en acuerdo con la cifra de George *et al.* [1], de  $\Delta E_{hid} = -38 kcal/mol$  (cfr. sección 5.1.1).

TABLA 5-7

Sistema	$H_2O$	$H_3PO_4$	$H_4P_2O_7$
1	-3818.09	-3816.62	-3798.16
2	-3819.46	-3818.04	-3801.31
3	-3822.49	-3819.40	-3803.51
4	-3823.13	-3817.10	-3818.56
5	-3819.06	-3811.90	-3814.65
Energía	-3820.45 ±1.99	-3816.61 ±2.54	-3816.65* ±2.11
$\Delta E_{hid}$	-11.14	-29.58	-40.76

Tabla 5-7. Promedios de las energías obtenidas para cada sistema simulado, tomados cada  $10^7$  configuraciones. En el último renglón se muestran los valores asignados a cada sistema, obtenidos con los promedios de los cinco renglones anteriores. Se reporta como incertidumbre la desviación estándar de la muestra de cinco valores. \*Puede notarse cómo la energía del sistema con pirofosfato aún muestra una tendencia a disminuir, por lo que el promedio se hizo únicamente con los resultados de las últimas  $2 \times 10^7$  configuraciones (renglones 4 y 5). Las unidades son kcal/mol.

La estimación de la contribución de la hidratación a la entalpía de la reacción (4-1),



se calcula como la diferencia

$$\Delta\Delta E_{hid} = 2 \times \Delta E_{hid}(H_3PO_4) - \Delta E_{hid}(H_4P_2O_7) - \Delta E_{hid}(H_2O) \quad (5-2).$$

El resultado es de:

$$\Delta\Delta E_{hid} = -7.26 \text{ kcal/mol} \quad (5-3).$$

Si se hace una propagación de errores con las desviaciones estándar de la Tabla 5-7, la incertidumbre asociada al valor en (5-3) es de  $\pm 9 \text{ kcal/mol}$ , que es mayor que la energía misma. Esto es común a las energías de hidratación de especies eléctricamente neutras (cfr. [30]). Sin embargo, hay que hacer notar que esta incertidumbre no corresponde a un error en la medida, en el sentido en que se aplica a un experimento. La desviación estándar de los resultados de una simulación proporciona un criterio para juzgar qué tan cerca del equilibrio está un sistema, y representa el tamaño de las fluctuaciones de las cantidades calculadas, con respecto a un valor promedio. De hecho, algunos parámetros termodinámicos se calculan precisamente a partir de esas fluctuaciones, como es el caso del calor específico [52].

De cualquier manera, en la Tabla 5-7 puede observarse que el sistema del  $H_4P_2O_7$  todavía no está completamente equilibrado. A pesar de ello, el valor de la energía ya comenzó

a presentar un comportamiento que no es monótonamente decreciente: la energía de las últimas  $10^7$  configuraciones es mayor que la de las penúltimas, y el promedio presentado es una estimación razonable del valor final.

Entonces la estimación de la entalpía de la reacción (4-1) en solución acuosa, es la suma de la entalpía en la fase gaseosa, más la contribución de la hidratación:

$$\begin{aligned}\Delta H_{ag} &= \Delta H_g + \Delta\Delta E_{hid} \\ &= -3.8 - 7.3 = -11.1 \text{ kcal/mol}\end{aligned}\quad (5-4).$$

Este es un valor exagerado, con respecto al dato experimental de George *et al.* [1],  $\Delta H_{ag} = -7.6 \text{ kcal/mol}$ , y amerita una discusión más amplia:

La estimación de la entalpía para la fase gaseosa es muy confiable, como se mostró en el capítulo 4. En ese caso, era de esperar que la contribución de la hidratación fuera de  $\Delta\Delta E_{hid} = -3.8 \text{ kcal/mol}$ , así que la exageración en la cifra de  $\Delta H_{ag} = -11.1 \text{ kcal/mol}$  debe provenir de las simulaciones. Éstas se efectuaron con tres sistemas: agua pura ( $H_2O$ ), agua con ortofosfato ( $H_3PO_4$ ), y agua con pirofosfato ( $H_4P_2O_7$ ). En la sección 5.5 ya se probó que los resultados del sistema  $H_2O$  ya han convergido, y en la Tabla 5-7 se muestra que el sistema  $H_3PO_4$  ya también llegó al equilibrio. Entonces hay que revisar al sistema  $H_4P_2O_7$ . Aunque este último no ha llegado al equilibrio, la energía de hidratación propuesta,  $\Delta E_{hid}(H_4P_2O_7) = -40.8 \text{ kcal/mol}$ , es un valor que ya presenta cierta convergencia.

En resumen, la cifra de  $\Delta\Delta E_{hid} = -7.3 \text{ kcal/mol}$  es efectivamente la estimación que se obtiene con el modelo MCHO, bajo las condiciones de la simulación. Por lo tanto el error debe originarse en las aproximaciones del modelo. Estas residen principalmente en: (1) el tipo de base usado en la generación de las superficies de energía y en la omisión de la correlación electrónica, cuando se hicieron los cálculos *ab initio*; (2) la posible incapacidad del modelo de reproducir correctamente el espectro de las energías calculadas *ab initio*, y (3) el uso de moléculas rígidas (aunque el modelo es polarizable, las posiciones relativas de los núcleos de los átomos en una misma molécula, son invariantes).

Las superficies de energía se calcularon con una base extensa (6-31G\*\*, ver la sección 4.2) y con el mismo nivel de aproximación para los sistemas  $H_3PO_4$  y  $H_4P_2O_7$ , así que es razonable suponer que los errores producidos en esta etapa sean muy similares para ambos sistemas, y que al calcular una cantidad como la diferencia de energías entre un sistema y el otro, los errores tiendan a cancelarse. Este no es necesariamente el caso, y puede ocurrir que el efecto de estos errores sea el causante de la discrepancia con el dato experimental. Ésta es una posibilidad que habría que explorar; pero por el costo computacional que representa, vale la pena examinar primero otras.

La segunda posibilidad: que el modelo no reproduzca el espectro de energías, parece muy poco probable, ya que los ajustes logrados son muy buenos (cfr. figs. 5-2, 5-3, 5-5 y 5-6). Además, la comparación de los resultados experimentales con los de las simulaciones del agua pura, que usan el modelo MCHO, es mejor que la de los demás modelos existentes (cfr. sección 5.5).

Queda entonces la aproximación de moléculas rígidas. En el caso del agua, funciona muy bien, porque es efectivamente una molécula muy poco flexible [53]. En las secciones 4.3 y 4.5 se demostró que el ortofosfato tiene esta misma característica. En cambio, en las mismas secciones se encontró que el pirofosfato sí es una molécula más flexible, con capacidad de giro alrededor del puente  $P - O - P$ , y de movimientos circulares de los átomos de hidrógeno. Para esta especie la aproximación de molécula rígida es inadecuada. Más aún, al relajar algunos grados de libertad, la energía de la simulación con el sistema  $H_4P_2O_7$  seguramente disminuiría, corrigiendo la energía  $\Delta\Delta E_{hid}$  en el sentido adecuado, i. e., hacia valores menos negativos (cfr. ec. (5-2)). Esta es la posibilidad de refinamiento más promisorio, no sólo por el costo computacional, sino porque los datos indican que el origen más probable de la exageración en la  $\Delta\Delta E_{hid}$ , está en la aproximación de moléculas rígidas.

Aun cuando las configuraciones del pirofosfato no están completamente convergidas, se utilizó el método descrito en la sección 3.5 para calcular las energías libres de Helmholtz de los sistemas simulados. Los resultados se presentan en la Tabla 5-8.

TABLA 5-8

Sistema	$H_2O$	$H_3PO_4$	$H_4P_2O_7^*$
1	-2247.00	-2261.81	
2	-2264.04	-2255.01	
3	-2286.69	-2247.48	
4	-2227.85	-2247.81	-2257.15
5	-2229.89	-2261.40	-2257.33
Energía	-2250.90	-2254.70	-2257.24*
$\Delta F_{hid}$	-6.56	-23.44	-32.51

Tabla 5-8. Promedios de las energías libres de Helmholtz, obtenidas para cada sistema simulado, tomados cada  $10^7$  configuraciones. En el último renglón se muestran los valores asignados a cada sistema, obtenidos con los promedios de los cinco renglones anteriores. \*Como el sistema con pirofosfato aún no está bien convergido, el promedio se hizo únicamente con los resultados de las últimas  $2 \times 10^7$  configuraciones (renglones 4 y 5). Las unidades son kcal/mol.

Así como ocurrió con la energía intermolecular, el valor que se obtuvo para la energía libre del agua,  $\Delta F = -6.56 \pm 0.06$  kcal/mol, es menor en casi 1 kcal/mol al dato experimental de  $\Delta F = -5.73$  kcal/mol. La diferencia en las energías libres de Helmholtz, que corresponde a la hidratación de los fosfatos es de:

$$\Delta\Delta F_{hid} = 2(-23.44) - (-32.51) - (-6.56) = -7.81 \text{ kcal/mol} \quad (5-5)$$

Este último valor está sujeto a las condiciones previamente discutidas, por lo que no es conclusivo. Sin embargo, es de una magnitud similar a la de la estimación de la energía

de hidratación,  $\Delta E_{hid} = -7.26 \text{ kcal/mol}$ . La diferencia entre las dos sería una estimación de la contribución del término entrópico:

$$T\Delta S = \Delta F - \Delta E_{hid} = -0.55 \text{ kcal/mol} \quad (5 - 6).$$

El resultado es apenas una tercera parte del esperado,  $T\Delta S = -1.9 \text{ kcal/mol}$ , y seguramente requiere un refinamiento. De hecho, como la magnitud de la entropía depende de dos cantidades que están sujetas a error, el valor obtenido no es conclusivo.

En todo caso, de los cálculos y las simulaciones hechos en esta tesis, pueden obtenerse varias conclusiones relevantes, que se presentan en el siguiente capítulo.

### 5.7 Referencias del capítulo 5

- [1].- P. George, R. J. Witonsky, M. Trachtman, C. Wu, W. Dorwart, L. Richman, W. Richman, F. Shurayh y B. Lentz, *Biochim, Biophys. Acta* **223**, 1 (1970)
- [2].- T. C. Waddington, *Adv. Inorg. Chem. Radiochem.* **1**, 158 (1959)
- [3].- M. D. Hayes, L. G. Kenyon y P. A. Kollman, *J. Am. Chem. Soc.* **100**, 4331 (1978)
- [4].- B. Pullman, A. Pullman, H. Berthod y N. Gresh, *Theor. Chim. Acta* **40**, 93 (1975)
- [5].- H. Berthod y A. Pullman, *J. Comput. Chem.* **2**, 87 (1981)
- [6].- E. Clementi, W. Kolos, G. C. Lie y G. Ranghino, *Intl. J. Quantum Chem.* **17**, 337 (1980)
- [7].- J. Detrich, G. Corongiu y E. Clementi, *Chem. Phys. Lett.* **112**, 426 (1984)
- [8].- E. Clementi y S. Chin, eds., *Structure and Dynamics of Nucleic Acids. Proteins and Membranes*, Plenum Publishing Corp. (1986)
- [9].- P. A. Bash, M. J. Field, R. C. Davenport, G. A. Petsko, D. Ringe y M. Karplus, *Biochemistry* **30**, 5826 (1991)
- [10].- G. Alagona, C. Ghio y P. A. Kollman, *J. Am. Chem. Soc.* **107**, 2229 (1985)
- [11].- W. L. Jorgensen, *J. Chem. Phys.* **77**, 4156 (1982)
- [12].- W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey y M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983)
- [13].- D. W. Woods, en *Water, a Comprehensive Treatise*, **6**, F. Franks, ed. cap. 6, Plenum Press, Nueva York (1980)
- [14].- G. M. Torrie y J. P. Valleau, *J. Comput. Phys.* **23**, 187 (1977)

- [15].- C. G. Gray y K. E. Gubbins, *Theory of Molecular Fluids*, vol. 1, cap. 3, Clarendon Press, Oxford (1984)
- [16].- F. H. Stillinger y A. Rahman, *J. Chem. Phys.* 60, 1545 (1974)
- [17].- H. J. C. Berendsen, J. P. M. Postma, W. F. von Gunsteren y J. Hermans, en *Intermolecular Forces*, B. Pullman, ed., Reidel, Dordrecht, Holland (1981), p. 331
- [18].- E. Clementi, *J. Phys. Chem.* 89, 4426 (1985)
- [19].- A. Wallqvist, P. Ahlström y G. Karlström, *J. Phys. Chem.* 94, 1649 (1990)
- [20].- U. Niesar, G. Corongiu, E. Clementi, G. R. Kneller y D. K. Bhattacharya, *J. Phys. Chem.* 94, 7949 (1990)
- [21].- H. Saint-Martin, C. Medina-Llanos e I. Ortega-Blake, *J. Chem. Phys.* 93, 6448 (1990)
- [22].- L. Gil-Adalid e I. Ortega-Blake, *J. Chem. Phys.* 94, 3748 (1991)
- [23].- U. Niesar, G. Corongiu, M. J. Huang, M. Dupuis y E. Clementi, *Intl. J. Quantum Chem., Quantum Chem. Symp.* 23, 421 (1989)
- [24].- J. Hernández-Cobos (en preparación)
- [25].- M. Moreno-Bello, M. Bonilla-Marín, J. Hernández-Cobos e I. Ortega-Blake, (en preparación)
- [26].- El programa potop, que se basa en la rutina VA05AD, *Harwell Subroutine Library*, Division of Computer Science and Systems, Harwell Laboratories, Harwell, U. K. (1974)
- [27].- N. Pastor, comunicación personal
- [28].- R. Sonnenschein y K. Heinzinger, *Chem. Phys. Lett.* 102, 550 (1983)
- [29].- M. J. Mitchell y J. A. McCannnon, *J. Comput. Chem.* 12, 271 (1991)
- [30].- B. Jayaram y D. L. Beveridge, *J. Phys. Chem.* 94, 7288 (1990)
- [31].- N. E. Dorsey, *Properties of Ordinary Water Substances*, Rheinhold, Nueva York (1940)
- [32].- P. H. Berens, D. H. J. Mackay, G. M. White y K. R. Wilson, *J. Chem. Phys.* 79, 2375 (1983)
- [33].- O. Matsuoka, E. Clementi y M. Yoshimine, *J. Chem. Phys.* 64, 1351 (1976)
- [34].- G. C. Lie, E. Clementi y M. Yoshimine, *J. Chem. Phys.* 64, 2314 (1976)

- [35].- M. Neumann, *Mol. Phys.* 50, 841 (1983)
- [36].- M. P. Allen y D. U. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford (1990)
- [37].- J. P. Valleau y S. G. Whittington, *A Guide to Monte Carlo for Statistical Mechanics, en Modern Theoretical Chemistry*, No. 5, B. J. Berne, ed., cap. 4, Plenum, Nueva York y Londres (1977)
- [38].- D. J. Adams, E. M. Adams y G. J. Hills, *Mol. Phys.* 38, 387 (1979)
- [39].- S. W. de Leeuw, J. W. Perram y E. R. Smith, *Proc. R. Soc. Lond.* A 373, 27 (1980)
- [40].- T. A. Andrea, W. C. Swope y H. C. Andersen, *J. Chem. Phys.* 79, 4576 (1983)
- [41].- J. Dore, *Physics World* 1, 25 (1988)
- [42].- A. H. Narten y H. A. Levy, *J. Chem. Phys.* 55, 2263 (1971)
- [43].- A. H. Narten, *J. Chem. Phys.* 56, 5681 (1972)
- [44].- W. E. Thiessen y A. H. Narten, *J. Chem. Phys.* 77, 2656 (1982)
- [45].- J. C. Dore, *Water Science Reviews* vol. 1, F. Franks, ed., Cambridge University Press (1985)
- [46].- A. K. Soper, *Chem. Phys.* 88, 187 (1984)
- [47].- A. K. Soper y M. G. Phillips, *Chem. Phys.* 107, 47 (1986)
- [48].- G. C. Lie, *J. Chem. Phys.* 85, 7495 (1986)
- [49].- A. Guinier, *X - Ray Diffraction*, W. H. Freeman, San Francisco (1963)
- [50].- G. S. Kell, *Thermodynamic and Transport Properties of Fluid Water. en Water: a Comprehensive Treatise*, vol. 1, F. Franks, ed., Plenum Press, Nueva York y Londres (1972)
- [51].- J. R. Reimers, R. O. Watts y M. L. Klein, *Chem. Phys.* 64, 95 (1982)
- [52].- A. Wallqvist, *Chem. Phys.* 148, 439 (1990)
- [53].- D. Eisenberg y W. Kauzmann, *The Structure and Properties of Water*, Oxford University Press, Nueva York (1969)
- [54].- I. Ortega-Blake, J. C. Barthelat, E. Costes, E. Oliveros y J. P. Daudey, *J. Chem. Phys.* 76, 4130 (1982)

DISCUSIÓN, CONCLUSIONES  
Y PERSPECTIVAS

## 6.1 Discusión y conclusiones

Ya en los capítulos 4 y 5 se han presentado los resultados obtenidos en esta tesis, se han comparado con otros resultados, y se ha discutido sobre su confiabilidad y sobre sus implicaciones en la bioquímica del pirofosfato. Aquí se hará una recapitulación de las conclusiones más relevantes a que se ha llegado.

El resultado más importante de esta tesis, es haber mostrado que el esquema cálculos *ab initio* → potenciales de interacción → simulaciones numéricas, puede aplicarse al estudio de los fenómenos bioquímicos, y producir una descripción basada en primeros principios, que contribuye a una mejor comprensión de los datos experimentales.

En el caso que aquí se estudió, la alta calidad de los experimentos sobre la hidrólisis del pirofosfato, ya ha permitido distinguir el efecto que cada condición fisicoquímica tiene sobre los parámetros termodinámicos de la reacción, en la solución acuosa. La explicación de la alta reactividad se había hecho en términos de las interacciones de los reactantes con los solventes, y suponía que la hidrólisis era termoneutra en la fase gaseosa [1]. Ya que no ha sido posible hacer el experimento en esa fase, debido a la inestabilidad de los anhídridos de fosfato, se ha empleado la alternativa de obtener una estimación de la entalpía por medio de cálculos *ab initio* [2 - 4]; pero los resultados no han permitido una conclusión clara.

Como parte de esta tesis, se efectuaron cálculos al nivel  $MBPT(2)/6 - 31G^{**}$  que, con la inclusión del efecto de la energía del punto cero (ZPE), muestran que la entalpía de la reacción de hidrólisis del pirofosfato es de  $\Delta H_g = -3.8 \pm 1.0 \text{ kcal/mol}$ . Se asignó el valor de  $\pm 1.0 \text{ kcal/mol}$  a la incertidumbre, basándose en que los resultados de cálculos al mismo nivel, aplicados a otras reacciones isodérmicas, tienen desviaciones incluso menores con respecto a datos experimentales [5, 6]. Además, aquí se tuvo el cuidado de llegar a una convergencia razonable al límite HF, y a la geometría óptima de cada molécula calculada.

La magnitud de  $\Delta H_g$  muestra que la hidrólisis del pirofosfato neutro es exotérmica en la fase gaseosa. Sin embargo, el valor obtenido es apenas la mitad del valor experimental en la solución acuosa [7],  $\Delta H_{ag} = -7.6 \text{ kcal/mol}$ . Por consiguiente, se concluye que la energía de hidratación contribuye en igual medida que los efectos intramoleculares, a la entalpía en solución acuosa. Además, el dato experimental del término entrópico,  $T\Delta S_{ag} = -1.9 \text{ kcal/mol}$ , muestra que éste también es un factor importante en la energía libre de la reacción. Los datos experimentales de De Meis [8] apoyan esta última aseveración, aunque se refieren a otras especies iónicas del pirofosfato. La determinación del efecto que se obtiene al cambiar la composición del solvente, indica que el término entrópico es el más afectado.

Hasta ahora, la interpretación de estos resultados se ha hecho con la propuesta de que hay un cambio en la actividad del solvente. Este concepto fisicoquímico se aplica a las desviaciones que presentan las soluciones reales con respecto al caso ideal, y se adjudica a las interacciones intermoleculares. Al igual que ocurre con otros parámetros empíricos, la actividad no define un modelo molecular único. El desarrollo de una imagen molecular

clara requiere de un método que permita estudiar la estructura del estado líquido, así como las interacciones intermoleculares en detalle. Esto puede conseguirse por medio de las simulaciones numéricas.

La aportación de esta tesis fue la elaboración y el refinamiento de un potencial analítico para describir las interacciones entre diversas moléculas: el potencial MCHO. Este modelo hace posible reproducir los momentos multipolares, tanto permanentes como inducidos, de las moléculas. Aunque originalmente se diseñó para describir al agua, la misma forma funcional puede aplicarse a otras moléculas. Ya que los parámetros del potencial se ajustan para reproducir superficies de energía calculadas *ab initio*, se logra mantener una descripción basada en primeros principios.

Para efectuar las simulaciones, las características del modelo requieren ajustar las posiciones de las cargas móviles a las nuevas interacciones, cada vez que se mueva una molécula. Esto hizo necesario elaborar un programa de simulación que incluyera este efecto. El seguimiento de las posiciones de los núcleos y de las cargas, permite que en el análisis de resultados, se puedan estimar cantidades relacionadas con el comportamiento de los momentos eléctricos inducidos. Por ello, además de rutinas para la evaluación de energías de interacción, de capacidad calorífica y de energías libres de Helmholtz; se elaboraron otras para evaluar la magnitud del momento dipolar promedio, su orientación y su correlación. También se hicieron programas para generar las funciones de correlación por pares, y las funciones de forma correspondientes a la dispersión de rayos X y de neutrones. Estos programas se mantienen en continua evolución, ampliándose para atacar los problemas que son de interés en el Laboratorio de Biofísica del IFUNAM. Esto hace difícil la elaboración de manuales, ya que se necesitaría actualizarlos constantemente. Sin embargo, ya se encuentra en proceso una guía básica.

El potencial MCHO se usó en simulaciones de tres sistemas: agua pura, agua con ortofosfato y agua con pirofosfato. En el curso de las simulaciones se encontró la necesidad de corridas muy largas para llegar al equilibrio, aun para el caso del agua pura. Una vez que se efectuó un número grande de configuraciones ( $\sim 2 \times 10^7$ ), se compararon los resultados de la simulación del agua pura con datos experimentales; la energía molecular está ligeramente sobrestimada, en tanto que la estructura es la mejor que se ha obtenido hasta la fecha. A continuación, las energías de hidratación de los fosfatos se compararon con cifras deducidas a partir de datos experimentales, y se encontró un acuerdo razonable. Sin embargo, la contribución a la entalpía de la hidrólisis del pirofosfato resultó sobrestimada. Esto llevó a revisar las limitaciones del modelo usado, y se concluyó que la aproximación de moléculas rígidas no es adecuada para la descripción del pirofosfato.

A pesar de las inexactitudes en los resultados, fue posible obtener una descripción que ayuda a aclarar las propuestas sobre la *alta energía* de los compuestos fosforilados. Al combinar los resultados teóricos con los datos experimentales, se encuentra evidencia suficiente para concluir que la energía libre de la hidrólisis del pirofosfato se debe a tres factores: la entalpía debida a las interacciones intramoleculares, la energía de hidratación ocasionada por las interacciones con el solvente, y el cambio de entropía entre los reactantes y los productos. Cada uno de estos factores tiene una contribución igualmente importante.

Entonces se puede proponer una explicación de los resultados experimentales de la hidrólisis en distintos solventes, en términos de las interacciones intermoleculares con el solvente y de los cambios entrópicos, debidos a la reestructuración ocasionada por la presencia de los solutos. Esta descripción es más precisa y tiene mayor significado físico, que aquélla que emplea conceptos como el de la actividad del agua.

En cuanto a las explicaciones sobre los resultados experimentales de De Meis [1, 8, 9], los resultados señalan que el término entrópico  $T\Delta S$ , puede dominar la energía libre de las reacciones de hidrólisis de los anhídridos de fosfato en distintos solventes. Cabe también la posibilidad de que la presencia de los solutos modifique las interacciones entre las moléculas del solvente.

La conclusión general del trabajo es que el uso del potencial MCHO, y de los programas hechos para la simulación y el procesamiento de datos, permite una descripción detallada de las soluciones acuosas, y los resultados que produce son razonables; aunque están sujetos a errores. La ventaja más importante del esquema usado, es que al conocer todas las aproximaciones hechas en el modelo, se puede estimar la magnitud del efecto de cada una, e incluso proponer la forma de superar las posibles deficiencias, refinando con ello la descripción del fenómeno de interés. Además, la construcción a partir de primeros principios hace posible distinguir cuáles factores son los más relevantes en un fenómeno dado, y cuáles son secundarios.

## 6.2 Perspectivas

Aquí se ha desarrollado un método de estudio sistemático sobre la energía de la reacción de hidrólisis del pirofosfato neutro en solución acuosa, partiendo de cálculos *ab initio*, para representar la fase gaseosa, y se ha llegado hasta la simulación numérica de la solución acuosa. De este estudio sistemático se obtuvieron criterios para juzgar la confiabilidad de los resultados de la aproximación teórica, a la descripción de los fenómenos bioenergéticos. Se llegó así a una imagen molecular sobre el origen de la *alta energía* de la reacción en términos de la entalpía, debida tanto a las interacciones intramoleculares como a las energías de hidratación, y a una especulación sobre la contribución del término entrópico, debido a la posible reestructuración del solvente. Para refinar esta imagen será necesario: (1) analizar los datos estructurales que se obtengan de la simulación y (2) hacer otros cálculos y simulaciones en que se permita el movimiento de los hidrógenos en los fosforilos, y de los residuos de fosfato, con respecto al puente  $P - O - P$ .

En tanto que los resultados obtenidos para la fase gaseosa son de una alta calidad, de forma que se puede proponer una incertidumbre de  $\pm 1 \text{ kcal/mol}$ , los que se obtuvieron de las simulaciones para la solución acuosa todavía son inexactos. En este sentido, la aportación de esta tesis, es la de señalar la necesidad de simulaciones muy largas para llegar a estimaciones confiables de las energías de hidratación.

Por otra parte, aquí se han empleado únicamente las especies eléctricamente neutras, mientras que en las condiciones fisiológicas son más relevantes las especies aniónicas; así que habrá de aplicarse el método aquí desarrollado al estudio de los aniones. Lo cual presenta

el problema de considerar moléculas de capa abierta, que tienen distintas multiplicidades y son por ello difíciles de manejar con cálculos *ab initio*, en especial las doblemente cargadas. En cuanto a las simulaciones, es probable que las energías de hidratación resulten más confiables, por el simple hecho de ser mayores. Pero habrá que tener cuidado con el manejo de las interacciones de largo alcance, ya que podrían producir efectos espurios.

Como se mencionó en la sección 1.5, el verdadero sustrato de la pirofosfatasa es el  $MgP_2O_7^{2-}$ , por lo cual también es necesario tomar en cuenta el magnesio. Además del problema que representa el cálculo del anión doblemente cargado, la inclusión del magnesio incrementa de manera significativa el número de funciones que debe agregarse a la base, para lograr una descripción correcta de la molécula.

La posibilidad de estudiar todas las situaciones arriba mencionadas, ya está siendo explorada en el Laboratorio de Biofísica del IFUNAM, en un proyecto conjunto con el Departamento de Química de la Universidad de Arizona y el Instituto de Química Teórica de la Universidad de Varsovia.

### 6.3 Referencias del capítulo 6

- [1].- L. De Meis, *Biochim, Biophys. Acta* **973**, 333 (1989)
- [2].- M. D. Hayes, L. G. Kenyon y P. A. Kollman, *J. Am. Chem. Soc.* **100**, 4331 (1978)
- [3].- M. O'Keeffe, B. Domengès y G. V. Gibbs, *J. Phys. Chem.* **89**, 2304 (1985)
- [4].- C. S. Ewig y J. R. Van Wazer, *J. Am. Chem. Soc.* **110**, 79 (1988)
- [5].- G. P. Ford y C. T. Smith, *J. Am. Chem. Soc.* **109**, 1325 (1986)
- [6].- W. J. Hehre, L. Radom, P. V. R. Schleyer y J. A. Pople, *Ab Initio Molecular Orbital Theory*. Wiley Interscience, N. Y. (1986).
- [7].- P. George, R. J. Witonsky, M. Trachtman, C. Wu, W. Dorwart, L. Richman, W. Richman, F. Shurayh y B. Lentz, *Biochim, Biophys. Acta* **223**, 1 (1970)
- [8].- L. De Meis, *J. Biol. Chem.* **259**, 6090 (1984)
- [9].- P. J. Romero y L. De Meis, *J. Biol. Chem.* **264**, 7896 (1989)

## APÉNDICE A

---

ALGUNOS CONCEPTOS  
FISICOQUÍMICOS  
USADOS EN LA BIOQUÍMICA

### A.1 Los parámetros termodinámicos

Como ocurre con todos los conceptos fundamentales de la Física, la definición de la cantidad llamada *energía*, presenta una serie de problemas, que muchos libros de texto evitan, al definirla simplemente como "la capacidad de realizar trabajo". La idea de trabajo, en cambio, tiene una definición operacional que la conecta directamente con los conceptos básicos de masa, posición y aceleración: "el trabajo  $dW$  hecho sobre un cuerpo, es el producto del desplazamiento  $dr$  producido, por una fuerza  $F$  ejercida sobre un cuerpo, en la dirección del desplazamiento", lo cual puede expresarse de la forma:

$$dW = F \cdot dr \quad (A - 1).$$

El libro de *Mecánica Clásica* del Dr. Juan de Oyarzábal [1], presenta una discusión muy amplia acerca del concepto de la energía en la Mecánica, así como su desarrollo histórico. El lector interesado puede recurrir al excelente libro *The Story of Physics* de Motz y Weaver [2].

Por otra parte, en sus *Lecciones de Física* [3], Feynman escribe que "en la Física actual no conocemos lo que es la energía, pero podemos pensar en la conservación de la energía como un principio matemático abstracto que es aplicable a todos los procesos conocidos, y que nos permite calcular una cantidad numérica que permanece constante, siempre que se tomen en cuenta todas las diferentes formas en que se presenta la energía: gravitacional, cinética, térmica, calórica, elástica, eléctrica, química, radiante, nuclear, de masa."

Es en este sentido matemático abstracto que la termodinámica emplea los conceptos de la energía  $U$  de un sistema, del trabajo  $W$  que puede realizar, y del calor  $Q$  que se le puede transferir, combinándolos en una expresión de la conservación de la energía que se conoce como la *primera ley de la Termodinámica*:

$$dU = dQ - dW \quad (A - 2).$$

La energía  $U$  es entonces un potencial  $U(x_1, \dots, x_n)$  en el espacio  $n$ -dimensional de estados de un sistema dado, en tanto que el trabajo termodinámico resulta ser una extensión natural del trabajo mecánico, de modo que siempre puede reducirse a la evaluación del producto de una fuerza por un desplazamiento, aun cuando en el equilibrio termodinámico, las cantidades que aparezcan en la expresión apropiada, se hayan transformado para representar cantidades termodinámicas [4].

Siguiendo con la definición matemática, el trabajo efectuado en un proceso reversible, es una integral de línea evaluada a lo largo de la curva que representa al proceso en el espacio  $(x_1, \dots, x_n)$ . El integrando tiene la forma

$$dW = \sum_{r=1}^m Y_r(x_1, \dots, x_n) dz_r(x_1, \dots, x_n) \quad (A - 3).$$

que consiste de  $m$  productos de propiedades intensivas  $Y_r$  (como por ejemplo la presión  $P$ ) con diferenciales exactas de propiedades extensivas  $z_r$  (como por ejemplo el volumen  $V$ ).

El concepto de calor no tiene un equivalente mecánico y, por lo tanto, debe definirse en términos puramente fenomenológicos: al poner en contacto dos sistemas con temperaturas distintas, al cabo de un cierto tiempo ambos terminarán con la misma temperatura, que será intermedia entre las dos originales. Entonces se dice que hubo una *transferencia de calor* del sistema que tenía mayor temperatura, hacia el que la tenía menor. La dificultad para definir el concepto de calor puede apreciarse al leer cualquier texto sobre termodinámica, o sobre mecánica estadística, en los que se encuentran aseveraciones como la siguiente: "heat is what is absorbed by a homogeneous system if its temperature increases, while no work is done" [5] (el calor es lo que absorbe un sistema homogéneo cuando su temperatura aumenta, sin que se efectúe trabajo). Una definición matemática del calor, se obtiene al considerar que cuando se transfiere calor a un sistema, su temperatura aumenta en forma proporcional:

$$dQ = C dT \quad (A - 4),$$

donde la constante  $C$  representa una propiedad extensiva de cada sistema particular, y se denomina *capacidad calorífica*. El trabajo y el calor no son formas de energía, sino que sólo representan las dos formas diferentes de transmisión de energía que se consideran en la Termodinámica y, por lo tanto, caracterizan a un proceso, y no a un estado. Sin embargo, el calor  $dQ$  transferido a un sistema, puede relacionarse con un potencial termodinámico  $S(x_1, \dots, x_n)$ , lo cual queda expresado al aplicar la *segunda ley de la Termodinámica* a un proceso reversible:

$$dS(x_1, \dots, x_n) = \frac{dQ}{T(x_1, \dots, x_n)} \quad (A - 5).$$

A este potencial  $S$  se le denomina *entropía*, y el factor integrante  $T$  es la temperatura termodinámica. La ec. (A-5) muestra que el potencial  $S$  es singular en  $T = 0$ , así que su comportamiento conforme  $T \rightarrow 0$  sólo puede inferirse a partir de un principio que es suplementario a la segunda ley (es decir, la tercera ley).

Además de la definición del potencial entropía  $S$ , la segunda ley asevera que en un sistema aislado, la entropía  $S_2$  de un estado final, es siempre mayor que la entropía  $S_1$  de un estado inicial, luego de que haya ocurrido un proceso irreversible:

$$S_2 > S_1 \quad (A - 6),$$

lo que significa que el cambio de entropía es positivo,  $\Delta S = S_2 - S_1 > 0$ .

La segunda ley expresa matemáticamente el hecho experimental de que los sistemas aislados evolucionan hacia estados con mayor entropía, al relajar las restricciones externas.

Al combinar las ecs. (A-1), (A-2) y (A-3) se obtiene una expresión para la energía en términos de diferenciales exactas:

$$dU = TdS - \sum_{r=1}^m Y_r dz_r \quad (A - 7),$$

en cuyo caso la energía interna  $U$  de un sistema es función de las  $m + 1$  variables  $S, z_1, \dots, z_m$ :

$$U = U(S, z_1, \dots, z_m) \quad (A - 8)$$

y la dimensión del espacio de estados es  $n = m + 1$ . Además, derivando parcialmente se obtienen las relaciones:

$$T = \left( \frac{\partial U}{\partial S} \right)_{z_r}, Y_r = \left( \frac{\partial U}{\partial z_r} \right)_{S, z_q} \quad (A-9)$$

y

$$\left( \frac{\partial T}{\partial z_r} \right)_{S, z_q} = - \left( \frac{\partial Y_r}{\partial S} \right)_{z_r} \quad (A-10).$$

Se puede demostrar [4] que todas las propiedades de equilibrio de cualquier sistema, se pueden deducir a partir de la ecuación de estado (A-8).

Al aplicar transformaciones de Legendre a la ec. (A-8), se pueden obtener otras ecuaciones de estado, con las que se definen otros potenciales termodinámicos. Comúnmente se parte de las variables que caracterizan a un gas:

$$Y_r = P, z_r = V, m = 1. \quad (A-11),$$

de donde

$$dU = TdS - PdV \quad (A-12),$$

con las transformadas de Legendre

$$dH = d(U + PV) = TdS + VdP \quad (A-13)$$

$$dF = d(U - TS) = -SdT - PdV \quad (A-14)$$

$$dG = d(H - TS) = -SdT + VdP \quad (A-15).$$

Los potenciales

$$H(S, P) = U + PV \quad (A-16)$$

$$F(T, V) = U - TS \quad (A-17)$$

$$G(T, P) = H - TS = U + PV - TS \quad (A-18)$$

se denominan la *entalpía*, la *función de Helmholtz* y la *función de Gibbs*, respectivamente. Para cada potencial se pueden escribir dos relaciones de reciprocidad análogas a las ecs. (A-9), y otra análoga a la ec. (A-10).

Cuando se consideran sistemas abiertos, la ec. (A-7) adopta la forma

$$dU = TdS - PdV + \sum_{j=1}^k \mu_j dn_j \quad (A-19),$$

con las respectivas transformadas

$$dH = d(U + PV) = TdS + VdP + \sum_{j=1}^k \mu_j dn_j \quad (A-20)$$

$$dF = d(U - TS) = -SdT - PdV + \sum_{j=1}^k \mu_j dn_j \quad (A-21)$$

$$dG = d(H - TS) = -SdT + VdP + \sum_{j=1}^k \mu_j dn_j \quad (A-22).$$

En este caso,  $n_j$  denota el número de moles de una de las  $k$  especies químicas presentes en el sistema, y  $\mu_j$  denota el potencial químico de la especie  $j$  en la mezcla. Las ecuaciones de estado del sistema, adoptan las formas siguientes:

$$U = U(S, V, n_1, \dots, n_k) \quad (A - 23)$$

$$H = H(S, P, n_1, \dots, n_k) \quad (A - 24)$$

$$F = F(T, V, n_1, \dots, n_k) \quad (A - 25)$$

$$G = G(T, P, n_1, \dots, n_k) \quad (A - 26).$$

El potencial  $G$  es extensivo y contiene dos parámetros intensivos,  $T$  y  $P$ , y  $k$  parámetros extensivos,  $n_j$ , por lo que debe ser homogéneo de orden 1 en las variables  $n_j$ :

$$G = \sum_{j=1}^k n_j \mu_j \quad (A - 27).$$

Al tomar la diferencial total  $dG$  y compararla con la ec. (A-22) se llega a la relación de Gibbs - Duhem - Margules :

$$-SdT + VdP - \sum_{j=1}^k n_j \mu_j = 0 \quad (A - 28).$$

## A.2 Los potenciales termodinámicos en las reacciones químicas

Aunque se han introducido los potenciales  $U$ ,  $S$ ,  $H$ ,  $F$  y  $G$  de una forma abstracta, cada uno puede aplicarse a la descripción de diversos fenómenos, con una interpretación física basada en cantidades mensurables. Una de estas cantidades es la temperatura, cuya variación puede relacionarse con la transferencia de calor: se puede aumentar la energía interna de un sistema al transferirle calor, sin que se efectúe ningún trabajo. El incremento de energía interna será:

$$dU = dQ \quad (A - 29).$$

Cuando se transfiere calor a un sistema, su temperatura aumenta en forma proporcional:

$$dQ = CdT \quad (A - 4).$$

La constante  $C$  se denomina la *capacidad calorífica* del sistema, y es una propiedad extensiva. Para que el sistema no efectúe trabajo, se mantiene con un volumen constante, y se obtiene entonces la capacidad calorífica a volumen constante,  $C_V$ . Alternativamente, la transferencia de calor puede tener lugar en un proceso con presión constante, permitiendo al sistema contraerse o expandirse; y entonces se habla de la capacidad calorífica a presión constante,  $C_P$ . Los dos procesos descritos son esencialmente distintos, puesto que cuando la presión se mantiene constante, el sistema ha realizado trabajo al cambiar de tamaño: una parte del calor transferido fue convertida en trabajo de expansión, por lo que no todo el calor dio lugar a un cambio en la temperatura del sistema.

Las dos distintas capacidades caloríficas se definen operacionalmente como:

$$C_V = \left( \frac{\partial Q}{\partial T} \right)_V, C_P = \left( \frac{\partial Q}{\partial T} \right)_P \quad (A - 30).$$

De la ec. (A-11) se llega a que:

$$C_V = \left( \frac{\partial U}{\partial T} \right)_V \quad (A - 31).$$

De forma similar, de (A-12) se obtiene:

$$C_P = \left( \frac{\partial H}{\partial T} \right)_P \quad (A - 32),$$

y entonces se puede identificar al cambio de entalpía de un sistema, con el calor transferido cuando la presión se mantiene constante:

$$\Delta H = Q_P \quad (A - 33).$$

La transferencia de calor producida por una reacción química, puede medirse en un calorímetro y así se determina el cambio de entalpía producido.

Usualmente, los valores tabulados para los cambios de entalpía en las reacciones químicas, se refieren a los reactantes y a los productos en ciertos estados específicos. La *entalpía estándar* de una reacción a una temperatura  $T$ , se denota como  $\Delta H^0(T)$  y es el valor de la diferencia  $H(\text{productos}) - H(\text{reactantes})$  a esa temperatura cuando todos los participantes están en sus *estados estándar*. Para cada sustancia, el estado *estándar* se define como el más estable bajo una presión de 1 atm, a la temperatura especificada. En las aplicaciones prácticas generalmente se considera  $T = 298$  K.

Ya que sólo los cambios en entalpía  $\Delta H$  tienen sentido físico, es conveniente adoptar la convención de que la entalpía de cada elemento en su estado estándar es nula.

La diferencia entre el cambio de entalpía  $\Delta H$  y el de energía interna  $\Delta U$  consiste en un término  $\Delta(PV)$  (cfr. ec. (A-15)),

$$\Delta H = \Delta U + \Delta(PV) \quad (A - 34),$$

que se reduce a  $P\Delta V$  en los procesos a presión constante. Si además el sistema conserva el mismo volumen, entonces  $\Delta H$  y  $\Delta U$  son iguales.

En el caso de moléculas, las entalpías de formación están relacionadas con las energías de los enlaces interatómicos. Estas energías pueden determinarse al conocer la entalpía de la reacción



en la que  $A$  y  $B$  representan grupos de átomos y  $(g)$  significa que la reacción ocurre en la fase gaseosa.

Cuando se consideran soluciones acuosas, se hace necesario tomar en cuenta las interacciones del soluto con el solvente, y se habla de la *entalpía de hidratación*.

Algunas reacciones químicas requieren del suministro de energía para ocurrir, y se denominan *endergónicas* o *endotérmicas*, en tanto que otras transfieren calor al ambiente, y se llaman *exergónicas* o *exotérmicas*. Dado que el cambio de entalpía  $\Delta H$  en una reacción a presión constante, es igual al calor transferido  $Q_p$ , el signo de  $\Delta H$  indica si la reacción es endergónica,  $\Delta H > 0$  (se requiere de energía), o exergónica,  $\Delta H < 0$  (se libera energía). Esta clasificación sólo indica si la transferencia de energía es hacia adentro o hacia afuera del sistema, pero no constituye un criterio para saber si la reacción ocurrirá de forma espontánea, i. e., si basta con colocar los reactantes juntos para obtener los productos. Para juzgar la espontaneidad de una reacción química, es necesario recurrir a la segunda ley de la Termodinámica y hacer uso del potencial entropía,  $S$ . Una forma de expresar esta segunda ley es la de decir que en un sistema aislado los procesos espontáneos ocurren en la dirección en que aumenta la entropía.

Es posible relacionar el cambio de entropía de un proceso a presión constante con el cambio de entalpía (cfr. [6]):

$$\Delta S_p = \int_{T_i}^{T_f} \frac{dH}{T} \quad (A-36)$$

En la química comúnmente se trata con sistemas que están en contacto con el ambiente (i. e., que no están aislados), así que el cálculo de la entropía debe considerar al "universo" (el sistema y sus alrededores). Sin embargo, para los procesos a volumen constante, se puede reescribir la ec. (A-5) de la forma:

$$TdS \geq dU \quad (A-37)$$

y para los procesos a presión constante,

$$TdS \geq dH \quad (A-38)$$

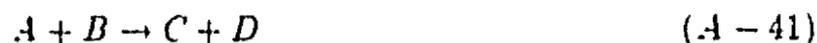
donde  $T$ ,  $dS$ ,  $dU$  y  $dH$ , se refieren únicamente al sistema. Usando los potenciales  $F$  (ec. A-16) y  $G$  (ec. A-17) esto es equivalente a escribir:

$$dF \leq 0 \quad (V, T, \text{ctes.}) \quad (A-39)$$

$$dG \leq 0 \quad (P, T, \text{ctes.}) \quad (A-40)$$

El cambio en la función de Helmholtz,  $\Delta F$  indica la máxima cantidad de trabajo que puede realizar un sistema, mientras que el cambio en la función de Gibbs,  $\Delta G$ , se refiere al máximo de trabajo disponible que no sea usado en expansión. A este trabajo se le llama *trabajo neto*.

En la química se usa con mayor frecuencia la función de Gibbs, pues generalmente se consideran reacciones que ocurren a presión constante. Entonces si en una reacción química



$\Delta G$  es negativa, significa que la reacción ocurrirá de forma espontánea como está escrita. En cambio, si  $\Delta G$  es positiva, entonces la reacción será espontánea en sentido contrario.

Es interesante notar que una reacción endotérmica puede ser espontánea. En efecto,

$$\Delta G = \Delta H - T\Delta S \quad (A - 42),$$

y aunque  $\Delta H > 0$ , puede ocurrir que el cambio en entropía sea tal que  $T\Delta S > \Delta H$ , de forma que  $\Delta G < 0$  y por lo tanto la reacción ocurra espontáneamente [7].

### A.3 El potencial químico, la fugacidad y la actividad

En la discusión anterior se ha hecho uso de la termodinámica para establecer criterios sobre la espontaneidad de una reacción. Sin embargo, la experiencia indica que aun cuando una reacción sea espontánea (*i. e.*,  $\Delta G < 0$ ), no basta con colocar los reactantes en un tubo de ensayo para obtener los productos en un tiempo razonable.

Entonces se necesita desarrollar una teoría de reactividad química más amplia, la cual requiere de consideraciones que no son puramente termodinámicas. Más aún, ya que las manipulaciones del formalismo termodinámico se hacen con ecuaciones de gas ideal y de soluciones ideales, para su aplicación a situaciones reales es necesario definir una serie de cantidades para corregir las desviaciones con respecto a los valores experimentales. Las cantidades que se definen tienen una cierta relación con una visión molecular de la materia, pero no parten de ella; es decir que se trata de conceptos empíricos, que no se obtienen directamente de primeros principios. Tal es el caso de la *fugacidad* y de la *actividad*.

En la sección A.1 se hizo una discusión de los potenciales termodinámicos de un gas ideal (ecs. (A - 2) a (A - 18)), y luego se introdujo el término  $\sum \mu_j dn_j$ , al considerar sistemas abiertos. Este término refleja el hecho de que al agregar una partícula a un sistema, se modifica su energía interna. A la variación de la energía con respecto al número de partículas de un tipo  $j$ , se le denomina el *potencial químico* de la especie  $j$ , y puede caracterizarse matemáticamente como:

$$\mu_j = \left( \frac{\partial U}{\partial n_j} \right)_{S,V,n_i} = \left( \frac{\partial F}{\partial n_j} \right)_{T,V,n_i} = \left( \frac{\partial G}{\partial n_j} \right)_{T,P,n_i} = \left( \frac{\partial H}{\partial n_j} \right)_{S,P,n_i} \quad (A - 43).$$

También se puede hablar del potencial químico de un gas ideal, por medio de la función de Gibbs. La dependencia de esta función con respecto de la presión  $P$  está dada por la expresión [6]:

$$G(P_f) = G(P_i) + \int_{P_i}^{P_f} V(P) dP \quad (A - 44).$$

de manera que para un gas ideal

$$G(P_f) = G(P_i) + n RT \ln (P_f/P_i) \quad (A - 45).$$

Al usar el concepto del estado estándar del gas (*vide supra*) y referirse a un mol de material, se define el *potencial químico* de un gas ideal como:

$$\mu(P) = \mu^0 + RT \ln (P/atm) \quad (A - 46).$$

Para aplicar la misma forma funcional a los gases reales, se define la *fugacidad*  $f$  con base en la ecuación

$$\mu(P) = \mu^0 + RT \ln (f/atm) \quad (A - 47).$$

de la cual se obtiene una receta para evaluar la fugacidad de un gas, si además se conoce el valor del *factor de compresión*  $C$  que permita la integración

$$f = P \exp \left[ \int_0^P \left\{ \frac{C(P,T) - 1}{P} \right\} dP \right] \quad (A - 48).$$

El factor  $C$  puede obtenerse si se conoce la ecuación de estado del gas, o bien a partir de un desarrollo del virial en potencias de la presión. En el primer caso, sería necesario contar con un modelo de las interacciones; de ser así, se podría integrar directamente la ec. (A - 44). La alternativa por el desarrollo del virial, requiere del conocimiento de los coeficientes  $B(T)$ ,  $C(T)$ , etc. Éstos se pueden obtener experimentalmente.

La extensión de este tipo de conceptos al caso de mezclas de gases, se hace introduciendo la idea de las *cantidades parciales molares*  $z_{A,m}$ , que se refieren a la modificación que sufre una variable extensiva  $z$  al agregar a una mezcla una cierta cantidad de la sustancia A:

$$z_{A,m} = \left( \frac{\partial z}{\partial n_A} \right)_{P,T,n_B,n_C,\dots} \quad (A - 49).$$

donde  $n_J$  es el número de moles de la sustancia J, presente en el sistema. También resulta útil emplear las *fracciones molares*  $x_J$ , definidas como

$$x_J = \frac{n_J}{\sum_{i=1}^N n_i} \quad (A - 50).$$

Entonces se puede evaluar el cambio de energía libre de una mezcla de dos gases (ideales) como:

$$\Delta G_{mez} = n RT (x_A \ln x_A + x_B \ln x_B) \quad (A - 51).$$

Asímismo, se puede considerar el equilibrio de un líquido con su vapor, usando el potencial químico en términos de la presión del vapor:

$$\mu(l) = \mu(g) = \mu^0 + RT \ln (P/atm) \quad (A - 52).$$

y, de forma más general, para una mezcla, en términos de la presión parcial de la sustancia A,  $P_A$ :

$$\mu_A(l) = \mu_A(g) = \mu_A^0 + RT \ln (P_A/atm) \quad (A - 53).$$

Si el líquido tiene un solo componente, entonces

$$\mu_A^*(l) = \mu_A^0 + RT \ln (P_A^*/atm) \quad (A - 54),$$

en que el asterisco se refiere a un componente puro. La combinación de (A-52) con (A-53) permite eliminar a  $\mu_A^0$  y relacionar el potencial químico de una sustancia en una mezcla,  $\mu_A$ , con el del líquido puro,  $\mu_A^*$ :

$$\mu_A(l) = \mu_A^*(l) + RT \ln (P_A/P_A^*) \quad (A - 55).$$

De esta última relación se define una *solución ideal*, como aquella en que la dependencia de  $P_A$  respecto de  $P_A^*$  es lineal (Ley de Raoult):

$$P_A = x_A P_A^* \quad (A - 56),$$

y por ello

$$\mu_A(l) = \mu_A^*(l) + RT \ln x_A \quad (A - 57).$$

En una mezcla de dos líquidos se define como *solvente* a la componente más abundante, y como *soluto* a la menos abundante. En soluciones que no son ideales, la presión de vapor del soluto se desvía significativamente de la ley de Raoult, pero puede aproximarse linealmente con la ley de Henry

$$P_A = x_A K_A \quad (A - 58),$$

en que  $K_A$  es una constante que se determina empíricamente y tiene las dimensiones de una presión. La ley de Henry define a una *solución diluida ideal*.

La energía libre de una mezcla de dos líquidos en una solución ideal, es entonces la misma que para la mezcla de dos gases ideales (cfr. ec. (A-51)), aunque el significado del vocablo *ideal* es distinto en cada caso; sin embargo, de nueva cuenta la aplicación de la ec. (A-57) a situaciones reales, requiere de una corrección para poder conservar la misma forma funcional. Aquí entra en juego la *actividad*  $a$ , que queda definida por la ecuación:

$$\mu_A(l) = \mu_A^*(l) + RT \ln a_A \quad (A - 59).$$

Así como se hizo con la fugacidad, para utilizar el concepto de la actividad, es necesario relacionarlo con la composición del sistema. Una forma de hacerlo es definir un *coeficiente de actividad*  $\gamma_S$ , que para un solvente es

$$a_S = \gamma_S x_S \quad (A - 60),$$

y que tiene la propiedad de que  $\gamma_S \rightarrow 1$  cuando  $x_S \rightarrow 1$ . Entonces el potencial químico de un solvente es

$$\mu_S = \mu_S^* + RT \ln x_S + RT \ln \gamma_S \quad (A - 61),$$

y el estado estándar se establece cuando el líquido solvente es puro. Ahora la energía libre de una mezcla de líquidos no ideales toma la forma:

$$\Delta G_m = RT(x_A \ln x_A + x_B \ln x_B + x_A \ln \gamma_A + x_B \ln \gamma_B) \quad (A - 62),$$

con lo que se define la función de Gibbs de exceso

$$\Delta G_{E,m} = RT(x_A \ln \gamma_A + x_B \ln \gamma_B) \quad (A - 63).$$

Aquí es importante hacer notar que en los casos ideales, la entropía de la mezcla se calcula como:

$$\Delta S_{mez} = -\left(\frac{\partial \Delta G_{mez}}{\partial T}\right)_{P,n_A,n_B} = -nR(x_A \ln x_A + x_B \ln x_B) \quad (A - 64)$$

y que, consecuentemente,  $\Delta H_{mez} = 0$ . En cambio, si se consideran líquidos no ideales, entonces la entropía de la mezcla se obtiene como:

$$\Delta S_{mez} = -\left(\frac{\partial \Delta G_m}{\partial T}\right)_{P,n_A,n_B} = -nR(x_A \ln x_A + x_B \ln x_B) - \left(\frac{\partial \Delta G_{E,m}}{\partial T}\right)_{P,n_A,n_B} \quad (A - 65),$$

ya que los coeficientes de actividad no son independientes de la temperatura. En este caso, la entalpía de exceso  $\Delta H_{E,m}$  no es necesariamente nula.

En cuanto al soluto, la definición de la actividad y del estado estándar requieren de mayor cuidado, ya que una solución se aproxima al comportamiento de la ley de Henry cuando hay una baja concentración de soluto, muy distante del líquido soluto puro. La igualdad de los potenciales químicos de las fases del líquido y del vapor se expresa como:

$$\mu_s(l) = \mu_s(g) = \mu_s^0 + RT \ln (K_s/atm) + RT \ln x_s \quad (A - 66).$$

Si el soluto siguiera la ley de Henry para toda concentración, entonces el estado estándar del soluto podría definirse cuando  $x_s = 1$ :

$$\mu^{\dagger} = \mu_s^0 + RT \ln (K_s/atm) \quad (A - 67),$$

en cuyo caso

$$\mu_s = \mu^{\dagger} + RT \ln x_s \quad (A - 68),$$

y la actividad del soluto se define con la ecuación

$$\mu_s = \mu^{\dagger} + RT \ln a_s \quad (A - 69),$$

que puede emplearse con una precisión razonable a bajas concentraciones del soluto [6], en que  $a_s \rightarrow x_s$ . En este caso también se define un coeficiente de actividad  $\gamma_s$ :

$$a_s = \gamma_s x_s \quad (A - 70),$$

pero con la propiedad de que  $\gamma_s \rightarrow 1$  cuando  $x_s \rightarrow 0$ .

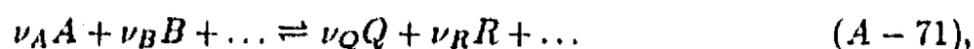
Se han definido las actividades y los estados estándar del solvente y del soluto con base en las fracciones molares  $x_S$  y  $x_s$ , pero también es posible hacerlo usando molaridades o molalidades; aunque los valores cambian, el formalismo es el mismo.

Aunque estrictamente sólo se trata de cantidades definidas empíricamente, se puede hablar de un significado físico de la fugacidad y de la actividad, en términos de una imagen molecular. En efecto, estos conceptos se hicieron necesarios para dar cuenta de las desviaciones respecto de situaciones ideales, que en el caso de los gases consiste en omitir tanto el tamaño como las interacciones de las moléculas del sistema, y en el caso de los líquidos consiste en considerar todas las interacciones intermoleculares iguales. Al aplicar el formalismo a fluidos reales, la fugacidad (o la actividad, según sea el caso) es una forma de cuantificar tales interacciones y compararlas con cantidades mensurables.

En el tratamiento de las soluciones iónicas, p. ej., las interacciones intermoleculares son de gran intensidad y de largo alcance, así que los coeficientes de actividad juegan un papel importante. Una forma de relacionarlos con las propiedades de los iones del sistema, es la teoría de Debye - Hückel, que no será tratada aquí, pues existe una abundante literatura al respecto (e. g. [6, 8]).

#### A.4 El equilibrio y las velocidades de las reacciones

Para el presente trabajo es más relevante abundar sobre la espontaneidad de una reacción química y sobre la rapidez con que ocurra. En este caso no bastarán los conceptos termodinámicos, sino que habrá que emplear algún tipo de modelo molecular. Para comenzar, hay que recordar que las reacciones químicas no continúan hasta la transformación del total de los reactantes en productos, sino que se llega a un equilibrio en el que existen tanto reactantes como productos, pero en el cual las concentraciones de unos y otros ya no varían. Así, para una reacción



donde  $\nu_J$  es el coeficiente estequiométrico de la especie  $J$ , se puede definir una función de Gibbs estándar para la reacción, en la forma:

$$\Delta G^0 = \nu_Q \mu_Q^0 + \nu_R \mu_R^0 + \dots - \nu_A \mu_A^0 - \nu_B \mu_B^0 - \dots \quad (A - 72),$$

y una constante de equilibrio

$$K_p = P_Q^{\nu_Q} P_R^{\nu_R} \dots / P_A^{\nu_A} P_B^{\nu_B} \dots \quad (A - 73),$$

en que se han usado las presiones parciales de vapor de cada especie química, con la consideración de gases ideales, con la cual se tiene la relación:

$$RT \ln K_p = -\Delta G_m^0 \quad (A - 74).$$

Cuando se trata con fluidos reales se emplea el concepto de fugacidad para los gases, y el de actividad para los líquidos, con las correspondientes constantes de equilibrio.

$$K_f = \{f_Q^{\nu_Q} f_R^{\nu_R} \dots / f_A^{\nu_A} f_B^{\nu_B} \dots\}_e \quad (A-75)$$

y

$$K_a = \{a_Q^{\nu_Q} a_R^{\nu_R} \dots / a_A^{\nu_A} a_B^{\nu_B} \dots\}_e \quad (A-76).$$

En las ecs. (A-75) y (A-76) el subíndice  $e$  denota que las cantidades indicadas corresponden a la situación de equilibrio. Ya que las actividades  $a_J$  se relacionan con las fracciones molares  $x_J$  por la ec. (A-60) (para el solvente, y (A-70) para el soluto),  $a_J = \gamma_J x_J$ , se llega a que la constante de equilibrio  $K_a$  puede expresarse como un producto de dos constantes,

$$K_a = K_\gamma K_x \quad (A-77),$$

una que depende de los coeficientes de actividad,  $K_\gamma$ , y la otra de las fracciones molares,  $K_x$ . Se puede llegar a expresiones similares usando las concentraciones molares,  $c_J$ , o las molales,  $m_J$ . En todo caso, la relación con la función de Gibbs depende de las actividades:

$$-\Delta G_m^0 = RT \ln K_a = RT \ln K_x + RT \ln K_\gamma \quad (A-78).$$

Comúnmente se omite el subíndice  $a$  y la constante  $K$  se denomina la *constante de equilibrio termodinámico*.

Hasta ahora se ha hablado únicamente del equilibrio, pero no del tiempo que tarda una reacción en alcanzarlo. Para ello es necesario estudiar las *velocidades* de las reacciones, con lo que además se pueden descubrir sus *mecanismos*. Este último término tiene dos connotaciones, correspondiendo la primera al análisis de una reacción química en una secuencia de pasos elementales. En este sentido, al establecer todos los pasos elementales se ha encontrado el mecanismo de la reacción. El otro significado se refiere a la descripción detallada de los estados energéticos de las moléculas participantes en cada uno de esos pasos.

El primer tipo de análisis sobre el mecanismo, es el tema de estudio de la *cinética química clásica*, en tanto que el segundo corresponde a la *dinámica química*.

Los datos experimentales con los que trabaja la cinética química, son las concentraciones de los reactantes y de los productos como función del tiempo. Para el monitoreo de las concentraciones se usan diversas técnicas, como la medición de cambios en la presión, los métodos espectroscópicos, los polarimétricos, los electroquímicos, los cromatográficos y los de sustitución por isótopos radiactivos. Estas técnicas se aplican de tres maneras distintas: a) el análisis en tiempo real, haciendo el monitoreo con la reacción en marcha; b) la detención (*quenching*), en que la reacción se *congela* después de un cierto tiempo y entonces se analizan las concentraciones, y c) el método del flujo, en que las soluciones con los reactantes se mezclan conforme fluyen hacia una cámara, la reacción continúa y puede ser monitoreada en diferentes posiciones a lo largo de un tubo que sale de la cámara de mezcla.

Para estudiar la tasa de una reacción, se pueden emplear las derivadas temporales de cualquiera de los reactantes o de los productos, usando un signo negativo para las especies que se consumen, y uno positivo para las que se producen. Una manera más adecuada, que incluye los coeficientes estequiométricos, se encuentra al escribir una reacción en la forma:

$$0 = \nu_A A + \nu_B B + \nu_C C + \nu_D D + \dots \quad (A-79),$$

donde los coeficientes  $\nu_J$  tienen signo negativo para los reactantes y positivo para los productos. Entonces se habla del grado de avance  $\xi$  de la reacción y de la tasa verdadera,  $d\xi/dt$ , de manera que el cambio en la especie  $J$  es  $\nu_J d\xi$ , y su tasa de cambio es  $d[J]/dt = \nu_J d\xi/dt$ .

Frecuentemente la tasa de una reacción es proporcional a las concentraciones de los reactantes, elevadas a alguna potencia. Por ejemplo, en una reacción como la (A-71), la tasa depende de las concentraciones de  $A$  y de  $B$ :

$$-\frac{d[A]}{dt} = k[A]^m[B]^n \quad (A-80),$$

donde  $k$  es el coeficiente de velocidad de la reacción, y no depende de las concentraciones. En este caso, la reacción es de orden  $m$  en  $A$ ,  $n$  en  $B$  y  $m+n$  en total. De hecho,  $m$  y  $n$  no son necesariamente números enteros.

Al establecer la ecuación de la tasa (*rate law, rl*) de una reacción, se cuenta con un criterio para juzgar un mecanismo de reacción propuesto: éste debe adecuarse a la *rl*. Aunque en muchos casos la *rl* refleja la estequiometría de la reacción, esto no es siempre cierto: la *rl* se obtiene a partir del experimento y no puede inferirse por el simple análisis de la ecuación de la reacción. Existen diversas técnicas para deducir la *rl* a partir de las medidas de las concentraciones de las especies participantes, y están descritas en cualquier texto de fisicoquímica (e. g. [6, 9]).

#### A.5 Los mecanismos de reacción con base en la teoría de colisiones

Para proponer algún mecanismo de reacción, es necesario continuar el análisis de los datos cinéticos, tratando de separar las reacciones en una serie de pasos que involucren bien una *reacción unimolecular*, en que una sola molécula se "rompe" o cambia de configuración, o bien una *reacción bimolecular*, en que un par de moléculas chocan e intercambian energía, átomos, o grupos de átomos.

Aunque la obtención de la *rl* de una reacción es un primer paso para elucidar el mecanismo, la obtención de este último no es trivial, principalmente debido a que diversos mecanismos pueden acomodarse a la misma *rl*. Esto es lo que se llama el *problema inverso* en el sentido de Hadamard [10], y resulta común a la interpretación de virtualmente todos los datos experimentales. Así, p. ej., el orden de una reacción, no es necesariamente lo mismo que la *molecularidad* de los pasos individuales en que se separe para analizarla. A pesar de ello, una gran cantidad de reacciones de segundo orden, puede describirse sobre la base de colisiones bimoleculares. A estas reacciones se les denomina *simples*.

La velocidad de la mayoría de las reacciones, aumenta al incrementar la temperatura en la forma funcional propuesta por Arrhenius,

$$k = A \exp(-E_a/RT) \quad (A - 81),$$

donde  $R$  es la constante de los gases y  $E_a$  se denomina la *energía de activación* de la reacción. Tanto  $E_a$  como el factor  $A$  (que es independiente de la temperatura) pueden determinarse a partir de una gráfica del  $\ln k$  contra  $1/T$ :

$$\ln k = \ln A - E_a/RT \quad (A - 82).$$

Se puede racionalizar la expresión de Arrhenius, al suponer que el coeficiente  $k$  es el resultado de una reacción bimolecular, en cuyo caso deben ocurrir dos cosas para que la reacción se realice: para comenzar, las moléculas deben reunirse. En un gas esto es una *colisión*, y en un líquido se llama *encuentro*. La tasa con la que ocurren las colisiones (o los encuentros) por unidad de volumen se designa por  $Z$ . Además, es necesario que las moléculas choquen con la suficiente energía  $E_a$  para reaccionar; la fracción de moléculas que tiene al menos esta energía, está dada por un factor de Boltzmann [6],  $e^{-E_a/RT}$ , de forma que el número de colisiones con suficiente energía para reaccionar por unidad de volumen es  $Z e^{-E_a/RT}$ . Esta es casi la relación de Arrhenius, salvo por el hecho de que el factor  $A$  no depende de la temperatura, en tanto que  $Z$  sí lo hace. La teoría de colisiones refleja las características principales que tienen las reacciones bimoleculares, y la energía de activación, es un parámetro útil que tiene un cierto significado físico, pero esta teoría no es completa para la descripción de las tasas de reacción. Una teoría completa debe basarse en la Mecánica Cuántica [11] y todavía está en desarrollo.

Ahora bien, usando el concepto de la energía de activación se puede observar que si ésta es alta, sólo una pequeña porción de las colisiones moleculares tiene suficiente energía para reaccionar; en cambio, si  $E_a$  es baja, la proporción aumenta, y con ella el coeficiente de velocidad. Entonces si la energía de activación puede abatirse de alguna manera, la reacción debe ocurrir más rápido. Esta es la función de un *catalizador*, el cual, por otra parte, no modifica el equilibrio de la reacción, sino sólo la velocidad con que se alcanza. Si el catalizador modificara la posición de equilibrio, entonces podría conseguirse una *máquina térmica* que no necesitaría estar en contacto con un objeto "frío", en contradicción con la segunda ley de la Termodinámica [6, 12]. En efecto, si a una reacción exotérmica que ya hubiera llegado al equilibrio, se le agregara tal catalizador, entonces procedería hasta llegar al nuevo equilibrio, liberando calor en el proceso. Parte de este calor podría transformarse en trabajo. Al retirar el catalizador, la reacción se regresaría al equilibrio anterior, tomando calor de sus alrededores. El sistema habría vuelto a su situación original, pero una parte de la energía del entorno se habría convertido en trabajo.

Por otra parte, el concepto de la energía de activación lleva al desarrollo de la *teoría del complejo activado*: en una reacción



las especies  $A$  y  $B$  se unen para formar un complejo activado  $(AB)^\ddagger$ , que origina los productos  $P$  con alguna tasa  $k^\ddagger$ . La velocidad de la reacción será

$$\frac{d[P]}{dt} = k^\ddagger [(AB)^\ddagger] \quad (A - 84).$$

El problema en este caso consiste en encontrar  $k^\ddagger$  y  $[(AB)^\ddagger]$ . Ya que para dar origen a los productos el complejo activado tiene que atravesar por una configuración crítica (llamada el estado de transición), en la que los enlaces están distorsionados, el paso del complejo activado a través de este estado puede identificarse con un modo de vibración, al que se asocia una frecuencia  $\nu$ . No necesariamente todos los complejos activados que llegan al estado de transición lo atraviesan, sino sólo una parte, que se cuantifica con el coeficiente de transmisión  $\kappa$ . Entonces la constante de velocidad resulta ser:

$$k^\ddagger = \kappa \nu \quad (A - 85).$$

Con esta teoría se puede llegar a definir una constante de equilibrio para el complejo activado,

$$K^\ddagger [A][B] = [(AB)^\ddagger] \quad (A - 86),$$

y a la ecuación de Eyring [6],

$$k_2 = \kappa (k_B T / h) \bar{K}^\ddagger \quad (A - 87),$$

en que  $k_2$  es una constante de velocidad de una reacción de segundo orden,  $k_B$  es la constante de Boltzmann,  $h$  la de Planck, y  $\bar{K}^\ddagger = \frac{h\nu}{k_B T} K^\ddagger$ . Con este tipo de expresiones se desarrolla la teoría del estado de transición para la catálisis enzimática (ver e. g. [13]), y además se puede definir una función de Gibbs de activación con la expresión:

$$\Delta G_m^\ddagger = -RT \ln \bar{K}^\ddagger \quad (A - 88).$$

Entonces el coeficiente de velocidad se escribe en la forma:

$$k_2 = \kappa (k_B T / h) \exp(-\Delta G_m^\ddagger / RT) \quad (A - 89),$$

y también se pueden definir una entropía y una entalpía de activación, al usar la ecuación:

$$\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger \quad (A - 90).$$

La introducción de (A-90) en (A-89) lleva a que:

$$k_2 = \kappa (k_B T / h) \exp(\Delta S^\ddagger / R) \exp(-\Delta H^\ddagger / RT) \quad (A - 91),$$

que es formalmente idéntica a la ecuación de Arrhenius, salvo por el hecho de que  $\Delta H^\ddagger \neq E_a$ ; sin embargo, puede concluirse que:

$$A \sim \kappa (k_B T / h) \exp(\Delta S^\ddagger / R) \quad (A - 92),$$

y se logra la conexión con la teoría de colisiones.

#### A.6 Referencias del apéndice A

- [1].- J. de Oyarzábal, *Mecánica Clásica*, Facultad de Ciencias, UNAM, México (1980)
- [2].- Ll. Motz y J. H. Weaver, *The Story of Physics*, Plenum Press, Nueva York y Londres (1989)
- [3].- R. P. Feynman, R. B. Leighton y M. Sands, *Lecciones de Física*, vol. I, Fondo Educativo Interamericano (1972)
- [4].- J. Kestin y J. R. Dorfman, *A Course in Statistical Thermodynamics*, Academic Press, Nueva York (1971)
- [5].- K. Huang, *Statistical Mechanics*, Wiley and Sons, Nueva York (1963)
- [6].- P. W. Atkins, *Physical Chemistry*, Oxford University Press (1979)
- [7].- G. E. MacWood y F. H. Verhoek, *J. Chem. Educ.* **38**, 334 (1961)
- [8].- J. O'M. Bokris y A. K. N. Reddy, *Modern Electrochemistry*, vol. 1, Plenum/Rosetta, Nueva York (1973)
- [9].- I. N. Levine, *Fisicoquímica*, McGraw - Hill, México (1981)
- [10].- H. Saint-Martin, *Tesis de Licenciatura*, Facultad de Ciencias, UNAM, México (1983)
- [11].- R. Daudel, *Quantum Theory of Chemical Reactivity*, D. Reidel, Dordrecht (1973)
- [12].- Ya. Guerasimov, V. Dreving, E. Eriomin, A. Kiseliyov, V. Lebedev, G. Panchenkov y A. Schliguin, *Curso de Química Física*, vol. I, 2a. ed., Mir, Moscú (1977)
- [13].- J. Kraut, *Science* **242**, 533 (1988)

APÉNDICE B

---

SUPERFICIES DE INTERACCIÓN  
DEL AGUA CON LOS FOSFATOS

## APÉNDICE B

---

En la Tabla B - 1 se presentan las coordenadas de la molécula de ortofosfato  $H_3PO_4$ , y las de las moléculas de agua  $H_2O$ , que se usaron para calcular las energías de interacción por pares, en la superficie a la que se ajustaron los parámetros del potencial. Las moléculas de agua están numeradas de acuerdo al dímero en el que se usaron, y que se indica en las columnas 1 y 6 de la Tabla 5 - 1. Las contribuciones no aditivas de tres cuerpos, se calcularon con la molécula del ortofosfatos y con combinaciones de las moléculas de agua mencionadas. En la Tabla B - 2 se muestran las combinaciones hechas con esas moléculas, indicándolas con el número que les corresponde en la Tabla B - 1.

De forma equivalente, en la Tabla B - 3 se presentan las coordenadas de la molécula del pirofosfato  $H_4P_2O_7$ , y de las moléculas de agua  $H_2O$ , que se usaron para calcular las energías de interacción. Otra vez, la numeración corresponde a la configuración indicada en las columnas 1 y 6 de la Tabla 5 - 3. Finalmente, la Tabla B - 4 muestra las combinaciones de moléculas de agua con las que se formaron los trímeros, para el cálculo de las contribuciones no aditivas.

TABLA B - 1

Átomo	x	y	z	Átomo	x	y	z
$H_3PO_4$							
$O_o$	0.000000	0.000000	0.000000				
$P_o$	0.000000	0.000000	2.745017				
$O_o$	2.680408	0.000000	4.021773				
$O_o$	-1.340194	2.321302	4.021773				
$O_o$	-1.340194	-2.321302	4.021773				
$H_o$	4.006429	1.075500	3.384369				
$H_o$	-2.934632	2.931911	3.384369				
$H_o$	-1.071853	4.007411	3.384369				
$H_2O$							
molécula 1				molécula 2			
$O_w$	-5.217127	0.000000	-2.002668	$O_w$	-5.746391	0.000000	-2.205833
$H_w$	-3.528424	0.000000	-1.354435	$H_w$	-4.057687	0.000000	-1.557600
$H_w$	-6.268044	0.000000	-0.530424	$H_w$	-6.797308	0.000000	-0.733590
molécula 3				molécula 4			
$O_w$	-6.099233	0.000000	-2.341277	$O_w$	-6.628496	0.000000	-2.544442
$H_w$	-4.410530	0.000000	-1.693044	$H_w$	-4.939793	0.000000	-1.896209
$H_w$	-7.150150	0.000000	-0.869032	$H_w$	-7.679413	0.000000	-1.072199
molécula 5				molécula 6			
$O_w$	-7.510604	0.000000	-2.883050	$O_w$	-5.253907	4.116029	6.132400
$H_w$	-5.821899	0.000000	-2.234817	$H_w$	-3.965830	3.106921	5.361374
$H_w$	-8.561519	0.000000	-1.410808	$H_w$	-6.148842	4.817140	4.783660
molécula 7				molécula 8			
$O_w$	-5.657610	4.432297	6.374050	$O_w$	-6.061312	4.748568	6.615699
$H_w$	-4.369532	3.423189	5.603025	$H_w$	-4.773234	3.739458	5.844674
$H_w$	-6.600453	5.170943	4.896150	$H_w$	-6.927665	5.427288	5.344011
molécula 9				molécula 10			
$O_w$	-6.195883	4.853994	6.696237	$O_w$	-6.263167	4.906706	6.736513
$H_w$	-4.907804	3.844883	5.925218	$H_w$	-4.975087	3.897596	5.965493
$H_w$	-7.106387	5.567303	5.305508	$H_w$	-7.173666	5.620012	5.345797
molécula 11				molécula 12			
$O_w$	-6.397734	5.012128	6.817063	$O_w$	-6.465018	5.064840	6.857337
$H_w$	-5.109655	4.003018	6.046043	$H_w$	-5.176939	4.055730	6.086317
$H_w$	-7.308235	5.725436	5.426343	$H_w$	-7.375515	5.778144	5.466630
molécula 13				molécula 14			
$O_w$	-6.599585	5.170262	6.937886	$O_w$	-7.272423	5.697379	7.340636
$H_w$	-5.311506	4.161152	6.166866	$H_w$	-5.984341	4.688269	6.569616
$H_w$	-7.510086	5.883570	5.547166	$H_w$	-8.182919	6.410683	5.949928
molécula 15				molécula 16			
$O_w$	-7.541558	5.908225	7.501735	$O_w$	-5.253919	4.116037	6.132371
$H_w$	-6.253476	4.899115	6.730715	$H_w$	-3.965830	3.106921	5.361374
$H_w$	-8.452058	6.621533	6.111014	$H_w$	-6.164404	4.829333	4.741633

TABLA B - 1 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 17				molécula 18			
O <sub>w</sub>	-5.657621	4.432307	6.374020	O <sub>w</sub>	-5.926754	4.643153	6.535121
H <sub>w</sub>	-4.369532	3.423189	5.603025	H <sub>w</sub>	-4.638667	3.634035	5.764124
H <sub>w</sub>	-6.568106	5.145601	4.983283	H <sub>w</sub>	-6.837241	5.356448	5.144384
molécula 19				molécula 20			
O <sub>w</sub>	-6.195889	4.853998	6.696220	O <sub>w</sub>	-6.599580	5.170258	6.937899
H <sub>w</sub>	-4.907800	3.844882	5.925225	H <sub>w</sub>	-5.311502	4.161150	6.166874
H <sub>w</sub>	-7.106374	5.567294	5.305484	H <sub>w</sub>	-7.510086	5.883570	5.547183
molécula 21				molécula 22			
O <sub>w</sub>	-7.003282	5.486529	7.179550	O <sub>w</sub>	-7.272417	5.697375	7.340651
H <sub>w</sub>	-5.715205	4.477418	6.408525	H <sub>w</sub>	-5.984340	4.688265	6.569626
H <sub>w</sub>	-7.913786	6.199838	5.788834	H <sub>w</sub>	-8.182921	6.410685	5.949933
molécula 23				molécula 24			
O <sub>w</sub>	6.519652	2.015465	5.617942	O <sub>w</sub>	6.891461	2.036978	5.682267
H <sub>w</sub>	7.832523	1.024005	4.866077	H <sub>w</sub>	8.204332	1.045518	4.930401
H <sub>w</sub>	7.385149	3.132965	6.746676	H <sub>w</sub>	7.756958	3.154477	6.811001
molécula 25				molécula 26			
O <sub>w</sub>	7.263271	2.058491	5.746591	O <sub>w</sub>	7.820986	2.090760	5.843079
H <sub>w</sub>	8.576141	1.067030	4.994726	H <sub>w</sub>	9.133855	1.099299	5.091213
H <sub>w</sub>	8.128767	3.175990	6.875325	H <sub>w</sub>	8.686483	3.208259	6.971811
molécula 27				molécula 28			
O <sub>w</sub>	8.378701	2.123029	5.939566	O <sub>w</sub>	8.750510	2.144541	6.003891
H <sub>w</sub>	9.691570	1.131568	5.187699	H <sub>w</sub>	10.063379	1.153079	5.252023
H <sub>w</sub>	9.244196	3.240528	7.068298	H <sub>w</sub>	9.616005	3.262041	7.132623
molécula 29				molécula 30			
O <sub>w</sub>	-5.210355	0.000000	0.000000	O <sub>w</sub>	-5.588300	0.000000	0.000000
H <sub>w</sub>	-3.401508	0.000000	0.000000	H <sub>w</sub>	-3.779454	0.000000	0.000000
H <sub>w</sub>	-5.663865	0.000000	1.751072	H <sub>w</sub>	-6.041810	0.000000	1.751072
molécula 31				molécula 32			
O <sub>w</sub>	-6.155218	0.000000	0.000000	O <sub>w</sub>	-6.533163	0.000000	0.000000
H <sub>w</sub>	-4.346372	0.000000	0.000000	H <sub>w</sub>	-4.724317	0.000000	0.000000
H <sub>w</sub>	-6.608728	0.000000	1.751072	H <sub>w</sub>	-6.986673	0.000000	1.751072
molécula 33				molécula 34			
O <sub>w</sub>	-7.100081	0.000000	0.000000	O <sub>w</sub>	-7.478027	0.000000	0.000000
H <sub>w</sub>	-5.291235	0.000000	0.000000	H <sub>w</sub>	-5.669180	0.000000	0.000000
H <sub>w</sub>	-7.553591	0.000000	1.751072	H <sub>w</sub>	-7.931537	0.000000	1.751072
molécula 35				molécula 36			
O <sub>w</sub>	-7.855972	0.000000	0.000000	O <sub>w</sub>	5.287764	5.287764	0.000000
H <sub>w</sub>	-6.047126	0.000000	0.000000	H <sub>w</sub>	4.008715	4.008715	0.000000
H <sub>w</sub>	-8.309482	0.000000	1.751072	H <sub>w</sub>	5.608445	5.608445	1.751072

TABLA B - 1 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 37				molécula 38			
O <sub>w</sub>	3.755309	3.451064	7.284445	O <sub>w</sub>	3.871697	3.776906	7.733529
H <sub>w</sub>	2.590445	4.136876	8.486393	H <sub>w</sub>	2.706833	4.462720	8.935477
H <sub>w</sub>	5.374768	4.037949	7.836567	H <sub>w</sub>	5.491157	4.363793	8.285651
molécula 39				molécula 40			
O <sub>w</sub>	3.949287	3.994134	8.032919	O <sub>w</sub>	4.143268	4.537204	8.781392
H <sub>w</sub>	2.784425	4.679948	9.234868	H <sub>w</sub>	2.978406	5.223018	9.983342
H <sub>w</sub>	5.568749	4.581021	8.585040	H <sub>w</sub>	5.762727	5.124091	9.333516
molécula 41				molécula 42			
O <sub>w</sub>	4.337248	5.080273	9.529868	O <sub>w</sub>	4.414840	5.297501	9.829257
H <sub>w</sub>	3.172386	5.766087	10.731815	H <sub>w</sub>	3.249979	5.983315	11.031205
H <sub>w</sub>	5.956708	5.667160	10.081989	H <sub>w</sub>	6.034300	5.884388	10.381378
molécula 43				molécula 44			
O <sub>w</sub>	8.071543	2.230222	4.736413	O <sub>w</sub>	8.963044	2.476549	4.929583
H <sub>w</sub>	6.364853	1.758652	4.366609	H <sub>w</sub>	7.256354	2.004979	4.559778
H <sub>w</sub>	9.050737	1.449451	3.431232	H <sub>w</sub>	9.942238	1.695778	3.624401
molécula 45				molécula 46			
O <sub>w</sub>	9.497944	2.624347	5.045484	O <sub>w</sub>	9.854545	2.722879	5.122753
H <sub>w</sub>	7.791255	2.152777	4.675681	H <sub>w</sub>	8.147856	2.251307	4.752948
H <sub>w</sub>	10.477139	1.843576	3.740302	H <sub>w</sub>	10.833739	1.942106	3.817571
molécula 47				molécula 48			
O <sub>w</sub>	10.211146	2.821409	5.200020	O <sub>w</sub>	9.452738	-0.562882	0.453687
H <sub>w</sub>	8.504457	2.349839	4.830215	H <sub>w</sub>	7.843241	-0.239665	1.213275
H <sub>w</sub>	11.190340	2.040638	3.894838	H <sub>w</sub>	9.077824	-1.602409	-0.978357
molécula 49				molécula 50			
O <sub>w</sub>	9.789032	-0.630415	0.294977	O <sub>w</sub>	10.293472	-0.731715	0.056911
H <sub>w</sub>	8.179533	-0.307200	1.054564	H <sub>w</sub>	8.683973	-0.408498	0.816498
H <sub>w</sub>	9.414118	-1.669942	-1.137068	H <sub>w</sub>	9.918558	-1.771243	-1.375133
molécula 51				molécula 52			
O <sub>w</sub>	10.797910	-0.833016	-0.181155	O <sub>w</sub>	0.000011	0.000000	9.520646
H <sub>w</sub>	9.188413	-0.509799	0.578432	H <sub>w</sub>	0.000008	0.000000	7.711799
H <sub>w</sub>	10.422996	-1.872542	-1.613197	H <sub>w</sub>	-1.751061	0.000004	9.974159
molécula 53				molécula 54			
O <sub>w</sub>	0.000013	0.000000	10.087564	O <sub>w</sub>	0.000013	0.000000	10.465509
H <sub>w</sub>	0.000009	0.000000	8.278717	H <sub>w</sub>	0.000009	0.000000	8.656663
H <sub>w</sub>	-1.751059	0.000004	10.541077	H <sub>w</sub>	-1.751059	0.000004	10.919023
molécula 55				molécula 56			
O <sub>w</sub>	0.000013	0.000000	10.843454	O <sub>w</sub>	-0.389252	0.304938	8.353594
H <sub>w</sub>	0.000011	0.000000	9.034608	H <sub>w</sub>	-0.996196	0.780436	6.717304
H <sub>w</sub>	-1.751059	0.000004	11.296968	H <sub>w</sub>	-1.484014	1.162603	9.510241

TABLA B - 1 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 57				molécula 58			
O <sub>w</sub>	-0.199026	0.155910	8.866430	O <sub>w</sub>	-0.072210	0.056558	9.208321
H <sub>w</sub>	-0.805970	0.631409	7.230140	H <sub>w</sub>	-0.679155	0.532056	7.572031
H <sub>w</sub>	-1.293788	1.013574	10.023077	H <sub>w</sub>	-1.166972	0.914223	10.364968
molécula 59				molécula 60			
O <sub>w</sub>	0.244833	-0.191822	10.063048	O <sub>w</sub>	0.435057	-0.340850	10.575886
H <sub>w</sub>	-0.362113	0.283676	8.426758	H <sub>w</sub>	-0.171888	0.134649	8.939594
H <sub>w</sub>	-0.849931	0.665843	11.219695	H <sub>w</sub>	-0.659706	0.516816	11.732531
molécula 61				molécula 62			
O <sub>w</sub>	0.561874	-0.440201	10.917776	O <sub>w</sub>	0.688690	-0.539553	11.259667
H <sub>w</sub>	-0.045070	0.035296	9.281486	H <sub>w</sub>	0.081746	-0.064056	9.623377
H <sub>w</sub>	-0.532888	0.417463	12.074423	H <sub>w</sub>	-0.406072	0.318111	12.416314
molécula 63				molécula 64			
O <sub>w</sub>	-3.183710	2.494182	8.522864	O <sub>w</sub>	-3.334659	2.612440	9.056368
H <sub>w</sub>	-2.702075	2.116862	6.820636	H <sub>w</sub>	-2.853026	2.235120	7.354138
H <sub>w</sub>	-4.601645	3.605032	8.357351	H <sub>w</sub>	-4.752597	3.723289	8.890855
molécula 65				molécula 66			
O <sub>w</sub>	-3.435295	2.691277	9.412036	O <sub>w</sub>	-3.686878	2.888374	10.301207
H <sub>w</sub>	-2.953660	2.313957	7.709807	H <sub>w</sub>	-3.205245	2.511054	8.598977
H <sub>w</sub>	-4.853230	3.802127	9.246522	H <sub>w</sub>	-5.104815	3.999223	10.135693
molécula 67				molécula 68			
O <sub>w</sub>	-3.837829	3.006631	10.834709	O <sub>w</sub>	-3.938463	3.085468	11.190378
H <sub>w</sub>	-3.356196	2.629311	9.132481	H <sub>w</sub>	-3.456830	2.708149	9.488148
H <sub>w</sub>	-5.255767	4.117480	10.669197	H <sub>w</sub>	-5.356400	4.196318	11.024864
molécula 69				molécula 70			
O <sub>w</sub>	-4.039096	3.164306	11.546047	O <sub>w</sub>	0.000000	4.619645	7.607127
H <sub>w</sub>	-3.557464	2.786986	9.843817	H <sub>w</sub>	0.000000	3.340597	6.328079
H <sub>w</sub>	-5.457034	4.275155	11.380534	H <sub>w</sub>	-0.819176	3.845974	9.022157
molécula 71				molécula 72			
O <sub>w</sub>	-5.257675	4.118981	6.122206	O <sub>w</sub>	-5.392352	4.224490	6.202457
H <sub>w</sub>	-3.967487	3.108219	5.356897	H <sub>w</sub>	-4.102164	3.213728	5.437148
H <sub>w</sub>	-6.164353	4.829293	4.727460	H <sub>w</sub>	-6.299032	4.934802	4.807711
molécula 73				molécula 74			
O <sub>w</sub>	-5.527031	4.330001	6.282709	O <sub>w</sub>	-5.661708	4.435510	6.362961
H <sub>w</sub>	-4.236843	3.319237	5.517399	H <sub>w</sub>	-4.371520	3.424746	5.597650
H <sub>w</sub>	-6.433709	5.040311	4.887964	H <sub>w</sub>	-6.568386	5.145822	4.968215
molécula 75				molécula 76			
O <sub>w</sub>	-5.796387	4.541019	6.443212	O <sub>w</sub>	-5.931064	4.646528	6.523463
H <sub>w</sub>	-4.506197	3.530257	5.677903	H <sub>w</sub>	-4.640876	3.635766	5.758154
H <sub>w</sub>	-6.703065	5.251332	5.048466	H <sub>w</sub>	-6.837742	5.356841	5.128717

TABLA B - 1 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 77				molécula 78			
O <sub>w</sub>	-6.065741	4.752037	6.603715	O <sub>w</sub>	-6.200420	4.857548	6.683967
H <sub>w</sub>	-4.775553	3.741275	5.838405	H <sub>w</sub>	-4.910230	3.846785	5.918658
H <sub>w</sub>	-6.972419	5.462350	5.208970	H <sub>w</sub>	-7.107098	5.567859	5.289221
molécula 79				molécula 80			
O <sub>w</sub>	-6.335097	4.963057	6.764218	O <sub>w</sub>	-6.469774	5.068567	6.844469
H <sub>w</sub>	-5.044909	3.952296	5.998909	H <sub>w</sub>	-5.179586	4.057805	6.079160
H <sub>w</sub>	-7.241775	5.673370	5.369472	H <sub>w</sub>	-7.376454	5.778879	5.449724
molécula 81				molécula 82			
O <sub>w</sub>	-6.739130	5.279587	7.004973	O <sub>w</sub>	-6.873807	5.385096	7.085224
H <sub>w</sub>	-5.448942	4.268823	6.239664	H <sub>w</sub>	-5.583619	4.374332	6.319915
H <sub>w</sub>	-7.645808	5.989897	5.610227	H <sub>w</sub>	-7.780487	6.095408	5.690479
molécula 83				molécula 84			
O <sub>w</sub>	-7.008486	5.490605	7.165477	O <sub>w</sub>	-7.143163	5.596114	7.245728
H <sub>w</sub>	-5.718298	4.479843	6.400168	H <sub>w</sub>	-5.852975	4.585352	6.480419
H <sub>w</sub>	-7.915164	6.200917	5.770730	H <sub>w</sub>	-8.049841	6.306426	5.850982
molécula 85				molécula 86			
O <sub>w</sub>	-7.277841	5.701623	7.325979	O <sub>w</sub>	-5.257675	4.118981	6.122206
H <sub>w</sub>	-5.987652	4.690861	6.560670	H <sub>w</sub>	-3.967487	3.108219	5.356897
H <sub>w</sub>	-8.184520	6.411936	5.931234	H <sub>w</sub>	-4.999830	3.913097	7.900704
molécula 87				molécula 88			
O <sub>w</sub>	-5.392352	4.224490	6.202457	O <sub>w</sub>	-5.527031	4.330001	6.282709
H <sub>w</sub>	-4.102164	3.213728	5.437148	H <sub>w</sub>	-4.236843	3.319237	5.517399
H <sub>w</sub>	-5.134507	4.018606	7.980955	H <sub>w</sub>	-5.269186	4.124115	8.061206
molécula 89				molécula 90			
O <sub>w</sub>	-5.661708	4.435510	6.362961	O <sub>w</sub>	-5.796387	4.541019	6.443212
H <sub>w</sub>	-4.371520	3.424746	5.597650	H <sub>w</sub>	-4.506197	3.530257	5.677903
H <sub>w</sub>	-5.403863	4.229624	8.141457	H <sub>w</sub>	-5.538540	4.335135	8.221710
molécula 91				molécula 92			
O <sub>w</sub>	-5.931064	4.646528	6.523463	O <sub>w</sub>	-6.065741	4.752037	6.603715
H <sub>w</sub>	-4.640876	3.635766	5.758154	H <sub>w</sub>	-4.775553	3.741275	5.838405
H <sub>w</sub>	-5.673219	4.440644	8.301961	H <sub>w</sub>	-5.807896	4.546154	8.382212
molécula 93				molécula 94			
O <sub>w</sub>	-6.200420	4.857548	6.683967	O <sub>w</sub>	-6.335097	4.963057	6.764218
H <sub>w</sub>	-4.910230	3.846785	5.918658	H <sub>w</sub>	-5.044909	3.952296	5.998909
H <sub>w</sub>	-5.942575	4.651663	8.462465	H <sub>w</sub>	-6.077252	4.757174	8.542716
molécula 95				molécula 96			
O <sub>w</sub>	-6.469774	5.068567	6.844469	O <sub>w</sub>	-6.604453	5.174076	6.924722
H <sub>w</sub>	-5.179586	4.057805	6.079160	H <sub>w</sub>	-5.314265	4.163314	6.159413
H <sub>w</sub>	-6.211929	4.862683	8.622967	H <sub>w</sub>	-6.346608	4.968192	8.703220

**Tabla B-1.** Coordenadas de la molécula de ortofosftao  $H_3PO_4$ , y de las moléculas de agua  $H_2O$ , que se usaron para el cálculo de las energías de interacción.

**TABLA B - 2**

<b>Trímero</b>	<b>molécula 1</b>	<b>molécula 2</b>
1	33	90
2	33	91
3	33	92
4	33	93
5	33	94
6	33	95
7	33	96
8	34	97
9	60	92
10	60	93
11	60	94
12	60	95
13	60	96
14	60	28
15	60	31
16	60	34
17	60	37
18	60	40
19	60	46
20	34	91
21	34	93
22	34	94
23	34	95

**Tabla B-2.** Trímeros usados en el cálculo de las contribuciones no aditivas a la energía de interacción del ortofosfato con el agua.

TABLA B - 3

Átomo	x	y	z	Átomo	x	y	z
<i>H<sub>4</sub>P<sub>2</sub>O<sub>7</sub></i>							
<i>O<sub>p</sub></i>	0.000000	0.000000	0.000000				
<i>P<sub>p</sub></i>	0.000000	0.000000	2.987482				
<i>P<sub>p</sub></i>	1.961504	0.000000	-2.253342				
<i>O<sub>p</sub></i>	-1.947322	1.525575	4.153122				
<i>O<sub>p</sub></i>	2.690686	2.473286	-3.164005				
<i>O<sub>p</sub></i>	2.798848	0.773339	3.593930				
<i>O<sub>p</sub></i>	0.650601	-1.815846	-4.198493				
<i>O<sub>p</sub></i>	-0.077216	-2.902694	3.593930				
<i>O<sub>p</sub></i>	4.159572	-1.654446	-1.143982				
<i>H<sub>p</sub></i>	3.173367	1.821853	5.039023				
<i>H<sub>p</sub></i>	1.317623	-1.995132	-5.886949				
<i>H<sub>p</sub></i>	-1.601264	-3.669613	4.239729				
<i>H<sub>p</sub></i>	5.381400	-0.920904	-0.005125				
<i>H<sub>2</sub>O</i>							
molécula 1				molécula 2			
<i>O<sub>w</sub></i>	-5.217127	0.000000	-2.002668	<i>O<sub>w</sub></i>	-5.746391	0.000000	-2.205833
<i>H<sub>w</sub></i>	-3.528424	0.000000	-1.354435	<i>H<sub>w</sub></i>	-4.057687	0.000000	-1.557600
<i>H<sub>w</sub></i>	-6.268044	0.000000	-0.530424	<i>H<sub>w</sub></i>	-6.797308	0.000000	-0.733590
molécula 3				molécula 4			
<i>O<sub>w</sub></i>	-6.099233	0.000000	-2.341277	<i>O<sub>w</sub></i>	-6.628496	0.000000	-2.544442
<i>H<sub>w</sub></i>	-4.410530	0.000000	-1.693044	<i>H<sub>w</sub></i>	-4.939793	0.000000	-1.896209
<i>H<sub>w</sub></i>	-7.150150	0.000000	-0.869032	<i>H<sub>w</sub></i>	-7.679413	0.000000	-1.072199
molécula 5				molécula 6			
<i>O<sub>w</sub></i>	-7.510604	0.000000	-2.883050	<i>O<sub>w</sub></i>	-5.253907	4.116029	6.132400
<i>H<sub>w</sub></i>	-5.821899	0.000000	-2.234817	<i>H<sub>w</sub></i>	-3.965830	3.106921	5.361374
<i>H<sub>w</sub></i>	-8.561519	0.000000	-1.410808	<i>H<sub>w</sub></i>	-6.148842	4.817140	4.783660
molécula 7				molécula 8			
<i>O<sub>w</sub></i>	-5.657610	4.432297	6.374050	<i>O<sub>w</sub></i>	-6.061312	4.748568	6.615699
<i>H<sub>w</sub></i>	-4.369532	3.423189	5.603025	<i>H<sub>w</sub></i>	-4.773234	3.739458	5.844674
<i>H<sub>w</sub></i>	-6.600453	5.170943	4.896150	<i>H<sub>w</sub></i>	-6.927665	5.427288	5.344011
molécula 9				molécula 10			
<i>O<sub>w</sub></i>	-6.195883	4.853994	6.696237	<i>O<sub>w</sub></i>	-6.263167	4.906706	6.736513
<i>H<sub>w</sub></i>	-4.907804	3.844883	5.925218	<i>H<sub>w</sub></i>	-4.975087	3.897596	5.965493
<i>H<sub>w</sub></i>	-7.106387	5.567303	5.305508	<i>H<sub>w</sub></i>	-7.173666	5.620012	5.345797
molécula 11				molécula 12			
<i>O<sub>w</sub></i>	-6.397734	5.012128	6.817063	<i>O<sub>w</sub></i>	-6.465018	5.064840	6.857337
<i>H<sub>w</sub></i>	-5.109655	4.003018	6.046043	<i>H<sub>w</sub></i>	-5.176939	4.055730	6.086317
<i>H<sub>w</sub></i>	-7.308235	5.725436	5.426343	<i>H<sub>w</sub></i>	-7.375515	5.778144	5.466630

TABLA B - 3 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 13				molécula 14			
O <sub>w</sub>	-6.599585	5.170262	6.937886	O <sub>w</sub>	-7.272423	5.697379	7.340636
H <sub>w</sub>	-5.311506	4.161152	6.166866	H <sub>w</sub>	-5.984341	4.688269	6.569616
H <sub>w</sub>	-7.510086	5.883570	5.547166	H <sub>w</sub>	-8.182919	6.410683	5.949928
molécula 15				molécula 16			
O <sub>w</sub>	-7.541558	5.908225	7.501735	O <sub>w</sub>	-5.253919	4.116037	6.132371
H <sub>w</sub>	-6.253476	4.899115	6.730715	H <sub>w</sub>	-3.965830	3.106921	5.361374
H <sub>w</sub>	-8.452058	6.621533	6.111014	H <sub>w</sub>	-6.164404	4.829333	4.741633
molécula 17				molécula 18			
O <sub>w</sub>	-5.657621	4.432307	6.374020	O <sub>w</sub>	-5.926754	4.643153	6.535121
H <sub>w</sub>	-4.369532	3.423189	5.603025	H <sub>w</sub>	-4.638667	3.634035	5.764124
H <sub>w</sub>	-6.568106	5.145601	4.983283	H <sub>w</sub>	-6.837241	5.356448	5.144384
molécula 19				molécula 20			
O <sub>w</sub>	-6.195889	4.853998	6.696220	O <sub>w</sub>	-6.599580	5.170258	6.937899
H <sub>w</sub>	-4.907800	3.844882	5.925225	H <sub>w</sub>	-5.311502	4.161150	6.166874
H <sub>w</sub>	-7.106374	5.567294	5.305484	H <sub>w</sub>	-7.510086	5.883570	5.547183
molécula 21				molécula 22			
O <sub>w</sub>	-7.003282	5.486529	7.179550	O <sub>w</sub>	-7.272417	5.697375	7.340651
H <sub>w</sub>	-5.715205	4.477418	6.408525	H <sub>w</sub>	-5.984340	4.688265	6.569626
H <sub>w</sub>	-7.913786	6.199838	5.788834	H <sub>w</sub>	-8.182921	6.410685	5.949933
molécula 23				molécula 24			
O <sub>w</sub>	6.519652	2.015465	5.617942	O <sub>w</sub>	6.891461	2.036978	5.682267
H <sub>w</sub>	7.832523	1.024005	4.866077	H <sub>w</sub>	8.204332	1.045518	4.930401
H <sub>w</sub>	7.385149	3.132965	6.746676	H <sub>w</sub>	7.756958	3.154477	6.811001
molécula 25				molécula 26			
O <sub>w</sub>	7.263271	2.058491	5.746591	O <sub>w</sub>	7.820986	2.090760	5.843079
H <sub>w</sub>	8.576141	1.067030	4.994726	H <sub>w</sub>	9.133855	1.099299	5.091213
H <sub>w</sub>	8.128767	3.175990	6.875325	H <sub>w</sub>	8.686483	3.208259	6.971811
molécula 27				molécula 28			
O <sub>w</sub>	8.378701	2.123029	5.939566	O <sub>w</sub>	8.750510	2.144541	6.003891
H <sub>w</sub>	9.691570	1.131568	5.187699	H <sub>w</sub>	10.063379	1.153079	5.252023
H <sub>w</sub>	9.244196	3.240528	7.068298	H <sub>w</sub>	9.616005	3.262041	7.132623
molécula 29				molécula 30			
O <sub>w</sub>	-5.210355	0.000000	0.000000	O <sub>w</sub>	-5.588300	0.000000	0.000000
H <sub>w</sub>	-3.401508	0.000000	0.000000	H <sub>w</sub>	-3.779454	0.000000	0.000000
H <sub>w</sub>	-5.663865	0.000000	1.751072	H <sub>w</sub>	-6.041810	0.000000	1.751072
molécula 31				molécula 32			
O <sub>w</sub>	-6.155218	0.000000	0.000000	O <sub>w</sub>	-6.533163	0.000000	0.000000
H <sub>w</sub>	-4.346372	0.000000	0.000000	H <sub>w</sub>	-4.724317	0.000000	0.000000
H <sub>w</sub>	-6.608728	0.000000	1.751072	H <sub>w</sub>	-6.986673	0.000000	1.751072

TABLA B - 3 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 33				molécula 34			
O <sub>w</sub>	-7.100081	0.000000	0.000000	O <sub>w</sub>	-7.478027	0.000000	0.000000
H <sub>w</sub>	-5.291235	0.000000	0.000000	H <sub>w</sub>	-5.669180	0.000000	0.000000
H <sub>w</sub>	-7.553591	0.000000	1.751072	H <sub>w</sub>	-7.931537	0.000000	1.751072
molécula 35				molécula 36			
O <sub>w</sub>	-7.855972	0.000000	0.000000	O <sub>w</sub>	5.287764	5.287764	0.000000
H <sub>w</sub>	-6.047126	0.000000	0.000000	H <sub>w</sub>	4.008715	4.008715	0.000000
H <sub>w</sub>	-8.309482	0.000000	1.751072	H <sub>w</sub>	5.608445	5.608445	1.751072
molécula 37				molécula 38			
O <sub>w</sub>	3.755309	3.451064	7.284445	O <sub>w</sub>	3.871697	3.776906	7.733529
H <sub>w</sub>	2.590445	4.136876	8.486393	H <sub>w</sub>	2.706833	4.462720	8.935477
H <sub>w</sub>	5.374768	4.037949	7.836567	H <sub>w</sub>	5.491157	4.363793	8.285651
molécula 39				molécula 40			
O <sub>w</sub>	3.949287	3.994134	8.032919	O <sub>w</sub>	4.143268	4.537204	8.781392
H <sub>w</sub>	2.784425	4.679948	9.234868	H <sub>w</sub>	2.978406	5.223018	9.983342
H <sub>w</sub>	5.568749	4.581021	8.585040	H <sub>w</sub>	5.762727	5.124091	9.333516
molécula 41				molécula 42			
O <sub>w</sub>	4.337248	5.080273	9.529868	O <sub>w</sub>	4.414840	5.297501	9.829257
H <sub>w</sub>	3.172386	5.766087	10.731815	H <sub>w</sub>	3.249979	5.983315	11.031205
H <sub>w</sub>	5.956708	5.667160	10.081989	H <sub>w</sub>	6.034300	5.884388	10.381378
molécula 43				molécula 44			
O <sub>w</sub>	8.071543	2.230222	4.736413	O <sub>w</sub>	8.963044	2.476549	4.929583
H <sub>w</sub>	6.364853	1.758652	4.366609	H <sub>w</sub>	7.256354	2.004979	4.559778
H <sub>w</sub>	9.050737	1.449451	3.431232	H <sub>w</sub>	9.942238	1.695778	3.624401
molécula 45				molécula 46			
O <sub>w</sub>	9.497944	2.624347	5.045484	O <sub>w</sub>	9.854545	2.722879	5.122753
H <sub>w</sub>	7.791255	2.152777	4.675681	H <sub>w</sub>	8.147856	2.251307	4.752948
H <sub>w</sub>	10.477139	1.843576	3.740302	H <sub>w</sub>	10.833739	1.942106	3.817571
molécula 47				molécula 48			
O <sub>w</sub>	10.211146	2.821409	5.200020	O <sub>w</sub>	9.452738	-0.562882	0.453687
H <sub>w</sub>	8.504457	2.349839	4.830215	H <sub>w</sub>	7.843241	-0.239665	1.213275
H <sub>w</sub>	11.190340	2.040638	3.894838	H <sub>w</sub>	9.077824	-1.602409	-0.978357
molécula 49				molécula 50			
O <sub>w</sub>	9.789032	-0.630415	0.294977	O <sub>w</sub>	10.293472	-0.731715	0.056911
H <sub>w</sub>	8.179533	-0.307200	1.054564	H <sub>w</sub>	8.683973	-0.408498	0.816498
H <sub>w</sub>	9.414118	-1.669942	-1.137068	H <sub>w</sub>	9.918558	-1.771243	-1.375133
molécula 51				molécula 52			
O <sub>w</sub>	10.797910	-0.833016	-0.181155	O <sub>w</sub>	0.000011	0.000000	9.520646
H <sub>w</sub>	9.188413	-0.509799	0.578432	H <sub>w</sub>	0.000008	0.000000	7.711799
H <sub>w</sub>	10.422996	-1.872542	-1.613197	H <sub>w</sub>	-1.751061	0.000004	9.974159

TABLA B - 3 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 53				molécula 54			
$O_w$	0.000013	0.000000	10.087564	$O_w$	0.000013	0.000000	10.465509
$H_w$	0.000009	0.000000	8.278717	$H_w$	0.000009	0.000000	8.656663
$H_w$	-1.751059	0.000004	10.541077	$H_w$	-1.751059	0.000004	10.919023
molécula 55				molécula 56			
$O_w$	0.000013	0.000000	10.843454	$O_w$	-0.389252	0.304938	8.353594
$H_w$	0.000011	0.000000	9.034608	$H_w$	-0.996196	0.780436	6.717304
$H_w$	-1.751059	0.000004	11.296968	$H_w$	-1.484014	1.162603	9.510241
molécula 57				molécula 58			
$O_w$	-0.199026	0.155910	8.866430	$O_w$	-0.072210	0.056558	9.208321
$H_w$	-0.805970	0.631409	7.230140	$H_w$	-0.679155	0.532056	7.572031
$H_w$	-1.293788	1.013574	10.023077	$H_w$	-1.166972	0.914223	10.364968
molécula 59				molécula 60			
$O_w$	0.244833	-0.191822	10.063048	$O_w$	0.435057	-0.340850	10.575886
$H_w$	-0.362113	0.283676	8.426758	$H_w$	-0.171888	0.134649	8.939594
$H_w$	-0.849931	0.665843	11.219695	$H_w$	-0.659706	0.516816	11.732531
molécula 61				molécula 62			
$O_w$	0.561874	-0.440201	10.917776	$O_w$	0.688690	-0.539553	11.259667
$H_w$	-0.045070	0.035296	9.281486	$H_w$	0.081746	-0.064056	9.623377
$H_w$	-0.532888	0.417463	12.074423	$H_w$	-0.406072	0.318111	12.416314
molécula 63				molécula 64			
$O_w$	-3.183710	2.494182	8.522864	$O_w$	-3.334659	2.612440	9.056368
$H_w$	-2.702075	2.116862	6.820636	$H_w$	-2.853026	2.235120	7.354138
$H_w$	-4.601645	3.605032	8.357351	$H_w$	-4.752597	3.723289	8.890855
molécula 65				molécula 66			
$O_w$	-3.435295	2.691277	9.412036	$O_w$	-3.686878	2.888374	10.301207
$H_w$	-2.953660	2.313957	7.709807	$H_w$	-3.205245	2.511054	8.598977
$H_w$	-4.853230	3.802127	9.246522	$H_w$	-5.104815	3.999223	10.135693
molécula 67				molécula 68			
$O_w$	-3.837829	3.006631	10.834709	$O_w$	-3.938463	3.085468	11.190378
$H_w$	-3.356196	2.629311	9.132481	$H_w$	-3.456830	2.708149	9.488148
$H_w$	-5.255767	4.117480	10.669197	$H_w$	-5.356400	4.196318	11.024864
molécula 69				molécula 70			
$O_w$	-4.039096	3.164306	11.546047	$O_w$	0.000000	4.619645	7.607127
$H_w$	-3.557464	2.786986	9.843817	$H_w$	0.000000	3.340597	6.328079
$H_w$	-5.457034	4.275155	11.380534	$H_w$	-0.819176	3.845974	9.022157
molécula 71				molécula 72			
$O_w$	-5.257675	4.118981	6.122206	$O_w$	-5.392352	4.224490	6.202457
$H_w$	-3.967487	3.108219	5.356897	$H_w$	-4.102164	3.213728	5.437148
$H_w$	-6.164353	4.829293	4.727460	$H_w$	-6.299032	4.934802	4.807711

TABLA B - 3 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 73				molécula 74			
O <sub>w</sub>	-5.527031	4.330001	6.282709	O <sub>w</sub>	-5.661708	4.435510	6.362961
H <sub>w</sub>	-4.236843	3.319237	5.517399	H <sub>w</sub>	-4.371520	3.424746	5.597650
H <sub>w</sub>	-6.433709	5.040311	4.887964	H <sub>w</sub>	-6.568386	5.145822	4.968215
molécula 75				molécula 76			
O <sub>w</sub>	-5.796387	4.541019	6.443212	O <sub>w</sub>	-5.931064	4.646528	6.523463
H <sub>w</sub>	-4.506197	3.530257	5.677903	H <sub>w</sub>	-4.640876	3.635766	5.758154
H <sub>w</sub>	-6.703065	5.251332	5.048466	H <sub>w</sub>	-6.837742	5.356841	5.128717
molécula 77				molécula 78			
O <sub>w</sub>	-6.065741	4.752037	6.603715	O <sub>w</sub>	-6.200420	4.857548	6.683967
H <sub>w</sub>	-4.775553	3.741275	5.838405	H <sub>w</sub>	-4.910230	3.846785	5.918658
H <sub>w</sub>	-6.972419	5.462350	5.208970	H <sub>w</sub>	-7.107098	5.567859	5.289221
molécula 79				molécula 80			
O <sub>w</sub>	-6.335097	4.963057	6.764218	O <sub>w</sub>	-6.469774	5.068567	6.844469
H <sub>w</sub>	-5.044909	3.952296	5.998909	H <sub>w</sub>	-5.179586	4.057805	6.079160
H <sub>w</sub>	-7.241775	5.673370	5.369472	H <sub>w</sub>	-7.376454	5.778879	5.449724
molécula 81				molécula 82			
O <sub>w</sub>	-6.739130	5.279587	7.004973	O <sub>w</sub>	-6.873807	5.385096	7.085224
H <sub>w</sub>	-5.448942	4.268823	6.239664	H <sub>w</sub>	-5.583619	4.374332	6.319915
H <sub>w</sub>	-7.645808	5.989897	5.610227	H <sub>w</sub>	-7.780487	6.095408	5.690479
molécula 83				molécula 84			
O <sub>w</sub>	-7.008486	5.490605	7.165477	O <sub>w</sub>	-7.143163	5.596114	7.245728
H <sub>w</sub>	-5.718298	4.479843	6.400168	H <sub>w</sub>	-5.852975	4.585352	6.480419
H <sub>w</sub>	-7.915164	6.200917	5.770730	H <sub>w</sub>	-8.049841	6.306426	5.850982
molécula 85				molécula 86			
O <sub>w</sub>	-7.277841	5.701623	7.325979	O <sub>w</sub>	-5.257675	4.118981	6.122206
H <sub>w</sub>	-5.987652	4.690861	6.560670	H <sub>w</sub>	-3.967487	3.108219	5.356897
H <sub>w</sub>	-8.184520	6.411936	5.931234	H <sub>w</sub>	-4.999830	3.913097	7.900704
molécula 87				molécula 88			
O <sub>w</sub>	-5.392352	4.224490	6.202457	O <sub>w</sub>	-5.527031	4.330001	6.282709
H <sub>w</sub>	-4.102164	3.213728	5.437148	H <sub>w</sub>	-4.236843	3.319237	5.517399
H <sub>w</sub>	-5.134507	4.018606	7.980955	H <sub>w</sub>	-5.269186	4.124115	8.061206
molécula 89				molécula 90			
O <sub>w</sub>	-5.661708	4.435510	6.362961	O <sub>w</sub>	-5.796387	4.541019	6.443212
H <sub>w</sub>	-4.371520	3.424746	5.597650	H <sub>w</sub>	-4.506197	3.530257	5.677903
H <sub>w</sub>	-5.403863	4.229624	8.141457	H <sub>w</sub>	-5.538540	4.335135	8.221710
molécula 91				molécula 92			
O <sub>w</sub>	-5.931064	4.646528	6.523463	O <sub>w</sub>	-6.065741	4.752037	6.603715
H <sub>w</sub>	-4.640876	3.635766	5.758154	H <sub>w</sub>	-4.775553	3.741275	5.838405
H <sub>w</sub>	-5.673219	4.440644	8.301961	H <sub>w</sub>	-5.807896	4.546154	8.382212

TABLA B - 3 (continuación)

Átomo	x	y	z	Átomo	x	y	z
molécula 93				molécula 94			
$O_w$	-6.200420	4.857548	6.683967	$O_w$	-6.335097	4.963057	6.764218
$H_w$	-4.910230	3.846785	5.918658	$H_w$	-5.044909	3.952296	5.998909
$H_w$	-5.942575	4.651663	8.462465	$H_w$	-6.077252	4.757174	8.542716
molécula 95				molécula 96			
$O_w$	-6.469774	5.068567	6.844469	$O_w$	-6.604453	5.174076	6.924722
$H_w$	-5.179586	4.057805	6.079160	$H_w$	-5.314265	4.163314	6.159413
$H_w$	-6.211929	4.862683	8.622967	$H_w$	-6.346608	4.968192	8.703220
molécula 97				molécula 98			
$O_w$	-6.739130	5.279587	7.004973	$O_w$	-6.873807	5.385096	7.085224
$H_w$	-5.448942	4.268823	6.239664	$H_w$	-5.583619	4.374332	6.319915
$H_w$	-6.481285	5.073701	8.783471	$H_w$	-6.615962	5.179210	8.863722
molécula 99				molécula 100			
$O_w$	-7.008486	5.490605	7.165477	$O_w$	-7.143163	5.596114	7.245728
$H_w$	-5.718298	4.479843	6.400168	$H_w$	-5.852975	4.585352	6.480419
$H_w$	-6.750641	5.284721	8.943975	$H_w$	-6.885318	5.390230	9.024226

Tabla B-3. Coordenadas de la molécula de pirofosfato  $H_4P_2O_7$ , y de las moléculas de agua  $H_2O$ , que se usaron para el cálculo de las energías de interacción.

**TABLA B - 4**

Trímero	molécula 1	molécula 2
1	77	38
2	77	39
3	77	40
4	77	41
5	77	42
6	77	43
7	77	58
8	77	59
9	77	61
10	77	62
11	77	63
12	77	68
13	77	69
14	77	70
15	59	68
16	59	69
17	59	70
18	59	86
19	59	66
20	59	39
21	39	66

**Tabla B-4.** Trímeros usados en el cálculo de las contribuciones no aditivas a la energía de interacción del pirofosfato con el agua.

**APÉNDICE C**

---

**ARTÍCULOS RELACIONADOS  
CON ESTA TESIS**

140

# Nonadditivity in an analytical intermolecular potential: The water-water interaction

H. Saint-Martin, C. Medina-Llanos,<sup>1)</sup> and I. Ortega-Blake

Instituto de Física, UNAM, Laboratorio de Cuernavaca, Apartado postal 139-B, Cuernavaca, Morelos 62191, México

(Received 26 September 1988; accepted 13 July 1990)

We present an analytical potential for intermolecular interactions which includes nonadditivity. We fitted this analytical form to an *ab initio* energy surface for the water-water interaction and looked into the fidelity of the fitting to two-body interactions, as well as three-body and four-body nonadditives. We tested this potential in a Monte Carlo simulation of liquid water. The results obtained led us to conclude that the model performs very well, giving the best reported fit to the radial distribution functions, the enthalpy per water molecule, and other parameters. The potential appears as an inexpensive, accurate, and flexible potential to be used in numerical simulation studies.

## INTRODUCTION

The use of numerical simulations (Monte Carlo or molecular dynamics) has become a powerful tool for the study of molecular systems in many fields of physics, but realistic potentials have to be used when considering real systems. One approach has been to use analytical forms fitted to *ab initio* energy surfaces (see e.g., Ref. 1). The analytical form should be simple enough to be used later on in simulation studies, but also flexible enough to enable a good reproduction of a costly *ab initio* energy surface. Furthermore, it is now clear that nonadditive contributions play a crucial role in many systems (e.g., Refs. 2 and 3) and thus have to be included. The attempts to modify two-body potentials, so as to account for nonadditivity in an averaged way, have not yielded good results (see, e.g., the review of Egelstaff<sup>4</sup>). Here we propose a new model which gives a better reproduction of the two-body energy surface, it is simple enough for its use in numerical simulations and, most important, is able to include nonadditivity.

The model proposed is of an atom-atom type and is presented in Fig. 1. To each atom correspond a center  $I$  and a net charge  $\delta_I$ , with the following expression for the atom-atom potential:

$$V_{IJ} = A_{IJ} e^{-n_1 R_{IJ}} + B_{IJ} e^{-n_2 R_{IJ}} + C \frac{\delta_I \delta_J}{r_{IJ}} + a r_{IJ}^2 + a' r_{IJ}^2 + \frac{k}{r_{IJ}^{n_3}} \quad (1)$$

where  $R_{IJ}$  is the distance between centers  $I$  and  $J$ ,  $r_{IJ}$  the distance between the center  $I$  and the net charge  $\delta_J$ , and  $r_{IJ}$  the distance between charges  $\delta_I$  and  $\delta_J$ . The parameters  $A$ ,  $B$ ,  $n_1$ , and  $n_2$  depend on the pair of atoms interacting and correspond roughly to a Lennard-Jones type; they do not appear in intramolecular interactions so no vibrations are included. The parameter  $C$  is equal to one for all intermolecular atom-atom interactions and can have different values for pairs of interacting atoms which belong to the same

molecule. The parameters  $a_i$  and  $a_j$  depend only on the type of atom. The last term in Eq. (1) provides a repulsive interaction between positive and negative net charges to prevent their collapse, and also permits the inclusion, to some extent, of exchange energy. The  $k$  and  $n_3$  parameters depend on the pair of interacting atoms. Since this potential considers mobile charges and harmonic oscillator terms we will be referring to it as the MCHO potential.

For the fitting of the above potential to an *ab initio* surface, it should be noted that while the distances  $R_{IJ}$  are given by the surface itself (from the nuclear positions), the  $r_{IJ}$  distances depend on the electric field around the  $\delta_i$  and the  $\delta_j$  charges, so that an iterative procedure has to be used to solve the coupled third degree equations on  $x_i$ ,  $y_i$ , and  $z_i$ . With an initial guess for the parameters, the zero force equations for the coordinates of the  $\delta$  particles are solved, first through a direct iteration and then through the Newton-Raphson method, or the convergence accelerator of Shanks *et al.*<sup>5</sup> With these coordinates the set of  $r_{ij}$  is obtained and a set of parameters is adjusted. Finally, the new parameters are used for solving again the zero force equations.

We use the above model for reproducing the two-body water-water surface of Matsuoka, Clementi, and Yoshimine,<sup>6</sup> from which the well established MCY water potential was derived. The fitting was done by using a direct minimum squares procedure for the Lennard-Jones-type parameters, and a nonlinear optimization algorithm<sup>7</sup> for the others. The norm minimized was  $\sum |V_{12} - \epsilon_{12}|^2$ , where  $V_{12}$  is the *ab initio* interaction energy between molecules 1 and 2, and  $\epsilon_{12}$  the corresponding value obtained from the model. We obtained a mean standard deviation value of  $\sigma = 0.13$  kcal/mol for the fit, which is lower than the corresponding value for the original MCY potential ( $\sigma = 0.18$  kcal/mol). This shows the good ability of the model for reproducing the intermolecular forces.

The model has built in nonadditivity and we can adjust it to reproduce simultaneously the two-, three-, and many-

<sup>1)</sup>On leave from Institute of Quantum Chemistry, University of Uppsala, S-75120, Uppsala, Sweden

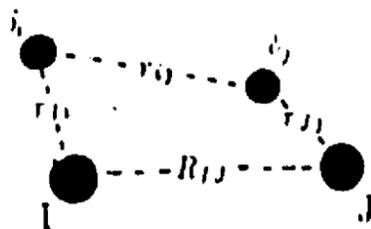


FIG. 1. Picture of the presented MCHO model.  $i$  and  $j$  correspond to the atomic nuclei and  $b_i$  and  $b_j$  to their respective mobile net charges.

body interactions. Unfortunately, there are no reported values of nonadditivity calculations with the basis set used in the MCY surface. For this reason we chose to adjust the model to the three-body nonadditivity of another *ab initio* surface, computed with an extended basis set by Clementi *et al.*,<sup>8</sup> who obtained the three-body nonadditivity using the many-body expansion

$$E_{123} = \sum_{i=1}^3 \sum_{j=i+1}^3 \epsilon_{ij} + \epsilon_{123}, \quad (2)$$

where  $\epsilon_{123}$  is the three-body nonadditivity. The reproduction of the three-body nonadditivity surface produced a value of  $\sigma = 0.22$  kcal/mol, without readjusting the parameters to fit it. This value compares rather well with that for the fitting of an induction-type model for this nonadditivity surface by Clementi *et al.*<sup>8</sup> ( $\sigma = 0.18$  kcal/mol), and with the fitting to a larger nonadditivity surface done by Clementi and Corongiu<sup>1</sup> ( $\sigma = 0.21$  kcal/mol). The fact that there is no need of readjusting parameters in our model is a very handsome feature, but there is no reason why this should not be done.

The model includes nonadditivity to more than three-body terms and it is important that higher terms, at least four-body, are reasonable. To check this, we computed some water tetramers using the same basis set as that of Clementi *et al.*<sup>8</sup> and made a many-body expansion

$$E_{1234} = \sum_{i=1}^4 \sum_{j=i+1}^4 \epsilon_{ij} + \sum_{i=1}^4 \sum_{j=i+1}^4 \sum_{k=j+1}^4 \epsilon_{ijk} + \epsilon_{1234}, \quad (3)$$

where  $\epsilon_{1234}$  is the four-body nonadditivity. In spite of a good reproduction of the four-body nonadditivity of these tetramers, their effect on the simulation seemed to be exaggerated. This encouraged us to look more deeply into the four-body nonadditivity in liquid water (see Ref. 9). As a result of this analysis, it was clear that a four-body surface of clusters corresponding to the structures that appear in the liquid was needed. Hence, simultaneous to the fit of the two-body and three-body surfaces mentioned, a fit to a four-body surface<sup>9</sup> was performed. The comparison between the model values and the corresponding *ab initio* values are presented in Figs. 2, 3, and 4. Note that the fitting to the two-body surface is still very good ( $\sigma = 0.15$  kcal/mol). The dipole and quadrupole moments of the isolated water molecule were included as extra weighted points in the surface to be fitted, giving a model value of  $\mu = 2.12$  D, very close to the *ab initio* value  $\mu = 2.26$  D.

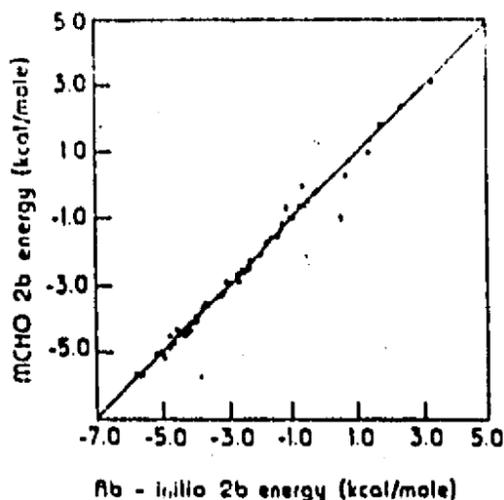


FIG. 2. Comparison between *ab initio* values from Matsuoka *et al.* (Ref. 6) and the corresponding values predicted by the MCHO model with the parameters of Table I. Values are in kcal/mol.

used as reference. The quadrupole components produced by the model are  $Q_{xx} = 2.65$ ,  $Q_{yy} = -2.64$ , and  $Q_{zz} = -0.01$  in units of  $10^{-26}$  esu cm<sup>2</sup>, which compare well with the experimental values<sup>10</sup> of  $Q_{xx} = 2.63$ ,  $Q_{yy} = -2.50$ , and  $Q_{zz} = -0.13$  in the same units. Contrary to what could be expected, the addition of the requirement to reproduce the dipole and quadrupole moments, far from making more difficult the fitting of the interaction, conduced to a much better one. The MCHO potential parameters corresponding to all these fittings are presented in Table I.

An immediate test for the potential is to look into the water dimer using the experimental geometry.<sup>11,12</sup> The interaction energy predicted by the model is  $-5.61$  kcal/mol, compared to the experimental value of  $-5.4 \pm 0.2$  kcal/mol. The discrepancy observed is the one to be expected from the molecular basis set used, following the

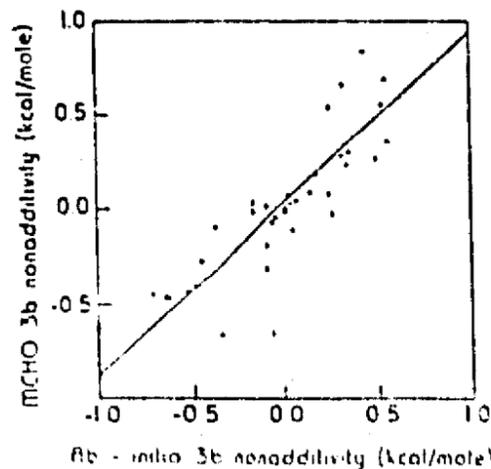


FIG. 3. Comparison between *ab initio* three-body nonaddivities from Clementi *et al.* (Ref. 8) and the corresponding values predicted by the MCHO model with the parameters of Table I. Values are in kcal/mol.

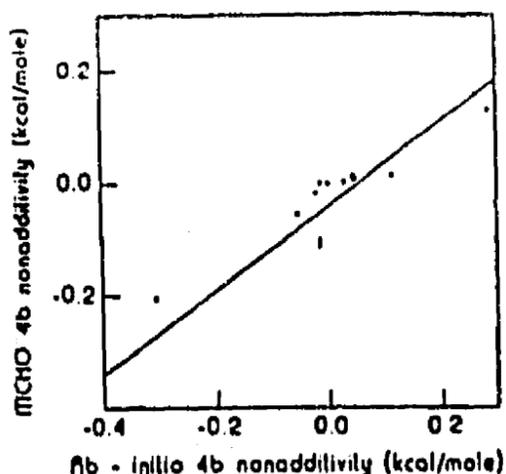


FIG. 4. Comparison between *ab initio* four-body nonadditivities from Gil-Adalid and Ortega-Blake (Ref. 9) and the corresponding values predicted by the MCHO model with the parameters of Table I. Values are in kcal/mol.

analysis made by Szalewicz *et al.*<sup>13</sup> Since the model permits the polarization of the water molecules it is important to see the resulting dipole moment of the dimer. A value of  $\mu = 3.23$  D against the experimental value<sup>11,12</sup> of  $\mu = 2.6$  D indicates that polarizability, which was only fitted through the energy interaction, is somewhat exaggerated, most certainly due to the quality of the basis set used in computing the two-body surface.

As a final and most important test, we performed a Monte Carlo simulation of liquid water, with 343 water molecules at a temperature of  $T = 298$  K and a density  $\rho = 0.9997$  g/cm<sup>3</sup> in a cubic box. Every time that a water molecule was moved, the coordinates of its  $\delta$  particles were computed by solving their zero force equations in the field produced by all other water molecules in a sphere of radius 6 Å. The enthalpy  $H$  after 1000  $K$  configurations, having started from a convergent configuration of the MCY po-

TABLE I. Parameters for the MCHO model when fitted to the two-body surface of Matsuoka *et al.* (Ref. 6), the three-body nonadditivity surface from Clementi *et al.* (Ref. 8), and the four-body nonadditivity surface of Gil-Adalid and Ortega-Blake (Ref. 9). Values give energies in kcal/mol when distances and charges are in a.u. The values for  $C_{1000}$  and  $C_{1001}$  have to be multiplied by 627.5

	O-O	O-H	H-H
A	$3.4714017 \times 10^7$	691458.52	15287.312
B	$-1.7487824 \times 10^6$	-34475.771	$1.1834726 \times 10^{12}$
$n_1$	3.5757985	1.4891927	0.63156768
$n_2$	3.9774684	0.70067690	15.202894
$C_{1000}$	1.000000	1.000000	1.000000
$C_{1001}$		0.46762433	0.000000
$\lambda_{1000}$		99.545247	
$n_{1000}$		2.4182009	
$\lambda_{1001}$		7118.5641	
$n_{1001}$		6.4901068	
	Oxygen	Hydrogen	
$q$	59.012287	455.05127	
$\delta$	-1.4261645	0.713108728	

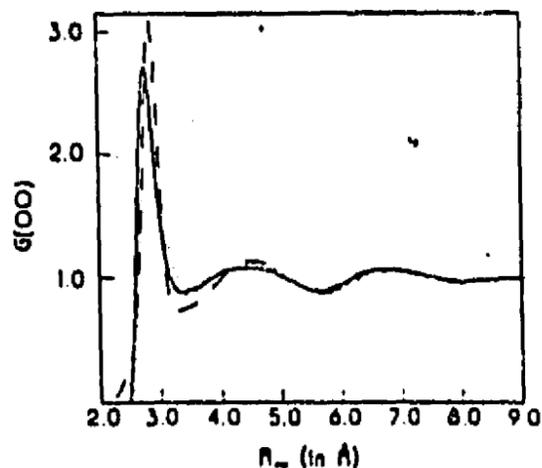


FIG. 5. Oxygen-oxygen radial distribution function. The dotted line corresponds to the experimental curve of Narten and Levy (Ref. 22), the broken line to the experimental curve of Soper and Phillips (Ref. 21), and the solid line to the theoretical prediction of the MCHO model.

tential<sup>20</sup> and rejecting the first 250  $K$  configurations, is  $-8.1 \pm 0.06$  kcal/mol, which compares rather well with the experimental value<sup>12</sup> of  $-8.1$  kcal/mol. The practically full agreement is coincidental, since here we are lacking vibrational motions and quantum correction.<sup>15</sup> However, both terms have been estimated to be small. A molecular dynamics simulation including vibrations<sup>17</sup> suggests a value of 0.416 kcal/mol for this correction. The quantal correction has been estimated to be 0.2 kcal/mol by Owicki and Scheraga<sup>16</sup> with a semiempirical procedure, but 0.93 kcal/mol by Lie and Clementi.<sup>17</sup>

Since we started from a minimum of the MCY potential, there is a possibility that our simulation could stay in a local minimum, so we decided to restart the simulation from a random configuration to test this possibility. We rejected the first 300  $K$  configurations and we found the same value of  $-8.1$  kcal/mol for the enthalpy after 200  $K$  configurations.

Apart from the enthalpy values one can obtain from the Monte Carlo several features to compare the model predictions with their experimental counterparts. Here we compare several of them.

#### Radial distribution functions

In Figs. 5, 6, and 7 we present the comparison between the theoretical and the experimental curves. There has been a lot of discussion (see, e.g., Refs. 18, 19, and 20) on the validity of these experimental curves, leading to a second experimental determination by Soper and Phillips<sup>21</sup>, which differs from the "classical" curves from Narten and Levy.<sup>22</sup> Here we can see that the theoretical curves are quite close to the experimental ones. The only discrepancy being in the oxygen-oxygen radial distribution function, a shift of the first peak towards closer distances; this could be due to the lack of vibrations in the model, hence not allowing for the stretching of the O-H bond upon hydrogen bonding. Apart from this deficiency, the radial distribution

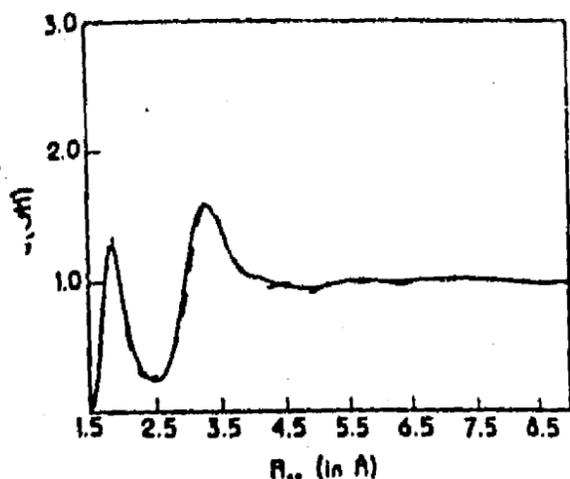


FIG. 6. Oxygen-hydrogen radial distribution function. The broken line corresponds to the experimental curve of Soper and Phillips (Ref. 21), and the solid line to the theoretical prediction of the MCHO model.

functions present the best reported approach of a nonempirical potential to the experimental curves. Certainly it neither shows an exaggerated structure,<sup>1</sup> nor lack of structure<sup>23</sup> when many body terms are included.

#### X-ray and neutron scattering structure functions

The x-ray and neutron structure functions, which are the experimentally observed curves, are a better test for the model. In Figs. 8 and 9 the comparison between experiment and theory is presented. Again, there is very good accordance with experiment, except at very small  $s$  values. This discrepancy could be due to the long range correlation, where a very small oscillation seems to occur. A larger simulation box is required to determine if this is an artifact of the potential or a consequence of border effects of a small box.

#### Thermodynamic parameters

Some other parameters to be compared are the heat capacity, the isothermal compressibility, and the pressure.

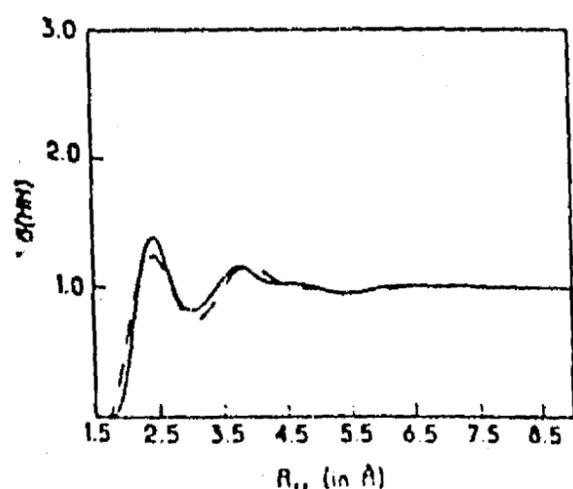


FIG. 7. Hydrogen-hydrogen radial distribution function. The broken line corresponds to the experimental curve of Soper and Phillips (Ref. 21), and the solid line to the theoretical prediction of the MCHO model.

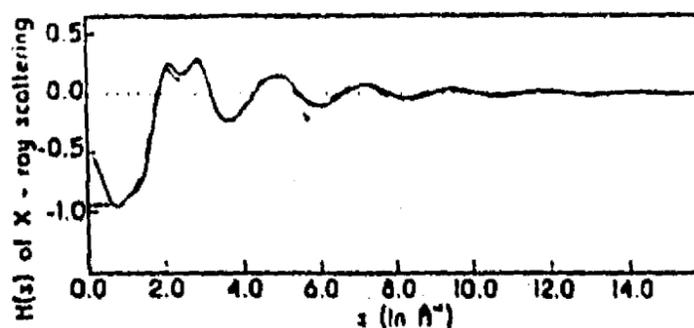


FIG. 8. Structure function  $H(s)$  from x-ray scattering. The broken line corresponds to the experimental curve (Refs. 21 and 22) and the solid line to the theoretical prediction of the MCHO model.

Following Lie *et al.*,<sup>24</sup> we estimated the configurational heat capacity<sup>25</sup> to be  $C_V = 8.1$  cal/mol K from the simulation. If we consider the nonconfigurational contribution (which is made up from mechanical degree of freedom<sup>25</sup>), with a value of either 9.6 or 6.0 cal/mol K, depending on whether mechanical motions are treated as Debye spectra or completely free, then the value lies between 17.7 and 14.1 cal/mol K, close to the experimental value of 17.9 cal/mol K. The isothermal compressibility yielded by the model is  $\kappa = 35 \times 10^{-5}$  atm<sup>-1</sup>, a rather exaggerated value compared to the experimental value of  $\kappa = 46 \times 10^{-5}$  atm<sup>-1</sup>, a consequence of the wrong behavior at  $s = 0$  of the structure functions. Finally, the pressure obtained from the virial expansion<sup>22</sup>

$$\frac{p}{\rho kT} = 1 - \frac{1}{6NkT} \left( \sum_i \sum_{j>i} r_{ij} \frac{\partial V_{ij}}{\partial r_{ij}} \right) \quad (4)$$

is  $p = 8822$  atm, which is similar to the 8500 atm computed for the original MCY potential.<sup>17</sup> This is no surprise, since both were adjusted to the same two-body surface, and it has been argued<sup>17</sup> that this high pressure is due to the lack of *ab initio* computed energies for configurations in the repulsive region.

#### CONCLUSIONS

Here we have presented a novel analytical model for intermolecular interactions. This model has been tested in

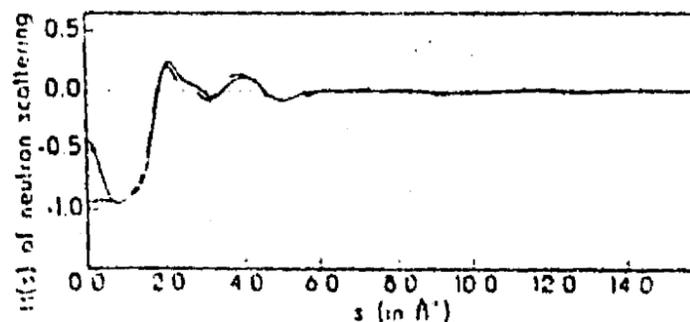


FIG. 9. Structure function  $H(s)$  from neutron scattering. The broken line corresponds to the experimental curve of Soper and Phillips (Ref. 21) and the solid line to the theoretical prediction of the MCHO model.

a stringent case, the water-water interaction, and proved to be successful. In addition, the model has several attractive features:

(a) It includes the charge rearrangements produced by the molecular interaction, thus being better able to fit costly *ab initio* surfaces. In fact, it has produced the best reported fit to the MCY surface.

(b) It has built in nonadditivity, mainly (but not only) of induction type. It has proven capable of a good—and simultaneous—reproduction of two-, three-, and four-body terms. This is not the only water-water potential which includes nonadditivity. Similar potentials have been proposed by Clementi *et al.*,<sup>1,2,3</sup> fitting the *ab initio* computed three- and four-body terms with a truncation of an induction-type series, and by Lybrand and Kollman,<sup>28</sup> with a semiempirical approach, considering the interaction of the electric field of the surrounding molecules with an induction-type term. There is an important difference with respect to the present model: here the nonadditivity is included not only through an effect in the potential energy calculation, but also in the evolution of the molecular configuration since the molecule charges respond and adjust to the electric field of the surrounding molecules and will reflect this in the next Monte Carlo move. Of course, what substantiates the present model is its good ability to reproduce experimental observations.

(c) It maintains a site-site formalism facilitating its use in numerical simulations and making it of general use in different fields of physics; furthermore, it has performed better than *ad hoc* analytical forms for the water-water interaction.

(d) It is not expensive in computational terms. For instance, Detrich *et al.*<sup>26</sup> mention that simulations including three-body potentials cost 20 times as much as those corresponding to two body, and those including four-body potentials are even more expensive by 2 orders of magnitude. By contrast, simulations with the model presented here are only approximately 10 times slower than simulations using the MCY potential. We are sure that all these features make the present model very useful for simulation studies.

*Note added in proof:* In a very recent work, Niesar *et al.*<sup>30</sup> extended the MCY potential to include the four-body nonadditivity due to polarization. The results of different

predicted properties show very good agreement with experiments. This stresses the importance of including the four-body effects.

#### ACKNOWLEDGMENT

We would like to acknowledge J. Hernández for all his help in the computational effort.

<sup>1</sup>E. Clementi and G. Corongiu, *Int. J. Quantum Chem. Quantum Biol. Symp.* 10, 31 (1983).

<sup>2</sup>M. M. Probst, T. Radnai, K. Heinzinger, P. Böppa, and B. M. Rode, *J. Phys. Chem.* 89, 753 (1985).

<sup>3</sup>I. Ortega-Blake, J. Hernandez, and O. Niwano, *J. Chem. Phys.* 81, 1294 (1984).

<sup>4</sup>P. A. Egelstaff, *Adv. Chem. Phys.* 53, 1 (1983).

<sup>5</sup>D. J. Shanks, *Math. Phys.* 34, 1 (1953).

<sup>6</sup>O. Matsuoka, E. Clementi, and M. Yoshimine, *J. Chem. Phys.* 64, 1351 (1976).

<sup>7</sup>VAOSAD, *Harwell Subroutine Library*, Division of Computer Science and Systems, Harwell Laboratories, Harwell, U. K. (1974).

<sup>8</sup>E. Clementi, W. Kolos, G. C. Lie, and G. Ranghino, *Int. J. Quantum Chem.* 17, 377 (1980).

<sup>9</sup>L. Gil-Adalid and I. Ortega-Blake, *J. Chem. Phys.* (submitted).

<sup>10</sup>J. Verhoeven and A. Dymanus, *J. Chem. Phys.* 52, 3222 (1970).

<sup>11</sup>T. R. Dyke, K. M. Mack, and J. S. Muenster, *J. Chem. Phys.* 66, 454 (1977).

<sup>12</sup>H. A. Gebbie, W. J. Burroughs, J. Chamberlain, J. W. Harries, and R. G. Jones, *Nature* 221, 143 (1969).

<sup>13</sup>K. Szalewicz, S. J. Coe, W. Kolos, and R. Bartlett, *J. Chem. Phys.* 89, 3662 (1989).

<sup>14</sup>N. E. Dorsey, *Properties of Ordinary Water-Substance* (Reinhold, New York, 1940).

<sup>15</sup>P. H. Berens, D. H. J. Mackay, G. M. White, and K. R. Wilson, *J. Chem. Phys.* 79, 2375 (1983).

<sup>16</sup>J. C. Owicki and H. A. Scheraga, *J. Am. Chem. Soc.* 99, 7403 (1977).

<sup>17</sup>G. C. Lie and E. Clementi, *Phys. Rev. A* 33, 2679 (1986).

<sup>18</sup>E. Clementi, *J. Phys. Chem.* 89, 4426 (1985).

<sup>19</sup>A. K. Soper, *Chem. Phys.* 88, 187 (1984).

<sup>20</sup>A. K. Soper, *Chem. Phys.* 107, 61 (1986).

<sup>21</sup>A. K. Soper and M. G. Phillips, *Chem. Phys.* 107, 47 (1986).

<sup>22</sup>A. H. Narten and H. A. Levy, *J. Chem. Phys.* 55, 2263 (1971).

<sup>23</sup>U. Niesar, G. Corongiu, M. J. Huang, M. Dupuis, and E. Clementi, *Int. J. Quantum Chem. Quantum Chem. Symp.* 23, 421 (1989).

<sup>24</sup>D. Eisenberg and W. Kauzmann, *The Structure and Properties of Water* (Oxford University, New York, 1969).

<sup>25</sup>G. C. Lie, E. Clementi, and M. Yoshimine, *J. Chem. Phys.* 64, 2314 (1976).

<sup>26</sup>J. Detrich, G. Corongiu, and E. Clementi, *Chem. Phys. Lett.* 112, 426 (1984).

<sup>27</sup>L. Verlet, *Phys. Rev.* 129, 98 (1967).

<sup>28</sup>T. P. Lybrand and P. A. Kollman, *J. Chem. Phys.* 83, 2923 (1985).

<sup>29</sup>This equilibrium configuration was kindly made available to us by Dr. G. Corongiu.

<sup>30</sup>U. Niesar, G. Corongiu, and E. Clementi, *J. Phys. Chem.* (in press).

Chapter 3

**THE MICROSCOPIC STRUCTURE AND DYNAMICS OF WATER  
AT A SURFACE**

**Humberto Saint-Martin and Iván Ortega-Blake**

**TABLE OF CONTENTS**

I.	Introduction.....	68
II.	Theoretical Microscopic Studies .....	68
III.	Bulk Water .....	69
IV.	Theoretical Studies of the Wall-Water Interphase .....	73
V.	Water Structure Near a Flat Dielectric Wall .....	78
A.	Results .....	79
1.	Water Density Profile Along the z Axis .....	79
2.	Hydrogen Bond Pattern at the Surface .....	81
3.	Dipole Orientation .....	81
4.	Dipole Value and Dipole Correlation .....	81
VI.	Discussion .....	85
VII.	Conclusions.....	88
	References.....	88

## I. INTRODUCTION

The interest in understanding the behavior of a solvent surrounding a nonpolar solute and, in particular, a surface is now well documented, leading to several books devoted to this topic (see, for example, References 1 and 2). The reasons are manifold: basic research in statistical mechanics of liquids to develop analytical approaches for complex problems; the use of new powerful techniques, Monte Carlo and Molecular Dynamics, for envisaging the microscopic role of the solvent in an interesting system; the biochemical interest because of the relevance to biological membranes and enzymic catalysis; and the industrial interest of micelles, microemulsions, and surfaces.

The very general appearance of the phenomena of water solvation of surfaces in physical, chemical, and biological systems, being metallic arrangements or proteins, has led to an enormous amount of research. Particularly, this has permitted the establishment of phase diagrams of mixtures of water, oil, and surfactant. These mixtures can have different forms of aggregation: lamellar, bicontinuous, or intercontiguous phases, and discrete phases as micelles or reverse micelles. The appearance of a particular phase depends on the ratios of the components of the mixture and also on the salinity, temperature, and other environmental factors that can change from one phase to another without apparent change in the macroscopic properties, but it can produce changes in the electric conductivity and it can certainly conduce to quite different phenomena in solvation, catalysis, etc. It is thus very important to understand the mesoscopic description of the system and much theoretical work has been devoted to this, but once it is clear that a particular phase has been attained, for instance reverse micelles, understanding the phenomena will come from a microscopic description.

Microscopic understanding of the phenomena can come from experiments directed to look into the molecular properties and also, with an increasing frequency, from theoretical methods. This work is devoted to an analysis of the information provided by microscopic theoretical methods and the comparison with available experimental work that will permit the elaboration of a molecular picture to start guessing the behavior of the solvent at the interphase.

## II. THEORETICAL MICROSCOPIC STUDIES

Very soon after the establishment of quantum mechanics, the straightforward application of physics, the Schrödinger equation, was pursued for the understanding of the microscopic properties of the matter. More than half a century afterward, and due mainly to the advance in computers, it is now possible to obtain molecular and intermolecular properties of small molecules with reliability and criteria to assess this reliability (see, for example, Reference 3). An extension of this idea of obtaining the microscopic behavior from the first principles is to compute the intermolecular potential of a substance, e.g., water, and then to construct a numerical model with a few hundred of these molecules to simulate the substance, e.g., liquid water. This approach has been championed by Enrico Clementi for many years and is now proving to be quite successful and is becoming popular. On the other hand, coming not from quantum mechanics, but from statistical mechanics, is the approach initiated by Bernal, Stillinger, and others of constructing classical potentials for the intermolecular interaction and adjusting their parameters to reproduce known macroscopic properties. Both manners are now used to obtain potentials to be employed in Monte Carlo and Molecular Dynamics simulations. In the Monte Carlo method (see, for example, Reference 4), the molecular movements are performed randomly, using the computed potential energy to sample the configuration space in an intelligent manner. Molecular Dynamics (see, for example, Reference 5) solves Newton's equations of motion with the intermolecular forces derived from the adjusted potential. In both methods, one looks into the statistical averages

to obtain the expected values of the molecular properties. Hence it is very important to have a large enough sample, either of the configurational space or of the phase space, to have confidence in the obtained values which are also critically affected by the number of elements considered to simulate the ensemble. Of course, the molecular potential used is crucial since the obtained properties will stem from it.

From our point of view, it is more convenient to develop the intermolecular potential from first principles, mainly because it allows a stepwise refinement of the model that should always conduce to a better representation of the system. For instance, there has been a permanent interest in reproducing liquid water by numerical simulations and many potentials have been used. However, it has been recently shown that the proper reproduction of structural properties requires a very accurate description of the intermolecular interaction;<sup>8,9</sup> that is, the physical model for the water molecule has to include effects such as polarizability and many body corrections. Since reproduction of bulk water is very important for the study of its behavior at the interphase, we will now present a Monte Carlo simulation of liquid water with a large number of waters and made with a very refined potential.

### III. BULK WATER

A Monte Carlo simulation was performed on a box of 1668 water molecules and dimensions  $\Delta x = 32 \text{ \AA}$ ,  $\Delta y = 32 \text{ \AA}$ , and  $\Delta z = 48 \text{ \AA}$  having periodic boundary conditions. The initial configuration was generated from the combination of several boxes of 343 waters that had attained equilibrium. The large box was allowed to equilibrate for  $5 \times 10^5$  Monte Carlo steps until the variation in the energy of the ensemble was  $\pm 0.05$  kcal/mol. One can obtain from the simulation thermodynamic and structural parameters that can be compared with their experimental counterparts. The potential energy per water molecule obtained in this simulation is  $7.96 \pm 0.05$  which compares rather well with  $8.1 \pm 0.4$ , the experimental value.<sup>9</sup> It is interesting to note that a value of  $8.2 \pm 0.05$  was obtained from a simulation with the same potential and 343 waters, showing that there is still a dependence of the molecular properties for this number of waters.

There is one structural parameter which is particularly important: the radial distribution function (rdf). This function is obtained by computing the average number of particular atoms, or molecules, that exist in a shell of radius  $r$  centered on the atom we are looking to correlate with, of course normalized to an average isotropic density. For instance, for a crystal, the rdf will have sharp peaks and regions of zero value since there are some distances at which there is a number of atoms, while for most of the distance values there is none. On the other hand, for a homogeneous gas, the distribution of atoms is isotropic and the rdf will have a unit value for all distances except the very close excluded volume. A liquid will have an rdf which is a case intermediate between the above two. The more structure that is in the solvent, the more structure that will appear in the function. A strong advantage of the rdf is that it can be obtained experimentally through X-ray or neutron diffraction, and therefore provides a test for the simulation. As a matter of fact, for a long time there has been a strong desire in reproducing the experimental rdf of liquid water as close as possible. In Figures 1 to 3, we present the rdf for oxygen-oxygen, oxygen-hydrogen, and hydrogen-hydrogen obtained from the simulation as well as their experimental counterparts. We can see that there is quite a good agreement, except for the first peak of the rdf for oxygen-oxygen, possibly due to the lack of vibrations in the physical water model. This is the first time that an *ab initio*-based potential correctly reproduces the second peak which is related to the tetrahedral structure and thus to the water network. The same potential used with 343 waters<sup>7</sup> does not present such an agreement, showing again that for 343 waters there is still an impact on the statistical properties. There are several comparative studies on the qualities of different water potentials (see, for example, References 12 and 13). It is clear now that

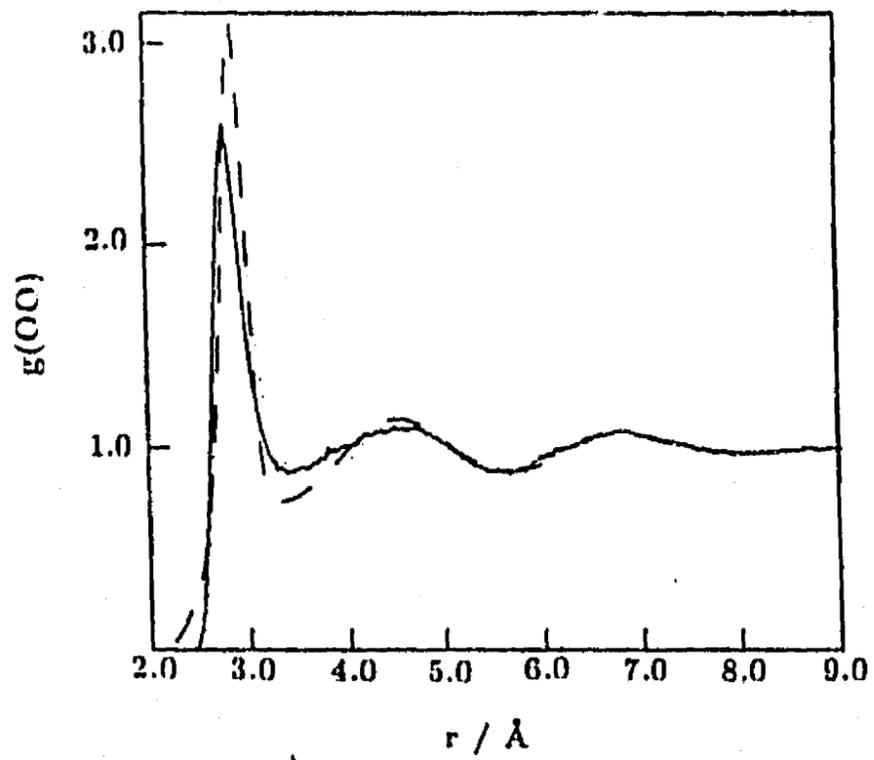


FIGURE 1. Oxygen-oxygen radial distribution functions. The full line represents the theoretical curve, the dotted line represents the experimental curve of Narten and Levy,<sup>10</sup> and the broken line represents the experimental curve of Soper and Phillips.<sup>11</sup>

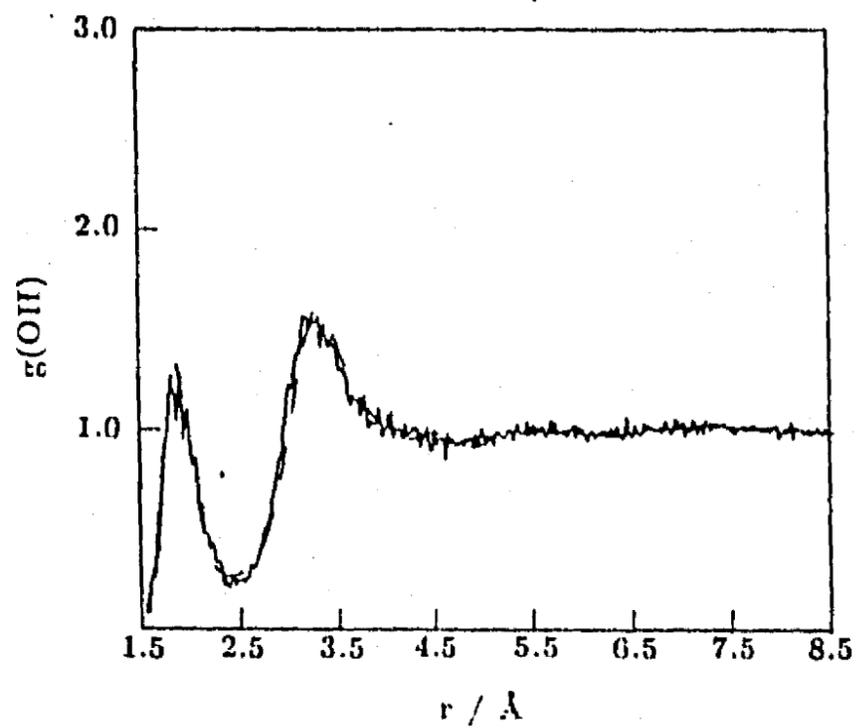


FIGURE 2. Oxygen-hydrogen radial distribution functions. The full line represents the theoretical curve, and the broken line represents the experimental curve of Soper and Phillips.<sup>11</sup>

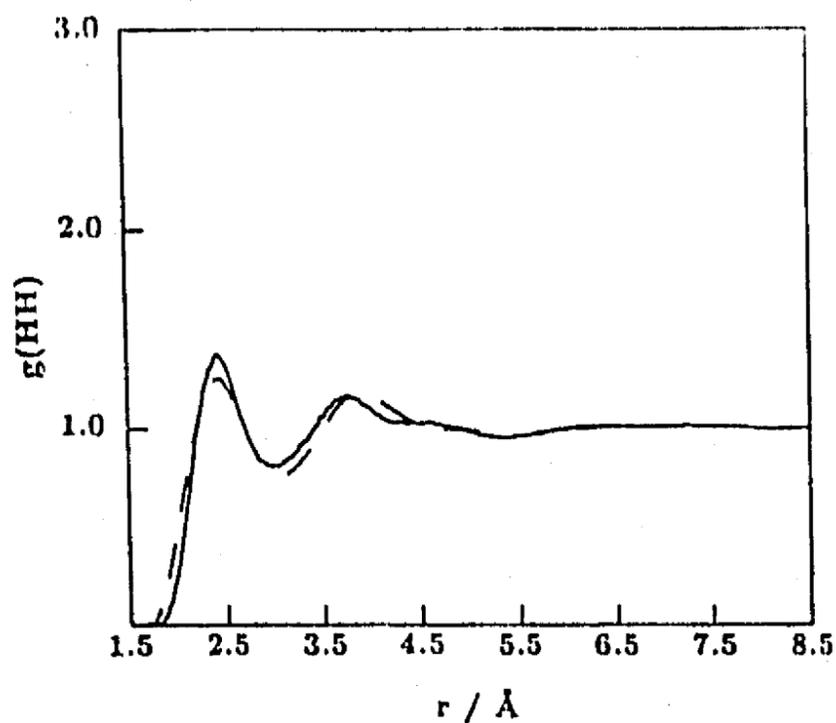


FIGURE 3. Hydrogen-hydrogen radial distribution functions. The full line represents the theoretical curve, and the broken line represents the experimental curve of Soper and Phillips.<sup>11</sup>

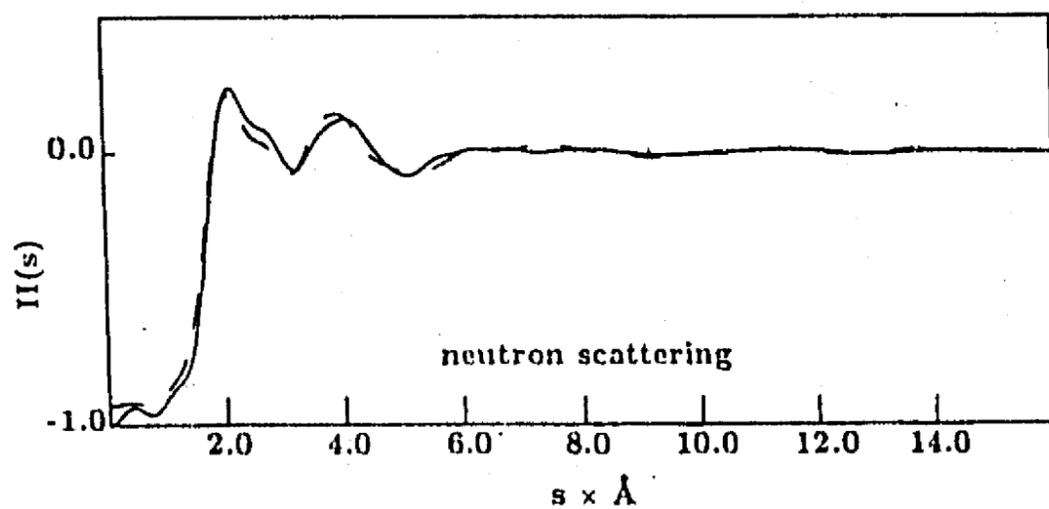


FIGURE 4. Structure function H for X-ray scattering. The full line represents the theoretical prediction and the broken line represents the experimental curve of Narten.<sup>12</sup>

the only manner to reproduce accurately this structural function is through the refinement of the physical model, and in this respect the first principle's potential is advantageous because of the knowledge of the approximations done and the way to improve them. In addition to this advantage, one can be confident that if the structure of the bulk is well predicted, not adjusted, then similarly good predictions of the water structure in other conditions will be obtained. A more stringent test for the predictions of water structure is in the actual experimental observation, the structure function H. In Figures 4 and 5, we present the theoretical and experimental curves for neutron and X-ray scattering. Again there is very good agreement and an improvement over the 343 waters simulation.<sup>7</sup>

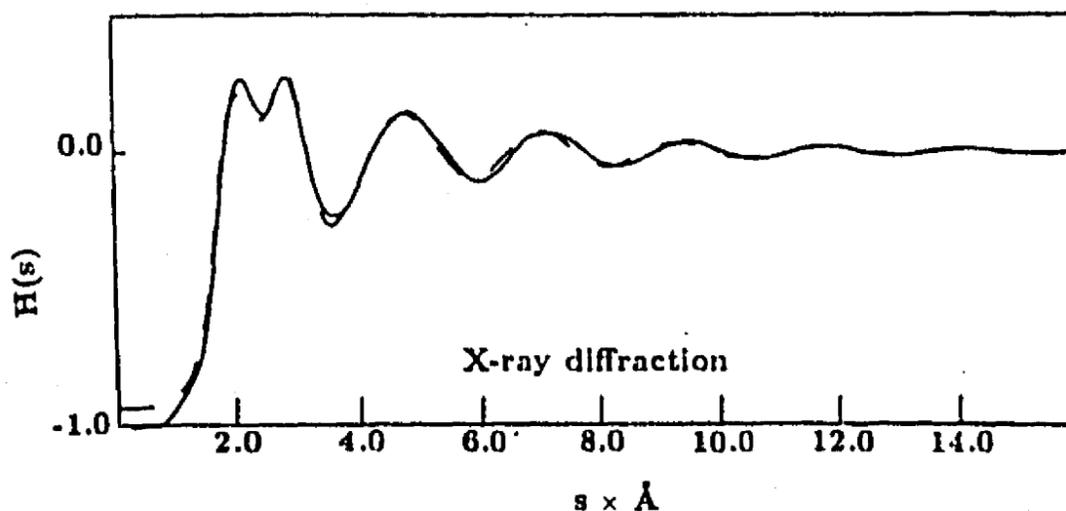


FIGURE 5. Structure function  $H$  for neutron scattering. The full line represents the theoretical prediction and the broken line represents the experimental curve of Narten.<sup>14</sup>

From the above results, it is quite clear that this water potential is able to reproduce the energetic and structural properties with a small loss of accuracy if the simulation is done with a reduced number of waters ( $\sim 300$ ). There is, however, an important deficiency in the potentials fitted to the *ab initio* values: the pressure is very far from the experimental value, and it is clear that this is a value difficult to reproduce in general (see, for example, Reference 15). Hence, with this last concern in mind, we can expect that the structural properties of water are well investigated with highly refined *ab initio*-based potentials. The reason for the improved refinement of the structural properties is the inclusion of physical phenomena in the water model, and since these phenomena are relevant for the discussion at the end of this work, we would like to extend a bit over them here.

Water polarization is obviously an important phenomenon in the water interaction, but there are many water models that are able to reproduce the dimer interaction with a rigid water in both the positive (nuclei) and the negative (electrons) charge positions. However, when water has to respond collectively, the rigid molecule is forced to do it as a whole molecule, and no intramolecular relaxation of dipoles is allowed. Thus, in spite of a good reproduction of the interaction energy, other properties, such as the dielectric constant, will be impaired. Another important phenomenon is the many-body nonadditivity, i.e., the effect that a third or fourth molecule will have in the interaction of the two or three other molecules. A decade ago it was assumed that for water this correction was small and that its inclusion would be just a fine point. Now it is quite clear that collective properties depend in a relevant manner on these corrections; furthermore, for the liquid state, this correction is much larger than expected.<sup>6</sup> Another feature that has been included in the water potential is the vibrational freedom (see, for example, Reference 16). The current idea is that this will have only a small consequence on the properties of the liquid, but this still has to be clearly established. Charge transfer has been circumvented by flexible analytical forms that can account for their obvious action, even at the pairwise interaction, but it could be that some particular phenomena will require its explicit inclusion. The potential that was used in the bulk simulation just reported, and the lamellar simulations that will be presented afterward, includes polarization and three and four body corrections, not only the induction component, but adjusted to the *ab initio* nonadditive values.

Before entering into the simulation of an interphase, we will present a summary of previous works and an analysis of the general panorama that emerges from them.

#### IV. THEORETICAL STUDIES OF THE WALL-WATER INTERPHASE

In Table I, we present a summary of the theoretical works made in the last decade for the wall-water interphase, mainly numerical simulations. Without being exhaustive, this table gives a good glance on the efforts made and the results that emerged from the simulations. First of all, it is clear that the theoretical simulations have become more frequent in recent years and have become more able to yield extensive information about the system. Early work, up to 1985, was performed with models of wall that had some simple forms of repulsion and with poorly refined water potentials. These latter were rigid and nonpolarizable and had proven to yield reasonable predictions of the bulk water structure, but still far from accurate descriptions. Nonetheless, the results are quite clear and consistent. The water density in the neighborhood of the wall has an oscillatory damped behavior, i.e., it shows an increase in density at the wall because of close packing, but this structural phenomenon is buffered pretty fast by the liquid. The density of the bulk is regained in less than 10 Å. Of course, one has to bear in mind the size of the computational cell which is of similar dimensions. There is also a general agreement on the structural effects near the wall, enhanced structure, similar to the ice Ih (where h refers to the normal hexagonal lattice as opposed to ice Ic, with a cubic structure), and an important role of the hydrogen bond. For the dynamical properties, there is no such agreement. Water self-diffusion is proposed to be both faster and slower at the interphase, and the only agreement is that variations are of less than an order of magnitude. A similar pattern appears for the dipole-dipole correlation. In a subsequent phase of the theoretical work, the last 5 years, a search for complexity in the model is characteristic, i.e., simulating the interphase with much more atomic detail and introducing different physical environments, such as charged surfaces or fields. Intermolecular potentials for water, as well as those for the other species now included, were kept simple by necessity. Of course, with much more variability, the modeling of the interphase is now quite complex and the predicted properties show also a certain degree of variability. However, a general result established in a very consistent manner is the absence of structural effects produced by the interphase at distances larger than a few Å (~8). This finding also appears in more recent works where the intermolecular potentials are more refined and will be discussed afterwards. In the vicinity of the interphase, we can see the following effects:

1. The water density presents an oscillatory damped behavior similar to that observed for primitive walls, except for two cases.<sup>27,30</sup> In these cases, the lack of oscillations is attributed to the presence of ions in the simulations, but there are other simulations with ions that show a strong oscillatory behavior.<sup>31,34</sup>
2. All the models with atomic detail of the interphase show a broad boundary, i.e., not a well-defined border, but a rough mixture of the two phases. A very good example can be seen in Figure 6.
3. A general and striking result is observed in the models that include electric fields either through charged surfaces or in a continuous manner. The electric field is almost immediately canceled by water, contrary to the textbook idea that the Gouy-Chapman phase extends over tens of Angstroms.
4. There seems to be a more-packed structure near the wall, similar to ice Ih.
5. The results on the dynamic properties are similar to those reported by the simulations with primitive walls; i.e., there is a discrepancy about the values being increasing or decreasing, but they retain the same order of magnitude as those in the bulk.
6. The inclusion of ions in the model allows us to look into their distribution. However, the results are contradictory, predicting condensation of ions near the wall<sup>34</sup> or repulsion by the surface.<sup>31</sup>

TABLE I  
Theoretical Studies of the Wall Interphase

Year	Model	Potentials	Results	Ref.
1981	Box with periodic conditions in x and y Two hard walls of infinite repulsion on z $\Delta z$ between walls: 20 Å 150 water molecules Monte Carlo, $3.5 \times 10^6$ conf. T = 300 K	Water-water MCY <sup>24</sup> Rigid, nonpolarizable <i>ab initio</i> Cubic cutoff	$\rho(z)$ , oscillatory damped behavior with 3 peaks; max amplitude = 1.25, max range = 7.5 Å No wetting present Enhanced hydrogen bonding in the surface Orientation of dipole near the wall	17
1983	Box with periodic conditions in x and y Hard repulsive walls, with an "elastic" potential, $\phi(z) = a \exp(\beta z)$ $\Delta z$ between walls: 20 Å 150 water molecules Molecular dynamics, 25 ps T = 302 K	Water-water ST2 <sup>25</sup> Rigid, nonpolarizable Empirical Spherical cutoff, r = 7.4 Å	$\rho(z)$ has an oscillatory, damped behavior, with 2 peaks; max amplitude = 0.2, max range = 5 Å No wetting present Water dipole orientation parallel to the wall Dipole-dipole correlation smaller than bulk rdf shows higher peaks (more structure) close to the wall $D_{\perp}$ in the wall is half of that in bulk $D_{\parallel}$ is not affected $D$ near the wall is 20% less than in bulk Dielectric relaxation in the wall 8.5 GHz and 17 GHz in the bulk	18
1983	Box with periodic conditions in x and y Lennard-Jones walls, with a potential similar to that of water Walls separated until zero pressure is obtained $\Delta z$ between walls: 25 Å 216 water molecules Molecular dynamics	Water-water ST2 <sup>25</sup>	$\rho(z)$ behavior is very dependent on the level of convergence to equilibrium and the number of configurations considered ( $\Delta t$ ); is it possible that there will be no oscillation in a large equilibrium sample?; in this one, the behavior is oscillatory, damped with 3 peaks; max amplitude = 0.1, max range = 10 Å No wetting present Water dipole orients parallel to wall Orientational correlation extends to 10 Å $D_{\perp}$ is 30% larger than $D_{\parallel}$ away from the wall, but they are equal at the interphase Hydrogen bonding is only 75% at the interphase	19
1984	Box with periodic conditions in x and y Lennard-Jones wall, simulating paraffin $\Delta z$ between walls: 24 Å 216 water molecules Molecular dynamics, 72 ps T = 290 K	Water-water ST2 <sup>25</sup> Water-water ST2 <sup>25</sup> without charges Spherical cutoff, r = 8.2 Å	$\rho(z)$ for a simple liquid (ST2 without charges) shows strong oscillations which extend for a long range, and there is wetting; for water (full ST2) the behavior is oscillatory and damped, max amplitude = 0.2, max range = 7 Å, and there is no wetting Ice Ih-like structure at the wall Orientational structure buffers at 6 Å The molecular orientation is governed by HB interactions Water molecules close to the wall preserve 75% of their HBs	20

TABLE I (continued)  
Theoretical Studies of the Wall Interphase

Year	Model	Potentials	Results	Ref.
1985	Sphere with restricting forces, $r = 9 \text{ \AA}$ 101 water molecules Molecular dynamics, 30 ps $T = 301 \text{ K}$	Water-water TIPS2 <sup>26</sup> Rigid, nonpolarizable Empirical Spherical cutoff, $r = 8.5 \text{ \AA}$	The second peak of the rdf is washed out compared to bulk, which implies loosening of the tetrahedral structure Explicit orientation in the wall Structural preference is not propagated deep into the liquid, and this is assumed to be due to the spherical topology Water molecules close to the interphase sacrifice one HB in order to maximize the total interaction	21
1986	A mixture of 1094 water molecules with 15 sodium octanoate in a $34 \text{ \AA}$ cubic box Rescaling to maintain constant energy Molecular dynamics, 30 ps (FC) and 72 ps (RC) $T = 294 \text{ K}$	Water-water SPC <sup>27</sup> Rigid, nonpolarizable empirical (FC) Lennard-Jones type interaction for other atoms Dihedral potential for rotating heads Truncation of long-range coulombic potential Spherical cutoff, $r = 10 \text{ \AA}$ Charges reduced to 50% (RC) to account for dielectric constant	$\rho(r)$ extends away from micelle with only one peak (RC) and no wetting present Broad boundary in the micelle, very dependent on the potential Same dynamics for water in the interphase and in the bulk, so there is no support for the concept of bound water	22
1986	Box with periodic conditions in x and y Walls with different dielectric constant and charges, 77.3 and 5.0, respectively $\Delta z$ between walls: $21 \text{ \AA}$ Monte Carlo, $2 \times 10^5$ conf. $T = 301 \text{ K}$	Water continuum model Electrostatic ion-ion, ion-wall, and ion-water interactions	The ions are repelled because of the low dielectric constant of the wall	23
1986	Box with periodic conditions in x and y Lennard-Jones type walls $\Delta z$ between walls: $24 \text{ \AA}$ 200 water molecules 8 lithium iodine molecules Molecular dynamics, 10 ps $T = 283 \text{ K}$	Water-water ST2 <sup>28</sup> Lennard-Jones type for Li-I, empirical Spherical cutoff, $r = 9.3 \text{ \AA}$	$\rho(z)$ has no oscillations near the wall, assumed to be produced by the presence of ions The Lennard-Jones wall is very short-ranged in structure making ice Ih-like structure near the wall Faster motion near the wall, due to the absence of shear force An influence on the whole lamina increases self-diffusion The water molecules preserve 75% of their HB close to the wall	24
1987	Box with periodic conditions in x and y Hard walls $\Delta z$ between walls: $20 \text{ \AA}$ 125 water molecules Monte Carlo, $2.02 \times 10^6$ configurations $T = 399 \text{ K}$	Water-water TIPS2 <sup>26</sup>	$\rho(z)$ oscillatory damped behavior but nonsymmetric; max amplitude = 0.2, max range = $8 \text{ \AA}$ No wetting present No electric polarization near the surface Ice Ih structure, modest	25

TABLE I (continued)  
Theoretical Studies of the Wall Interphase

Year	Model	Potentials	Results	Ref.
1987	Box with periodic conditions in x and y Metallic wall with image charge-interactions Gouy-Chapman double layer on one wall $\Delta z$ between walls: 32 Å 200 water molecules Monte Carlo, $2.0 \times 10^6$ conf. T = 399 K	Water-water TIPS2 <sup>26</sup>	The field has a minor effect on the structure described in the previous entry The effect of the field is negligible at the neutral wall Hydrogen bonding between waters controls the structure A metallic surface freezes a thin layer of water in a rigid ice 1h structure There is wetting for the metallic surface There are no effects on polarization and angular distribution	26
1988	Mixture of water and sodium octanoate in a 3-D box to form micelles Length of box = 34 Å 1068 water molecules 15 sodium octanoate Molecular dynamics, 50 ps T = 300 K	Water-water SPC <sup>27</sup> Lennard-Jones type for other atoms, with a flexible surfactant Ewald sums for long-range coulombic interactions	The micelles formed are nonspherical but stable $\rho(r)$ outside the micelle does not show any structure, but just an increase up to a plateau The micelle remains partially charged because it is not neutralized completely by the counterion SPC water has a dielectric constant of 65, and the scaling of Reference 22 is not justified The cutoff used for long-range interactions in Reference 22 could yield a difference with this work	27
1988	Periodic boundary conditions in 3-D mixture of: 526 water molecules 52 Na-decanoate 76 decanol Variable size of the unit cell Molecular dynamics, 30 ps T = 300 K	Water-water SPC <sup>27</sup> Lennard-Jones type for other atoms; generic potentials from combination rules Approximation of forces at 3 different ranges: short, with direct cutoff at 7.5 Å medium, updated every 10 steps long, approximated with a charge density	$\rho(r)$ , a peak in density is suggested, but there is not enough water for a second one Water penetrates into the bilayers Na <sup>+</sup> also penetrates, but less markedly The interphase is very broad and diffuse There is not a diffuse electrical double layer D inside the membrane is 10 times lower than in bulk D in the interphase is 3 times lower than in bulk Reorientation and rotation of water is slower than in bulk Na <sup>+</sup> concentration is maximum at bulk	28
1989	A mixture of water, monovalent ions, and a double layer through macro ions Statistical mechanics with the RHNC closure	Water: sphere with dipole and quadrupole moments Ions with charge and different sizes Electrostatic interactions	$\rho(z)$ has a maximum close to the wall and depending on the surface charge, it can show oscillations which extend deep into the bulk; these oscillations are also dependent on the size of the macro ions The results are very sensitive to the interaction potential Ice 1h-like structure at the surface The ice structure is not broken by the presence of ions The surface charge is screened extremely fast	29

TABLE I (continued)  
Theoretical Studies of the Wall Interphase

Year	Model	Potentials	Results	Ref.
1989	Box with periodic conditions in x and y Wall with a repulsive Lennard-Jones type potential Wall with atomic Lennard-Jones in 3 layers $\Delta z$ between walls: 35 to 50 Å Molecular dynamics	Water-water TIP4P (updated version of TIP2)	Ice Ih-like structure near the surface $D_{\parallel}$ is greater in the surface than in bulk The residence time in the surface is somewhat greater than in bulk, but still small For the hydrophobic surface there is no wetting and the deviations from the bulk density do not extend more than 3 molecular layers While in the hydrophobic surface the water molecules preserve 75% of their hydrogen bonds, close to the silica wall they keep only 25%, and the density profile has only 2 peaks	30
1989	Sphere with a radius of 19 Å, having a repulsive wall at 19.4 Å Mixture of: 1000 water molecules 50 Na 50 carboxylic groups Low density to obtain zero pressure Molecular dynamics, 50 ps	Water-water MCY <sup>24</sup> For ions empirical Lennard-Jones Coulombic interactions are scaled (1/√2) to account for the dielectric constant	$\rho(z)$ has an oscillatory damped behavior with 3 peaks, max amplitude = 0.5, max range = 8 Å Wetting possible Enhanced hydrogen bonding at the wall surface when ions were considered without charge When ions are charged, the hydrogen bonding at the wall surface is similar to that in bulk, but water molecules in contact with ions have a reduced HB Water structure more densely packed up to 9 Å Zero Na <sup>+</sup> concentration at bulk $D_{\parallel}$ is reduced around the ions and the interphase Ions and surface compensate effects Strong orientation of water dipoles in the first hydration shells of ions	31
1990	3-D periodic conditions Mixture of "water" and "oil" particles for a total of 39,304 particles Molecular dynamics, 10 <sup>7</sup> steps T = E/K	Lennard-Jones truncated at 2.5	$\rho(r)$ oscillations in the interphases of surfactant water Formation of micelles and lamellae Depletion layer after interphase	32
1990	Box with periodic conditions in x and y Repulsive walls $\Delta z$ between walls: 24 Å 216 water molecules Molecular dynamics, 96 ps T = 302 K	Water-water NEMO <sup>29</sup> <i>ab initio</i> polarizable spherical cut-off, $r = 8.5$ Å	$\rho(z)$ has an oscillatory damped behavior with 3 peaks, max amplitude = 0.75 max range = 10 Å Wetting can be seen from $\langle z \rangle$ Ice Ih-like structure in the first two layers More structure than with a pair-wise additive potential Induced dipole aligned to the surface Permanent dipole not aligned Molecules at the surface preserve 75% of their hydrogen bonds Polarization of water molecules does not play an important role in the structure	33

TABLE 1 (continued)  
Theoretical Studies of the Wall Interphase

Year	Model	Potentials	Results	Ref.
1991	Box with periodic conditions in x and y Wall with gas-crystal potential, neutral Wall with a DC field = 0.1 V/Å Δz between walls: 20 Å 256 water molecules Molecular dynamics, 30 ps, 70 ps T = 298 K	Water-water SPC-FF, SPC plus polarization and vibrations, adjusted to thermodynamic properties Long-range electrostatic interactions are treated with the reaction field geometry	ρ(z) has an oscillatory damped behavior with 2 peaks, max amplitude = 0.5, max range = Å Wetting is seen in both cases, with and without the DC field, but is more marked in the latter The HB structure is broken by the reduction of the O-H distance at the surface, but the HB is enhanced in the bulk; hence there is low entropy at the wall, but high at bulk D⊥ is enhanced with respect to bulk water	34
1991	Box with periodic conditions in x and y At one z surface COOH groups and tails The other z surface free Surfactant in hexagonal lattice 200 water molecules 64 surfactant molecules Monte Carlo, 1.8 × 10 <sup>6</sup> conf. T = 300 K	Water-water TIP4M Surfactant with generic Lennard-Jones Spherical cutoff with different radii	ρ(z) has oscillatory damped behavior with 2 peaks, and not enough water for a third peak, max amplitude = 0.5, max range = 6 Å Thermal motion of the heads allows water penetration Water dipoles are oriented from the wall to the bulk The first hydration monolayer screens almost completely the membrane force field of the head groups	35

Note: ρ(z) or ρ(r) corresponds to the water density; D⊥ represents the diffusion of water perpendicular to the surface; D∥ represents the diffusion of water parallel to the surface; and D represents the total diffusion.

The results which show agreement in a wide range of surface models and intermolecular potential attain a degree of confiability, but the appearance of discrepancies stresses the fact that many properties are critically dependent on the intermolecular potential.

The most recent works revisit the problem using more refined potentials for water. The results of those simulations are similar to those obtained previously: an oscillatory damped behavior of the density and the ice-like structure at the wall, but, a new physical phenomenon not pointed out by the authors seems to appear: the presupposed bulk density, i.e., that at the middle of the box, where no oscillations are present, shows a decrement suggesting wetting of the surface. A similar result had been observed,<sup>20</sup> but for a simple liquid. This is an important finding since it is a long-range effect and could affect the chemical behavior of water in systems as reverse micelles. In order to shed some light on this point, here we present a simulation of the interphase with a very refined water potential. We want also to clarify the possibility that some of the structure is just a consequence of the lack of equilibration in the simulation, as questioned by Sonnenschein and Heinzinger.<sup>19</sup>

## V. WATER STRUCTURE NEAR A FLAT DIELECTRIC WALL

The model considered here is a water box with 343 waters and periodic conditions in x and y. In the z direction, there are two walls with a distance of 22 Å between them. The wall simulates a medium of dielectric constant  $\epsilon = 3$  that should resemble the membrane. The water molecules interact with the images of all the molecules which appear in the dielectric slab. The cutoff radius for water-water is 8 Å and for the water-water images is

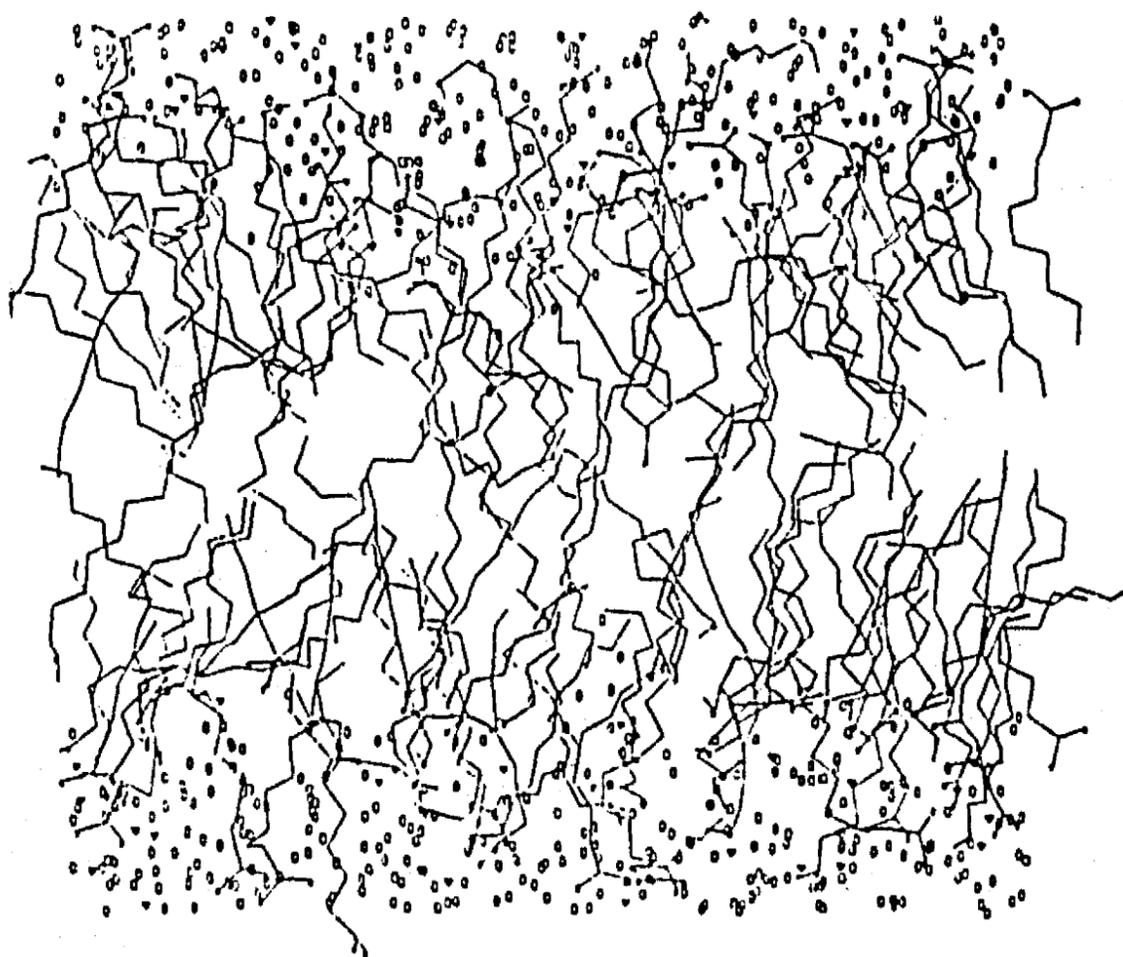


FIGURE 6. A molecular dynamic simulation result for the interphase water, Na-decanoate, decanol. Projection on the  $xz$  plane of a snapshot of the system after 30 ps simulation: decanoate ions and decanol molecules (—), sodium ions ( $\square$ ), water oxygens ( $\circ$ ), and oxygens in the lipids ( $\bullet$ ). (From Egberts, E. and Berendsen, H. J. C., *J. Chem. Phys.*, 89, 1718, 1988. With permission.)

also 8 Å. There is a zone in the immediate vicinity of the wall (0.5 Å) where the image interaction remains constant, i.e., with the value attained at  $z = 0.5$  Å. This zone has the purpose of considering the atomic range where the image model will fail. In any case, this zone only modulates the water-wall interaction; the value of 0.5 Å yields a strong enough interaction in the closest vicinity, and we will be looking to see if such a large perturbation propagates into the bulk.

#### A. RESULTS

First of all, we want to mention that there is indeed a very strong dependence of the observed structure on the number of configurations used to attain equilibrium. In our simulation, it was possible to obtain what looked like equilibrium in the potential energy for as many as a million configurations, but if further configurations were allowed to go, one could observe that a very slow convergence rather than equilibrium had been obtained. Unfortunately, in these not quite equilibrated runs, one could see a behavior of the water density and other structural phenomena similar to those reported previously in the literature. In the results that we are reporting now, we threw away 8 millions of configurations and then made statistics on 3.2 millions of configurations; this length of runs had not been considered before.

##### 1. Water Density Profile Along the $z$ Axis

In Figure 7, we can see the water density profile of the simulated lamella; in Figure 8,

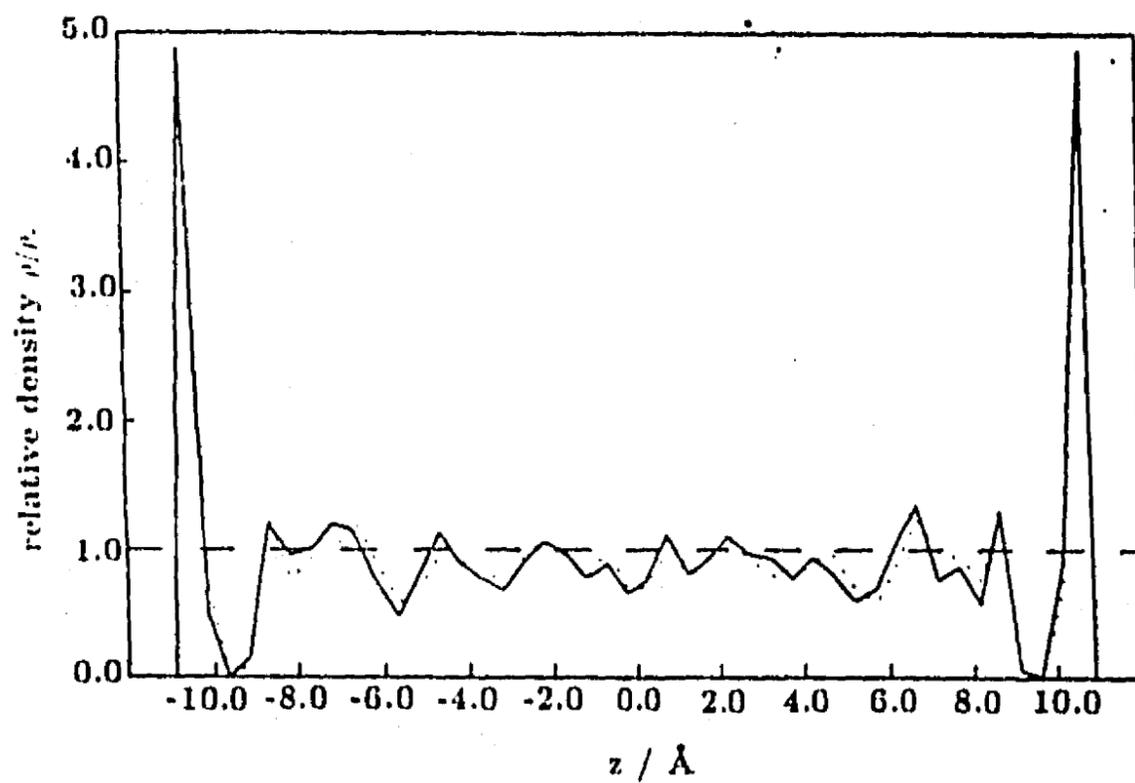


FIGURE 7. The density of water along the  $z$ -axis for the simulation of the lamella (full line). Also shown is the symmetrized function constructed from  $[\rho(z) + \rho(-z)]/2$ , represented by the dotted line.

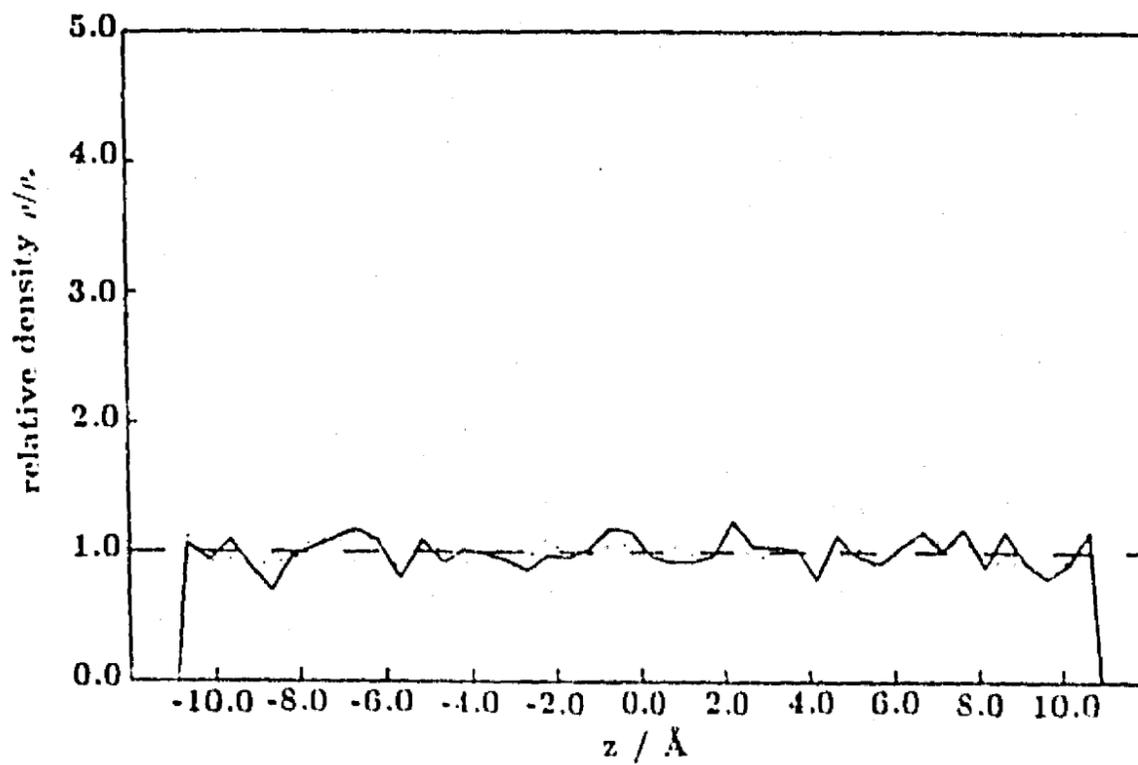


FIGURE 8. The density of water along the  $z$ -axis for the simulation of the bulk (full line). Also shown is the symmetrized function constructed from  $[\rho(z) + \rho(-z)]/2$ , represented by the dotted line.

the corresponding bulk density is presented; and, in both figures, the symmetrized density obtained from the average value between  $\rho(z)$  and  $\rho(-z)$  is shown. We can see in Figure 7 that there is a huge peak of almost five times the ordinary density showing a strong close packing produced by the charges' images. However, in spite of this strong surface effect, the bulk density is attained extremely fast ( $\sim 6 \text{ \AA}$ ). It was quite clear during the long equilibration runs that to destroy any local structure, evidenced by the persistence of peaks in  $\rho(z)$ , required a large number of configurations, indicating an energetic hindrance; this could also be reflected in the need for large relaxation times in molecular dynamics. Apart from being an interesting observation on the capability of water to retain local structure, these findings highlight the need for quite long runs. The symmetrized function does a form of common mode rejection of noise and it is thus more reliable to look into. In this function, we can see that, apart from the fast buffering by the liquid, wetting does indeed appear, in agreement with the other simulations of polarizable models.

## 2. Hydrogen Bond Pattern at the Surface

A good indication of the energetic properties of water at the surface would be the ability to form hydrogen bonds and the distribution of strengths compared to the bulk. In order to estimate this, we considered the interaction energy of close neighboring waters (less than  $3.5 \text{ \AA}$ ); if the energy exceeded  $-4.0 \text{ kcal/mol}$ , we counted them as hydrogen bonded. An analysis of the results as a function of  $z$  is presented in Table 2 along with the distribution of hydrogen bond energy values. From this table, we can see that in the bulk there is a rather homogeneous distribution of hydrogen bonds and a constant average energy. In the lamella, there is a strong modification near the wall, as expected, but no increase in the average strength. We can also see a depletion layer of hydrogen bonds, corresponding to the depletion in  $\rho(z)$ . A more important result is that, on recovering to the bulk, the number of hydrogen bonds is slightly increased while the energy remains almost the same. This result is in agreement with the observation that wetting could be occurring, and the increased ability to form hydrogen bonds is a reflection of the reduced density.

## 3. Dipole Orientation

The dipole orientation in lamella has been one of the most common phenomena looked into in the theoretical simulation. A strong orientation of the dipole will suggest quite a different electric environment than that at the bulk. Here we want to remark that, in a model with a rigid molecule, the permanent dipole should orient by the orientation of the whole molecule, hence being subject to structural hindrance, but in the polarizable models, there is an induced dipole that can orient without the concomitant movement of the molecule. In Figures 9 and 10, we present the dipole orientation of the lamella and the bulk, respectively, that is counting the projection of the dipole along the  $z$ -axis. An isotropic distribution will give a zero value which is observed in the bulk's simulation. In the simulation of the lamella, we can see that, at a few angstroms from the wall, the dipole orientation is the same as that in the bulk.

## 4. Dipole Value and Dipole Correlation

A much-desired parameter of the simulation is the dielectric constant as a function of the  $z$  coordinate since this will let us know to some extent what sort of chemical environment will be occurring. One can obtain the dielectric constant as the response to some field and some model that involves several approximations. The value will be obtained from the expectation values from the dipole and the dipole-dipole correlation. Here we will be looking to these expectation values along the  $z$  axis and compare them to those in the bulk. This will permit us to see any difference that could appear without introducing the approximations of the reaction field model, but, of course, this prevents us from obtaining numerical values for the dielectric constant.

TABLE 2  
Hydrogen Bonding in Bulk Water and in Water  
Between Dielectric Plates

$z$ (Å)	Bulk		Lamella	
	$N_{hb}$	$E_{hb}$ (kcal/mol)	$N_{hb}$	$E_{hb}$ (kcal/mol)
-10.63	1.82	-4.56	0.80	-4.43
-10.13	1.78	-4.54	0.00*	-3.92
-9.64	1.80	-4.55	0.00	0.00
-9.15	1.81	-4.54	1.74	-4.54
-8.65	1.83	-4.55	1.64	-4.49
-7.66	1.87	-4.53	1.96	-4.73
-7.17	1.71	-4.52	1.78	-4.69
-6.67	1.85	-4.57	2.00	-4.57
-6.18	1.80	-4.57	2.08	-4.58
-5.68	1.67	-4.53	1.93	-4.65
-5.19	1.78	-4.56	2.04	-4.63
-4.70	1.77	-4.58	2.02	-4.57
-4.20	1.83	-4.57	2.00	-4.53
-3.71	1.73	-4.53	2.07	-4.57
-3.21	1.82	-4.54	2.10	-4.60
-2.72	1.88	-4.55	1.98	-4.60
-2.22	1.82	-4.55	1.86	-4.56
-1.73	1.74	-4.53	2.00	-4.60
-1.24	1.61	-4.53	1.93	-4.59
-0.74	1.83	-4.54	1.99	-4.59
-0.25	1.84	-4.54	2.15	-4.60
0.25	1.92	-4.56	2.07	-4.60
0.74	1.88	-4.55	2.11	-4.60
1.24	1.71	-4.58	2.10	-4.61
1.73	1.71	-4.55	1.96	-4.53
2.22	1.88	-4.57	2.09	-4.54
2.72	1.82	-4.66	2.18	-4.56
3.21	1.72	-4.55	2.10	-4.55
3.71	1.73	-4.61	2.14	-4.59
4.20	1.65	-4.53	2.04	-4.57
4.70	1.71	-4.60	2.02	-4.59
5.19	1.64	-4.55	1.89	-4.56
5.68	1.70	-4.55	1.96	-4.55
6.18	1.67	-4.55	2.13	-4.60
6.67	1.72	-4.59	1.89	-4.57
7.17	1.74	-4.56	1.83	-4.58
7.66	1.85	-4.55	2.01	-4.62
8.16	1.75	-4.57	1.58	-4.58
8.65	1.70	-4.57	1.73	-4.67
9.15	1.55	-4.53	1.03	-4.41
9.64	1.84	-4.53	0.00	0.00
10.13	1.83	-4.57	0.47	-4.18
10.63	1.77	-4.55	0.86	-4.45

Note:  $N_{hb}$  corresponds to the number of hydrogen bonds, and  $E_{hb}$  corresponds to the average strength.

\* These values are averages over  $3.2 \times 10^7$  configurations; there were very few which had a hb in this position, so that the average  $N_{hb}$  is  $2 \times 10^{-7}$ ; nevertheless, the average strength of those few is -3.92 kcal/mol.

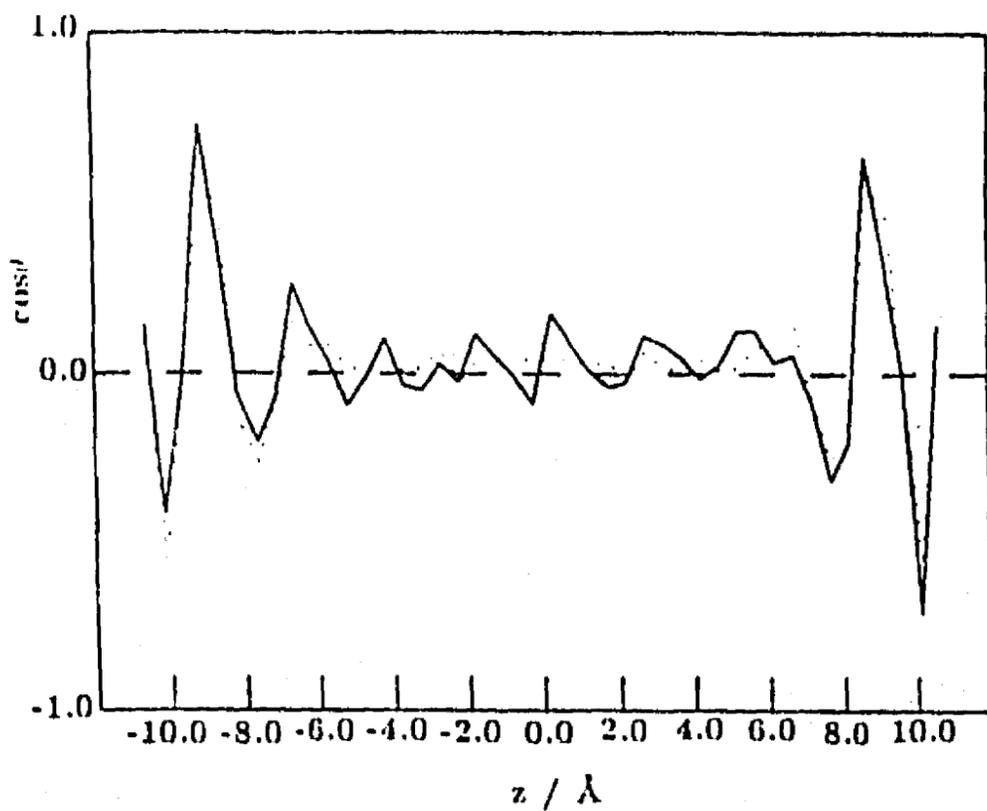


FIGURE 9. The dipole orientation of water along the  $z$ -axis for the simulation of the lamella (full line). Also shown is the symmetrized function constructed from  $[f(z) + f(-z)]/2$ , represented by the dotted line

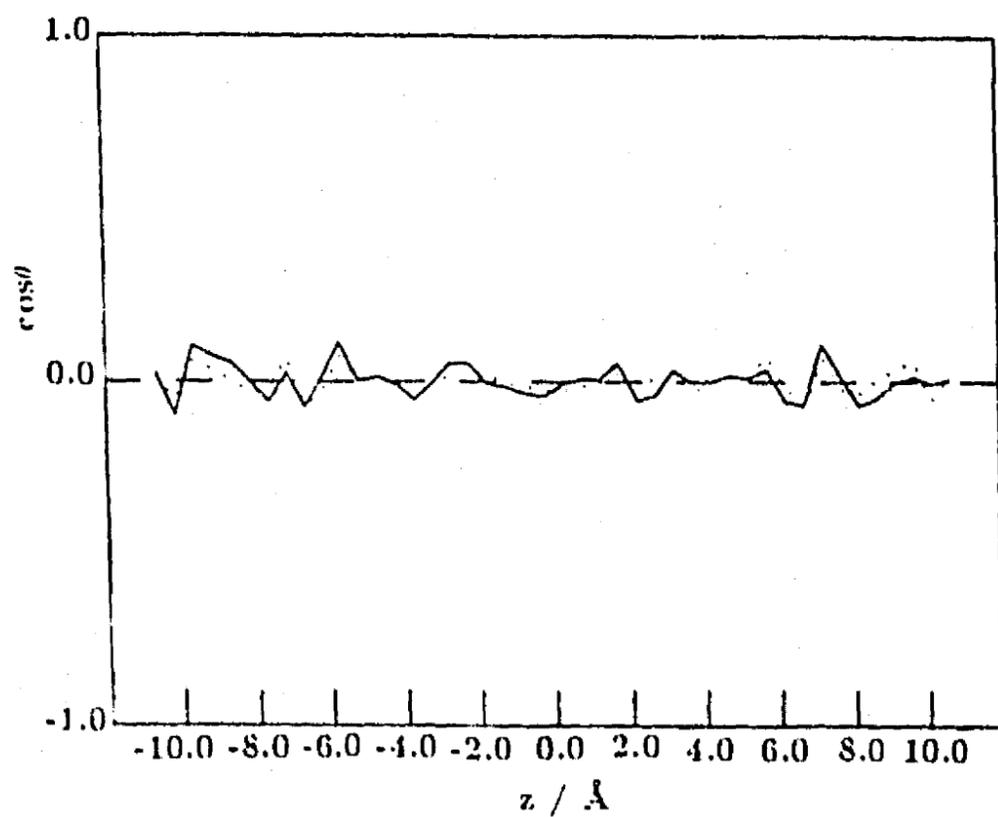


FIGURE 10. The dipole orientation of water along the  $z$ -axis for the simulation of the bulk (full line). Also shown is the symmetrized function constructed from  $[f(z) + f(-z)]/2$ , represented by the dotted line.

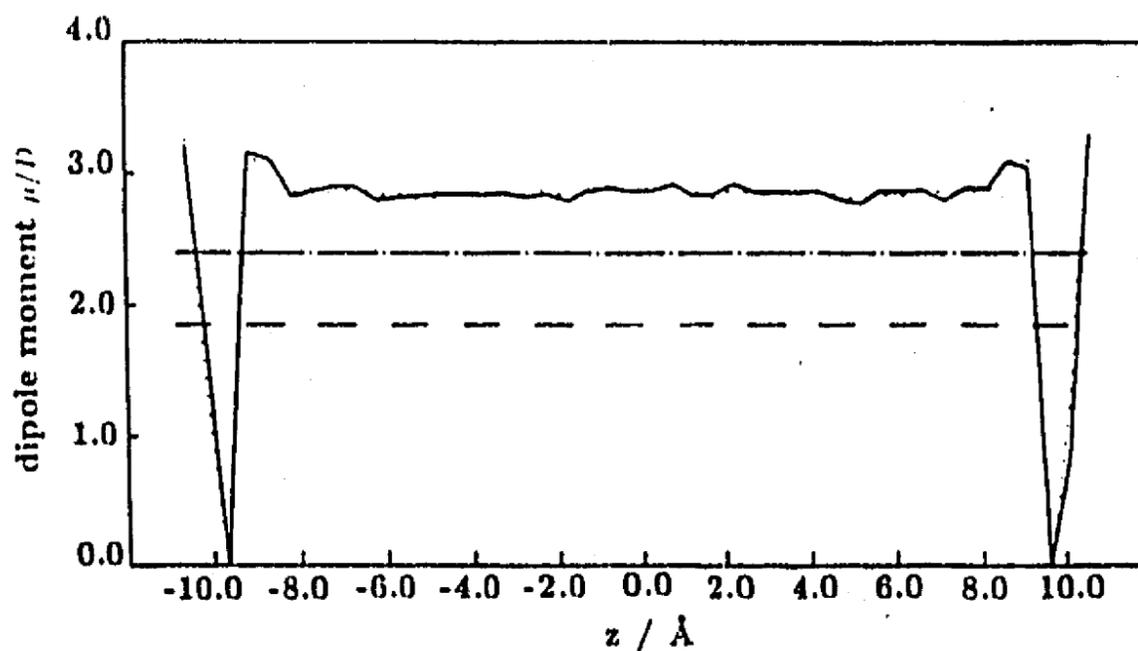


FIGURE 11. The average dipole value of water along the  $z$ -axis for the simulation of the lamella (full line). The broken line shows the dipole of the water molecule in the gas phase, and the dash-dot line shows the dipole of the MCHO model for a single water molecule. Also shown is the symmetrized function constructed from  $[f(z) + f(-z)]/2$ , represented by the dotted line.

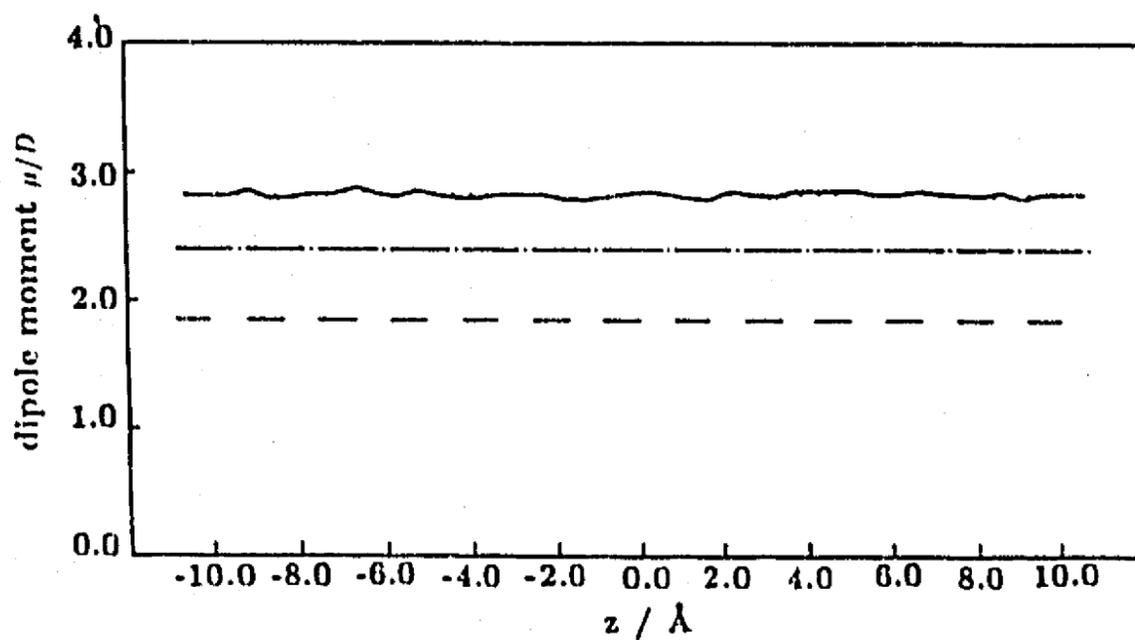


FIGURE 12. The average dipole value of water along the  $z$ -axis for the simulation of the bulk (full line). The broken line shows the dipole of the water molecule in the gas phase, and the dash-dot line shows the dipole of the MCHO model for a single water molecule. Also shown is the symmetrized function constructed from  $(f(z) + f(-z))/2$ , represented by the dotted line.

The result for the dipole value of the lamella and the bulk are presented in Figures 11 and 12, and the dipole correlations are presented in Figures 13 and 14. We can see again that there are no long-range modifications — just at a few angstroms from the wall — and there are again bulk properties.

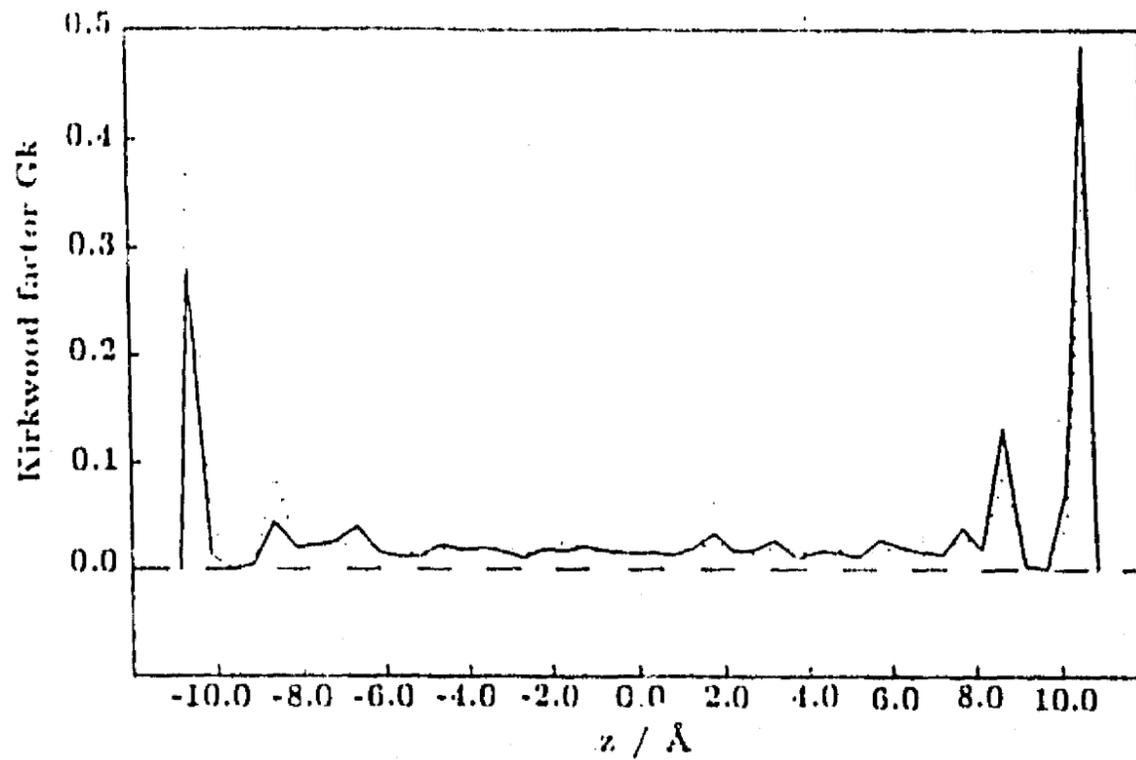


FIGURE 13. The dipole-dipole correlation of water along the  $z$ -axis for the simulation of the lamella (full line). Also shown is the symmetrized function constructed from  $[f(z) + f(-z)]/2$ , represented by the dotted line.

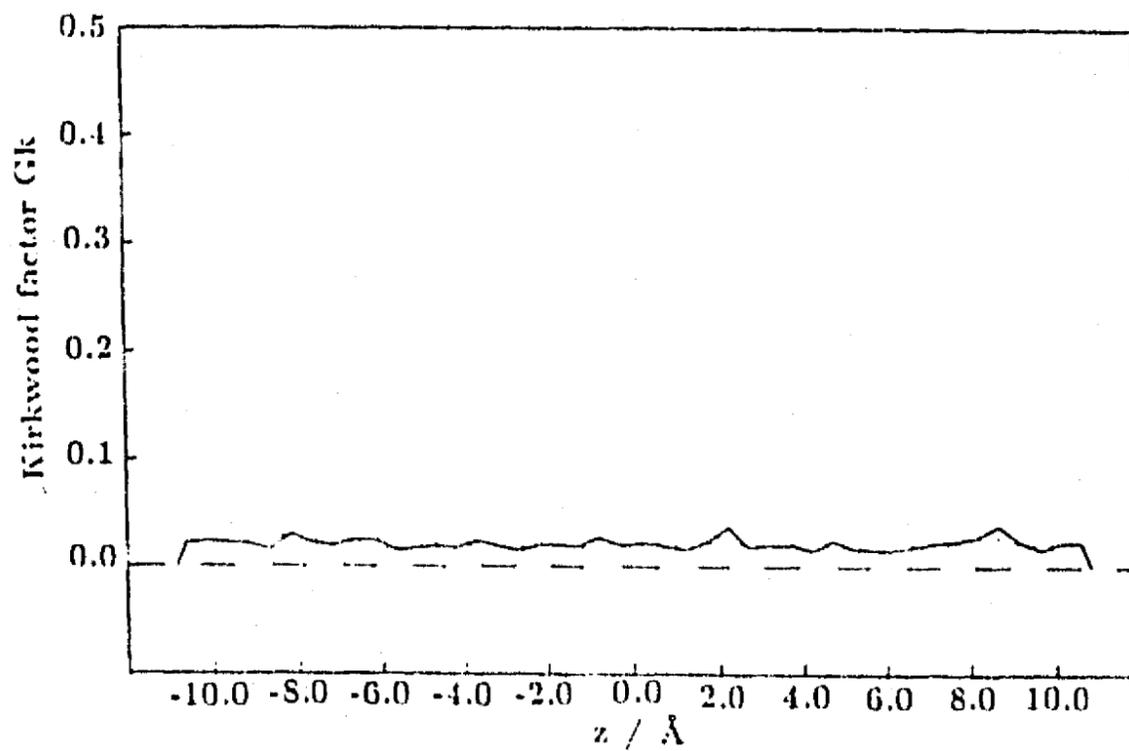


FIGURE 14. The dipole-dipole correlation of water along the  $z$ -axis for the simulation of the bulk (full line). Also shown is the symmetrized function constructed from  $[f(z) + f(-z)]/2$ , represented by the dotted line.

## VI. DISCUSSION

Here we have presented a detailed view of the efforts made by theoretical molecular methods to study the interphase between water solvent and a wall. We have seen the inherent

problems of the methodology: increasing complexity of the system vs. refinement of the intermolecular potential. We have also seen how the efforts over a decade have converged to a structural picture, but not yet to the dynamical one. The most striking and consistent result is the very short range of the structural changes produced by the wall, always less than 10 Å. Sometimes this observation goes as far as suggesting that electric perturbances as Gouy-Chapman double layer are also short ranging, a contradiction to the textbook ideas on these phenomena. The shortness of the perturbation is indeed an important matter. A decade ago, Clegg<sup>12</sup> proposed that water inside a cell was so structured, because of the long-range effect of protein surfaces and membranes, that a completely different biochemistry than that observed *in vitro* was working. Even without taking this radical position, the idea of considerable water structure and relative long-range effect of surfaces has been a prevailing one. The theoretical results suggest quite a different picture. How do they relate to experimental evidence?

In Table 3, we present a summary of some recent experimental results that can be compared to the theoretical findings. The first thing that we observe from the experimental panorama is a strong discrepancy on the water properties assumed to exist near the surface. Steytler et al.<sup>41</sup> point out that a general problem of the experimental information is that to separate the effects due to molecules close to the surface from those more distant, this has to weigh in the conclusions drawn from the experiments. One should also remember that the wall-water conditions can change, e.g., from reverse micelle to an intercontiguous phase, by environmental changes that could be produced by experimental manipulation.

From an analysis of Table 3, one can see that there is indeed experimental evidence supporting the general idea behind this review, i.e., that structural effects of the solvent are constrained to very short distances from the wall. Steytler et al.,<sup>41</sup> from the analysis of neutron diffraction studies, stress the idea that structural changes have to be restricted to few angstroms. Similar support can be obtained from Raman spectroscopy,<sup>52</sup> where it is also pointed out that structural changes could be coming from interstitial water rather than from direct surface effects. This last observation is particularly interesting, taking into account that theoretical simulations show a broad undefined boundary, i.e., a roughness of the membrane that will produce such interstices. Rheology and proton NMR<sup>53</sup> also support the idea with the finding of similar mobility of water in bulk and surface. Finally, there are several observations that in an indirect manner support the idea. There seems to be a normal pH value in reverse micelles,<sup>55</sup> i.e., one can expect a bulk dielectric constant for most of the water in the micelle. There seems to be no marked modification of enzyme structure when observed by circular dichroism.<sup>55</sup>

However, there is also a substantial interpretation of experiments along the idea of a strongly modified structure of the solvent at the interphase. Piculell et al.<sup>47-49</sup> report, after NMR studies, a strongly modified water mobility near the surface and the existence of a bound water which exchanges very slowly. None of the theoretical simulations predict anything similar since variations in water diffusion are always small and water exchange is rapid. Low mobility of water was also reported by quasielastic neutron scattering,<sup>46</sup> but it was associated with assumed huge ionic concentrations, a condition that has not been explored by simulations. Dielectric response studies<sup>51</sup> have also been used to advance a tremendous modification of the water structure inside a reverse micelle, leading to the proposal of a general ice-like phase. However, one should note that the reduction in dielectric relaxation observed in this work was also obtained in a molecular dynamics simulation<sup>18</sup> where only a small water layer suffered from structural modifications.

Experimental observations which are ascribed to the interfacial region are rather scarce, and again contradictory. Neutron diffraction suggests enhanced hydrogen bonding<sup>41</sup> whereas fluorescence picosecond studies suggest a more aprotic environment.<sup>44</sup> Unfortunately, more recent reports on these lines, to the best of our knowledge, have not appeared.

TABLE 3  
Some Experimental Work on Physical Properties of the Water-Wall Interphase

Year	System	Method	Results	Ref.
1983	Silica samples Spherisorb and Gasil	Neutron diffraction	There is an implicit problem in separating the effect due to the molecules close to surface and those more distant It is suggested, cautiously, that the interfacial region has enhanced hydrogen bonding There appears to be some local ordering of the water molecules in the interfacial region The interfacial region which is strongly hydrogen bonded and ice-like is restricted to few angstroms, $<10 \text{ \AA}$	43
1984	Reverse micelles	Fluorescence picosecond studies	More aprotic environment near the surface	44
1985	Micelles and microemulsions	Neutron small-angle scattering; differences between protonated and deuterated molecules	Intercontiguous structures of water and oil could be produced by salt	45
1985	Micelles and microemulsions	Quasielastic neutron scattering	In small water droplets, the local concentration of ion may be extremely high (12 M) This high concentration would be responsible for the decrease in the diffusion constant of water: four times less than bulk	46
1986	Protein solutions	Spin relaxation $^{17}\text{O}$ and $^2\text{H}$	There is a class of water that exchanges very slowly	47
1986	Colloidal silica	Same	A marked reduction in water diffusion	48
1986	Same	Low-frequency dispersion	Reduction of the water radial diffusivity by two orders of magnitude and of the lateral by one order	49
1986	Micelles and microemulsions	Dielectric response	Water behaves quasi-frozen in the microemulsion phase A dielectric relaxation lower than that observed in free water: 40 GHz vs. 99 GHz	50
1986	Micelles and microemulsions	Dielectric response	The temperature dependence of the dielectric response suggests an ice-like structure of water in a micelle Due to higher connectivity of the hydrogen bond, dipole relaxation decreases by six orders of magnitude, comparable to liquid crystal	51
1989	Lamellar and inverted micelles	Raman spectrum O-H bond stretching	Near multilamellar phospholipids there is a small fraction of the water molecules in crystalline structure The shape of the O-H bonds depends on supramolecular structure adopted by phospholipids rather than on the surface interactions In colloidal particles, the Raman spectrum of water is almost identical to bulk The presence of small interstices also prevents water from crystallizing, and the Raman spectra depend on the size of the water domains rather than the nature of surface It is believed that reduced long-range intermolecular vibrational coupling is the origin of the Raman change	52

TABLE 3 (continued)  
Some Experimental Work on Physical Properties of the Water-Wall Interphase

Year	System	Method	Results	Ref.
1989	Bilayer aqueous colloids	Rheology H-NMR	The mobility of water seems to be the same as that in the bulk The mobility seems to decrease with micelle formation Thus, there are two types of water, but they are not sharply defined	53
1990	AOT micelles	Thermochemistry	Water in a micelle has a freezing temperature well below zero H of freezing decreases with decreasing amount of water The unfreezable water is strongly bound to AOT and prevents ordinary ice Water in the micelle is present mostly in a bound form: hydration water Solubility drastically decreases with increasing amount of water Enzyme activity increases with decreasing amount of water Thermal stability of the protein increases with decreasing amount of water	54
1991	Reversed micelles with enzyme	<sup>31</sup> P-NMR circular dichroism enzyme activity	The ionization behavior of the phosphate groups is similar to that in bulk; no anomaly of the micellar pH No marked perturbation on the enzyme structure by micellar water The physical properties of water in micelles of AOT or lecithin are very similar, suggesting that special properties are more determined by physical constraint than by chemical composition	55

## VII. CONCLUSIONS

A combined theoretical and experimental effort during the last decade in determining the properties of water at an interphase has advanced our understanding of this crucial problem. Even if there is no complete agreement, we feel that the fact that water attains the bulk properties, or very similar ones, after a few angstroms is emerging. The only exception observed so far, and with the most refined potentials, is wetting, which has shown a longer range of action than the dimensions of the boxes used in the simulations — just above 10 Å. However, this leaves some critical observations as open questions. Why does water in a reversed micelle have a low freezing point?<sup>54</sup> Why are solubility and enzyme thermal stability affected?<sup>54</sup> And mainly, why are catalytic reactions so much affected?

One could think that water crystallization is hindered by a local surface structure that does not allow for the ice topography; this is quite possible, mainly through the idea of interstitial water. One could think that a physical phenomenon such as wetting could conduce to local ionic distributions that would change the microscopic picture so far obtained. Further theoretical and experimental work is needed and will certainly help us to understand an important and intriguing problem.

## REFERENCES

1. Luisi, P. L. and Straub, B. E., Eds., *Reverse Micelles, Biological and Technological Relevance of Amphiphilic Structures in Apolar Media*. Plenum Press, New York, 1984.
2. Degrogio, V. and Corti, M., Eds., *Physics of Amphiphiles — Micelles, Vesicles and Microemulsions*. North-Holland, Amsterdam, 1985.
3. Sauer, J., Molecular models in *ab initio* studies of solids and surfaces: from ionic crystals and semiconductors to catalysis, *Chem. Rev.*, 89, 199, 1989.
4. Binder, K., Ed., *Topics in Current Physics, Application of the Monte Carlo Method in Statistical Physics*. Springer-Verlag, Heidelberg, 1984.
5. Ciccoli, G., Frenkel, D., and McDonald, I. R., *Simulation of Liquids and Solids, Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics*. North-Holland, Amsterdam, 1987.
6. Gil-Adalid, L. and Ortega-Blake, I., Four body nonadditivity in liquid water, *J. Chem. Phys.*, 94, 3748, 1991.
7. Saint-Martin, H., Medina-Ilanos, C., and Ortega-Blake, I., Nonadditivity in an analytical intermolecular potential: the water-water interaction, *J. Chem. Phys.*, 93, 6448, 1990.
8. Niesar, U., Corongiu, G., Clementi, E., Kneller, G. R., and Battachayra, D. K., Molecular dynamics simulation of liquid water using the NCC *ab-initio* potential, *J. Phys. Chem.*, 94, 7949, 1990.
9. Weast, R. C. and Astle, M. J., Eds., *CRC Handbook of Chemistry and Physics*, 63rd ed., CRC Press, Boca Raton, FL, 1983.
10. Narten, A. H. and Levy, H. A., Liquid water: molecular correlation functions from X-ray diffraction, *J. Chem. Phys.*, 55, 2263, 1971.
11. Super, A. K. and Phillips, M. G., A new determination of the structure of water at 25°C, *Chem. Phys. Letts.*, 107, 47, 1986.
12. Watanabe, K. and Klein, M., Effective pair potentials and the properties of water, *Chem. Phys.*, 131, 157, 1989.
13. Cieplak, P. and Kollman, P., A new water potential including polarization: application to gas-phase, liquid and crystal properties of water, *J. Chem. Phys.*, 92, 6755, 1990.
14. Narten, A. H., Liquid water: atom pair correlation function from neutron and X-ray diffraction, *J. Chem. Phys.*, 56, 5081, 1972.
15. Levitt, M., Molecular Dynamics of macromolecules in water, *Chemica Scr.*, 29A, 197, 1989.
16. Lie, G. C. and Clementi, E., Molecular Dynamics simulation of liquid water with an *ab-initio* flexible water-water interaction potential, *Phys. Rev. A*, 33, 2679, 1986.
17. Jönsson, B., Monte Carlo simulations of liquid water between two rigid walls, *Chem. Phys. Letts.*, 82, 520, 1981.
18. Marchesi, M., Molecular Dynamics simulations of liquid water between two walls, *Chem. Phys. Letts.*, 97, 224, 1983.
19. Sonnenschein, R. and Heinzinger, K., A molecular dynamics study of water between Lennard-Jones walls, *Chem. Phys. Letts.*, 102, 550, 1983.
20. Lee, C. Y., McCammon, J. A., and Rossky, P. J., The structure of liquid water at an extended hydrophobic surface, *J. Chem. Phys.*, 80, 4448, 1984.
21. Belch, A. and Berkowitz, M., Molecular Dynamics simulations of TIP2S water restricted by a spherical hydrophobic boundary, *Chem. Phys. Letts.*, 113, 278, 1985.
22. Jönsson, B., Edholm, O., and Teleman, O., Molecular Dynamics simulation of a sodium octanoate micelle in aqueous solution, *J. Chem. Phys.*, 85, 2259, 1986.
23. Bratko, D., Jönsson, B., and Wenneström, H., Electrical double layer interactions with image charges, *Chem. Phys. Letts.*, 128, 449, 1986.
24. Spuhr, E. and Heinzinger, K., A molecular dynamics study of an aqueous LiI solution between Lennard-Jones walls, *J. Chem. Phys.*, 84, 2304, 1986.
25. Valleau, J. P. and Gardner, A. A., Water-like particles at surfaces. I. The uncharged unpolarizable surface, *J. Chem. Phys.*, 86, 4162, 1987.
26. Gardner, A. A. and Valleau, J. P., Water-like particles at surfaces. II. In a double layer and at a metallic surface, *J. Chem. Phys.*, 86, 4171, 1987.
27. Watanabe, K., Ferrario, M., and Klein, M., Molecular Dynamics study of sodium octanoate micelles in aqueous solution, *J. Phys. Chem.*, 92, 819, 1988.
28. Egherts, E. and Berendsen, H. J. C., Molecular Dynamics simulation of a smectic liquid crystal with atomic detail, *J. Chem. Phys.*, 89, 1718, 1988.
29. Patey, G. N. and Torrie, G. M., Water and salt water near a charged surface: a discussion of some recent theoretical studies, *Chemica Scr.*, 29A, 39, 1989.
30. Rossky, P. J. and Lee, S. H., Structure and dynamics of water at interfaces, *Chemica Scr.*, 29A, 93, 1989.

31. Linse, P., Molecular Dynamics study of the aqueous core of a reversed micelle. *J. Chem. Phys.*, 96, 4992, 1989.
32. Smit, B., Hilbers, P. A. J., Esselenk, K., Rupert, L. A. M., Vanos, N. M., and Schlijper, A., Computer simulations of a water interface in the presence of micelles, *Nature*, 348, 624, 1990.
33. Wallqvist, A., Polarizable water at a hydrophobic wall, *Chem. Phys. Lett.*, 165, 437, 1990.
34. Zhu, S. B. and Robinson, G. W., Structure and dynamics of liquid water between plates, *J. Chem. Phys.*, 94, 1403, 1991.
35. Makovsky, N. N., Structure of the surfactant water interface, *Mol. Phys.*, 72, 235, 1991.
36. Matsuoka, O., Clementi, E., and Yushimine, M., Water dimer potential surface, *J. Chem. Phys.*, 64, 1351, 1976.
37. Stillinger, F. H. and Ramun, A., Improved simulation of liquid water by molecular dynamics, *J. Chem. Phys.*, 60, 1545, 1974.
38. Jönsson, B., Wenneström, H., and Halle, B., Ion distribution in lamellar liquid crystals. A comparison between results from Monte Carlo simulations and solutions of the Poisson-Boltzmann equation, *J. Phys. Chem.*, 84, 2179, 1980.
39. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J., Interaction models for water in relation to protein hydration, in *Intermolecular Forces*, Pullman, B., Ed., Reidel Dordrecht, 1981, 331.
40. Jörgenson, W. L., Chandrasekar, J., Madura, J. D., Impey, R. W., and Klein, M. L., Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 79, 926, 1983.
41. Wallqvist, A. and Karlström, G., A new non empirical force field for computer simulations, *Chemica Scr.*, 29A, 131, 1989.
42. Clegg, J. S., Alternative views on the role of water in cell function, in *Biophysics of Water*, Franks, F. and Mathias, S., Eds., John Wiley & Sons, Chichester, 1982.
43. Steytler, D. C., Dore, J. C., and Wright, C. J., Neutron diffraction studies of the water in meso and micropores, *Mol. Phys.*, 48, 1031, 1983.
44. Rodgers, M. A. J., Picosecond studies of rose bengal fluorescence in reverse micellar systems, in *Reverse Micelles. Biological and Technological Relevance of Amphiphilic Structures in Apolar Media*, Luisi, P. L. and Straub, B. E., Eds., Plenum Press, New York, 1984, 165.
45. De Geyer, A. and Tabony, J., Evidence for intercontiguous structures in concentrated microemulsions, neutron small angle scattering results, *Chem. Phys. Lett.*, 113, 83, 1985.
46. Tabony, J., Quasi elastic scattering measurements of molecular motions in micelles and microemulsions, *Chem. Phys. Lett.*, 113, 75, 1985.
47. Piculell, L., Water spin relaxation in colloidal systems. I.  $^{17}\text{O}$  and  $^2\text{H}$  relaxation in dispersions of colloidal silica, *J. Chem. Soc. Faraday Trans.*, 82, 387, 1986.
48. Piculell, L. and Halle, B., Water spin relaxation in colloidal systems. II.  $^{17}\text{O}$  and  $^2\text{H}$  relaxation in protein solutions, *J. Chem. Soc. Faraday Trans.*, 82, 401, 1986.
49. Halle, B. and Piculell, L., Water spin relaxation in colloidal systems. III. Interpretation of the low frequency dispersion, *J. Chem. Soc. Faraday Trans.*, 82, 415, 1986.
50. Nimitz, G., Blaggeli, B., Börngen, L., Marquardt, P., Schäffer, U., and Zorn, R., Dielectric and structural properties of a water-oil emulsion at the gel microemulsion transition, *Europhys. Lett.*, 2, 103, 1986.
51. Marquardt, P. and Nimitz, G., Critical dielectric behaviour of micellar water below the cloud point of a water-oil microemulsion, *Phys. Rev. Lett.*, 8, 1036, 1986.
52. Lafleur, M., Pigeon, M., Pérolet, M., and Caillé, J. P., Raman spectrum of interstitial water in biological systems, *J. Phys. Chem.*, 93, 1522, 1989.
53. Matsumoto, T., Ito, D., and Kohno, H., Interaction between amphiphiles and water molecules in concentrated bilayers aqueous colloids, *Colloid Polym. Sci.*, 267, 946, 1989.
54. Luisi, P. L., Häring, G., Maestro, M., and Rialdi, G., Proteins solubilized in organic solvents via reverse micelles: thermodynamic studies, *Thermochim. Acta*, 162, 1, 1990.
55. Peng, Q. and Luisi, P. L., The behaviour of proteases in lecithin reverse micelles, *Eur. J. Biochem.*, 188, 471, 1990.

BBAPRO 34033

## Ab initio calculations of the pyrophosphate hydrolysis reaction

Humberto Saint-Martin<sup>1</sup>, Iván Ortega-Blake<sup>1</sup>, Andrzej Leś<sup>2,\*</sup>  
and Ludwik Adamowicz<sup>2</sup>

<sup>1</sup> Instituto de Física, UNAM, Laboratorio de Cuernavaca, Cuernavaca, Morelos (Mexico)  
and <sup>2</sup> Department of Chemistry, University of Arizona Tucson, AZ (U.S.A.)

(Received 3 June 1991)

Key words: Pyrophosphate hydrolysis; Gas-phase reaction; Ab initio calculation; Enzyme mechanism

Ab initio quantum mechanical calculations were used to study the hydrolysis reaction  $H_4P_2O_7 + H_2O \rightleftharpoons 2H_2PO_4^-$ , as well as some molecular properties of the reactants and products. SCF calculations with several basis sets ranging from minimal to extended with polarization functions were used to look at the basis dependency of the reaction enthalpies and optimized geometries. Although the minimal basis sets yield erratic predictions of the enthalpy, when a more extended basis (3-21G\*) was used for the geometry optimization, and the total energies of the reactants and products were computed with this and larger basis sets, we obtained more consistent predictions of the structural properties of the P-O-P bridge and of the heat of the hydrolysis reaction ( $\Delta E = -7.39$  kcal/mol at the SCF/6-31G\*\* level). A comparison is made with previous estimates performed with smaller basis sets and without taking into account the electron correlation effects, which are calculated in the present work. The inclusion of the zero point energy calculated using the harmonic approximation, and of the electronic correlation energy determined at the MBPT(2) level, raised the computed heat of the reaction to  $-3.83$  kcal/mol, and when an estimate for the thermal energy was added, the value obtained was of  $-3.38$  kcal/mol. In conclusion, we found that the hydrolysis of pyrophosphate should be exothermic in the gas phase. The implications of this result in relation to some recent theories about enzyme catalysis are discussed.

### Introduction

Different forms of energy are interconverted in biological systems. The most common medium for energy exchange in cells are phosphate compounds. The exothermic process of the hydrolysis, which leads to the energy release, is catalyzed by enzymes. The chemical energy derived from the cleavage of the P-O-P linkage is used by the cells to perform various biological functions. For a long time, the energetic capabilities of the so called 'high-energy' phosphate compounds were considered to be due to purely intramolecular effects, such as opposing resonance, electrostatic repulsions and electron distribution along the P-O-P backbone

[1-4]. Recently, the general view about the enzyme catalyzed reactions has changed towards the consideration of the interactions between the reactants, the solvent and the enzymes. The idea that enzymes work by the principle of transition-state stabilization [5] is gradually becoming accepted; this principle asserts that an enzyme binds the reactants in their transition state with a higher affinity than it binds them in their ground state. The way in which this is done is not clear. A proposal advanced by Dewar is the *desolvation hypothesis* [6], the basic idea being that the enzyme catalyzed reactions are mainly *solvent-activated*, i.e., the activation energy barriers are due entirely to the solvation of the reactants; thus, the role of the enzyme is to 'squeeze out' the water surrounding the reactants from the active site, allowing the substrates to interact in a gas-phase-like environment. This hypothesis has been subject to criticism because it overlooks the fact that at physiological pH most of the organic molecules are deprotonated, and a substantial amount of energy would be needed to bring ions to the gas phase [7].

George et al. [8] proposed that for the high-energy phosphate compounds the reaction energy was due not

\* On leave from the Quantum Chemistry Laboratory, Department of Chemistry, University of Warsaw, Pasteura 1, PL-02-098 Warsaw, Poland.

Correspondence: H. Saint-Martin, Instituto de Física, UNAM, Laboratorio de Cuernavaca, Apartado Postal 139 - B, Cuernavaca, Morelos 62191, Mexico.

only to purely intramolecular effects, but that the difference in the hydration energies of reactants and products had an important contribution to the enhanced reactivity of the phosphate anhydrides over simple orthophosphate esters. They estimated the hydration energies of several molecules and speculated about the possible thermoneutrality of some reactions in the gas phase. There are no experimental data for the gas phase, but some recent experiments in which water is substituted by aprotic organic solvents [9-13] seem to support the idea that the contribution of solvation is highly relevant (see Discussion below).

Estimates for the heat of the pyrophosphate hydrolysis, based on *ab initio* calculations have already been reported [14-16] with contradictory conclusions. In this work we review these results and perform new calculations, to study the hydrolysis reaction



We examine the dependence of the predicted heat of reaction (1) on the basis set quality and arrive at a quite well converged value. Also, the zero point energy and the correlation energy contributions to the heat of the hydrolysis are calculated.

#### Computational details

Geometry optimizations were carried out with the GAUSSIAN-86 [17] program for the molecules  $\text{H}_2\text{O}$ ,  $\text{H}_3\text{PO}_4$  and  $\text{H}_2\text{P}_2\text{O}_7$ . Optimizations were also performed with the MNDO method included in the MOPAC [18] package. All single-point energy calculations were done with the GAUSSIAN-88 [19] program, using five different basis sets: STO-3G\* (basis I), 3-21G\* (basis II), and 6-31G\* (basis III), all of these with d functions only on the P atoms; a 6-31G\* extended with d polarization functions on the O atoms (basis IV), and another one (basis V), which adds to the latter p functions on the H atoms. We also made computations using the pseudopotentials of Barthelat et al. [20] included in the MONSTERGAUSS [21] program with a minimal basis set plus d functions on P (basis VI) and a double zeta basis plus d functions on P (basis VII). The basis VI has a similar quality to basis I, and basis VII is equally related to basis III.

#### Computed structures

(a)  $\text{H}_3\text{PO}_4$  (Fig. 1). Geometry optimizations of this molecule have been performed by Hayes et al. [14], and more recently by Ewig and Van Wazer [22] and by O'Keeffe et al. [15].

Ewig and Van Wazer [22] performed a full optimization with a basis equivalent to our basis I and a  $C_3$  symmetry was attained (structure PII), whereas

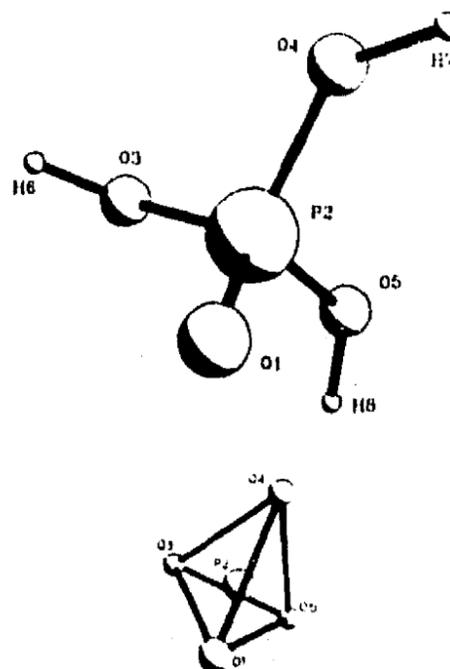


Fig. 1. Tetrahedral structure of the  $\text{H}_3\text{PO}_4$  molecule (geometry PIII).

O'Keeffe et al. [15] employed a basis equivalent to our basis IV and constrained the molecule to the  $C_3$  symmetry around the P=O bond in their geometry optimizations (structure PIV). In the present work, the structure PI was obtained with the MNDO method and the PIII was generated with the SCF method using basis II. Both structures were constrained to the  $C_3$  symmetry around the P=O bond. A comparison of the four geometries mentioned above is presented in the Appendix. The similarities in all the parameters obtained suggest that the  $\text{H}_3\text{PO}_4$  structure has been quite well determined. To test this conclusion we released all the constraints of geometry PIII and performed a reoptimization. Once again, the  $C_3$  symmetry was attained (geometry PV), and the structural parameters remained practically the same (see Appendix). The comparison of *ab initio* optimized geometries for molecules containing P-O bonds with crystallographic data has been discussed earlier [15] and good agreement was found with the  $\text{P}_2\text{O}_5$  crystal structure, whereas for the  $\text{P}_2\text{O}_5$ -III and  $\text{H}_3\text{PO}_4$  crystals the lengths of the double bonds P=O are somewhat larger than those of the SCF optimization. Worth mentioning is that for the  $\text{H}_3\text{PO}_4$  crystal the situation is complicated by strong intermolecular hydrogen bondings, which do not occur in the dilute gas.

The SCF energies corresponding to the PI, PII and PIII geometries were computed with basis sets I, II and IV in all-electron calculations, and with basis VI and VII, utilizing the pseudopotentials (see Appendix). It is interesting to note that although the PII structure has

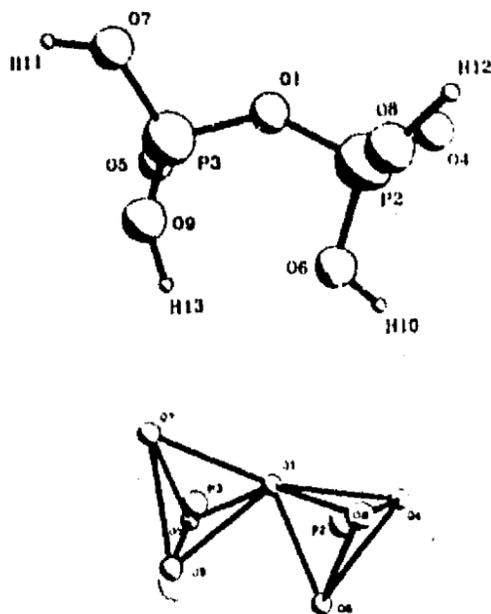


Fig. 2. Structure of the  $H_4P_2O_7$  molecule (geometry PPIII), showing its resemblance to two phosphoric acid molecules conjoined by a 'ball-and-socket' joint.

the lowest energy with basis I, its energy is higher for all other basis sets. This result indicates that the structures optimized with minimal basis sets should be used with caution, as also stated by O'Keeffe et al. [15]. To see the effect of the geometry when a larger basis set is used, we computed the energy of the PV structure with basis V, and it resulted 0.02 kcal/mol higher than that computed for the PIII geometry. This strengthens the conclusion that the PIII structure is quite reliable.

(b)  $H_4P_2O_7$  (Fig. 2). Quantum calculations on this molecule were first performed by Hayes et al. [14] using three different basis sets, and, in more recent studies, by O'Keeffe et al. [15] and by Ewig and Van Wazer [16].

O'Keeffe et al. [15] used a 6-31G\* basis set and obtained two optimal geometries, one staggered and one eclipsed (which will be referred to as PPIV). The energy of the former is 1.8 kcal/mol lower than that of the latter. They also studied the P-O-P bridge using a basis augmented with additional polarization functions on the bridging O atom. Surprisingly enough, the P-O-P angles obtained with such large basis are far from the average experimental value [15] of  $133^\circ$ .

Ewig and Van Wazer [16] employed a STO-3G\* basis set and also got two optimal geometries (which will be referred to as PPII), one staggered and the other eclipsed, the energy of the former being 2.3 kcal/mol lower than that of the latter. Their staggered conformation was optimized with almost no restrictions, whereas for the eclipsed structure they constrained the molecule to the  $C_2$  symmetry. The P-O-P

angle of the eclipsed geometry is closer to the experimental value than that of the staggered conformation.

In this work we performed the optimization for pyrophosphate with basis III, starting from the PPII eclipsed geometry. Some imaginary vibrational frequencies resulted at first, showing that some initially imposed constraints should be released. After the subsequent optimization with some constraints still remaining, a staggered structure was found (referred to as PPIII), and there were no more imaginary frequencies. In total, 16 out of 33 degrees of freedom were varied. The structural features of the PPIII geometry are very similar to those of the PPII staggered geometry, including that the phosphoryl group is staggered to the other phosphoryl and to POH, in contrast with the PPIV configuration, which has the phosphoryl groups staggered to both POH groups. As it was observed by Ewig and Van Wazer [16], diphosphoric acid closely resembles a pair of conjoined phosphoric acid molecules.

Although the basis set III does not comprise polarization functions on the bridging oxygen, the P-O-P angle is very reasonable, as well as the whole structure, corroborating Ewig and Van Wazer's observation that an improvement in the geometry obtained is not due to the inclusion of polarization functions.

At the SCF level using basis IV, we found a small energy difference between both PPII structures and the PPIII structure, in agreement with the results of O'Keeffe et al. [15] for their staggered and eclipsed conformations. Hence we decided to take a look at the rotational barrier around the P-O-P backbone. We constrained the eclipsed conformation to the  $C_2$  symmetry, and we forced the phosphoryl oxygens to be in an opposite position to that found in the PPII eclipsed structure. The obtained geometry is referred to as PPI. The differences in SCF energies obtained suggest that the minimum with respect to this degree of freedom is shallow. We also found that with basis sets II, VI and VII the PPI structure has a lower SCF energy than both PPII structures; however, with the larger basis set IV the situation is reversed.

Since in our optimization we varied only half of the parameters, we decided to go a step further, releasing all the constraints, and found the geometry PPV, whose structural parameters resulted very similar to those of the PPIII geometry (see Appendix), and yielded an energy 2.3 kcal/mol lower at the SCF level with basis II. However, with the larger basis V the SCF energy of the geometry PPV is 1.1 kcal/mol higher than that of PPIII. On behalf of this result, we can assess the reliability of the latter structure when large basis sets are used.

To investigate further the flexibility of the P-O-P angle, that seems to be the most important internal degree of freedom of the  $H_4P_2O_7$  molecule, we per-

formed calculations with basis V for the PPIII geometry with the mentioned angle being varied by  $-10^\circ$ ,  $-5^\circ$  and  $5^\circ$ . At the SCF level the PPIII optimized angle was found to correspond to the energy minimum. Then the MBPT(2) electronic correlation energies were computed for all the above structures. The results presented in Table I show that the PPIII angle is nearly optimal, within  $3^\circ$  at this level of the theory. Moreover, with the MBPT(2) contribution the P-O-P optimal angle lies very close to the experimental value [15] of  $133^\circ$ . To get an estimate for the force constant of the P-O-P bent,  $K_\theta = (1/r^2)d^2E/d\theta^2$ , we fitted a second order polynomial to the energies in Table I, and by differentiating it twice we obtained  $K_\theta = 36.6 \text{ Nm}^{-1}$  when considering only the SCF level, and  $K_\theta = 30.5 \text{ Nm}^{-1}$  at the MBPT(2) level. In effect, the electronic correlation was found to reduce the SCF P-O-P bent force constant by 17%.

The force constants related to the P-O-P angle were computed by O'Keeffe et al. [15] with an STO-3G\* basis set, and a value of  $K_\theta = 18.8 \text{ Nm}^{-1}$  was obtained. Our result is twice as large, but still in the same range, supporting the reliability of their results and hence their remark about a strong bond length-bond angle correlation.

#### The reaction energy of the pyrophosphate hydrolysis

The most interesting point to be discussed is that of the heat of the hydrolysis reaction of pyrophosphate. In the present case, it is of paramount importance to find out whether or not the  $\Delta E$  is negative, and to assert upper and lower bounds, since some recent theories about the high-energy of phosphates require it to be practically zero or even positive [6,8]. The ab initio results reported up to date are all negative, although Ewig and Van Wazer's [16] result is of only  $-0.42 \text{ kcal/mol}$ , when they include an estimate of the thermal contribution.

TABLE I

Energies computed for the  $\text{H}_2\text{P}_2\text{O}_7$  and various POP bridging angles  $\phi$ , close to the SCF/3-21G\* optimized value  $\phi_0 = 138.96^\circ$

The first column shows the variation of the angle  $\phi$ , the second presents the differences in SCF/6-31G\*\* energy, relative to the value corresponding to  $\phi_0 = 1208.037309 \text{ a.u.}$ , and the third, the contribution of the electronic correlation obtained with the MBPT(2)/6-31G\*\* method (reference value  $-1.475901 \text{ a.u.}$ ). The last column shows the cumulative (SCF+MBPT(2)) relative energy. All energies are given in kcal/mol.

$\phi$	SCF	MBPT(2)	SUM
$\phi_0 - 10^\circ$	2.15	-1.11	1.04
$\phi_0 - 5^\circ$	0.42	-0.46	-0.04
$\phi_0$	0.00	0.00	0.00
$\phi_0 + 5^\circ$	0.27	0.32	0.59

TABLE II

The heat of the hydrolysis reaction (I) obtained with the SCF method using different geometries and computed with different basis sets

The structures used for pyrophosphate and orthophosphate are described in the text. The geometries used for the water molecule were: (a) the experimental structure and (b) the one optimized with the corresponding basis set. Energies are given in kcal/mol.

(A)

Geo- metry	Basis sets *					pseudo- potentials	
	all-electron					VI	VII
	I	II	III	IV	V		
I	2.36	-6.72		-11.86		-1.62	-3.09
II	4.22	-6.36		-11.72			-2.94
III	-2.63	-1.99		-4.51		8.35	-1.04
IV	-0.76	-1.63		-4.37			-0.60
V	-4.89	-2.84		-5.21		6.50	-0.91
VI	-3.03	-2.48		-5.08			-0.47
VII	0.71	-3.22	-6.21	-6.60	-7.67	1.97	-5.98
VIII	2.57	-2.86	-5.35	-6.46	-7.38		-5.54
IX		-0.92			-8.71		
X		-0.56			-8.43		

(B) The geometries used for the different molecules are the following:

Geometry	$\text{H}_2\text{P}_2\text{O}_7$	$\text{H}_3\text{PO}_4$	$\text{H}_2\text{O}$
I	PPI	PI	(a)
II	PPJ	PI	(b)
III	PPII staggered	PII	(a)
IV	PPII staggered	PII	(b)
V	PPII eclipsed	PII	(a)
VI	PPII eclipsed	PII	(b)
VII	PPIII	PIII	(a)
VIII	PPIII	PIII	(b)
IX	PPV	PV	(a)
X	PPV	PV	(b)

\* As described in the text.

In this work we computed the energy differences between reactants and products at the SCF level, using the five different basis sets already mentioned, and four different geometries for the  $\text{H}_2\text{P}_2\text{O}_7$  molecule (PPI, both PPII and PPIII). The results are presented in Table II, where it can be seen that the energies computed with minimal basis sets have an erratic behaviour, allowing for both negative and positive values. On the contrary, when larger basis sets are used, the  $\Delta E^{\text{SCF}}$  is definitely negative, with a value which is somewhat independent from the  $\text{H}_2\text{P}_2\text{O}_7$  conformation as long as its most important degrees of freedom are nearly optimal (the PPI geometry presented is clearly inadequate, since the P-O-P bridge is far from optimal and the phosphate moieties are forced to lie in unfavourable positions).

Using basis set V we computed the  $\Delta E^{\text{SCF}}$  for the PIII and PPIII geometries (see Table II) and the same trend of negative values was observed. Our best esti-

mate at the SCF level for the hydrolysis enthalpy is, therefore,  $-7.39$  kcal/mol ( $-7.67$  kcal/mol when using the experimental geometry for water). Next we computed the MBPT(2) correlation energies, which gave  $\Delta E^{\text{MBPT}(2)}$  of  $2.25$  kcal/mol ( $2.96$  kcal/mol, respectively), whence our estimated heat for the hydrolysis reaction becomes  $\Delta E = \Delta E^{\text{SCF}} + \Delta E^{\text{MBPT}(2)} = -5.14$  kcal/mol (and  $-4.71$  kcal/mol, respectively). Also, an estimate for the contribution of the zero point energy was made, using the harmonic vibrational frequencies computed with basis II. The  $\Delta E^{\text{ZPE}}$  obtained is  $1.31$  kcal/mol, that added to the former values yields  $\Delta E = -3.83$  kcal/mol ( $-3.40$  kcal/mol). The correlation contribution alone amounts to roughly 30% (40%) of the SCF energy, and when combined with the ZPE the percentage grows to 50% (55%); yet, the value is still negative and, gauging for the trend shown in Table II, it seems to constitute a well converged estimate for the heat of the reaction at  $T = 0^\circ$  K. We further considered an estimate of the thermal contribution at a temperature of  $T = 298^\circ$  K, by means of an approximation to the molecular partition function of each species, using the harmonic frequencies related to the degrees of freedom of the reactants and the products. Since we are dealing with isolated molecules and we are not considering the interactions between them, this amounts to a somewhat crude estimate of the entropic term  $T\Delta S$  (see e.g. [23]). The value obtained is of only  $0.45$  kcal/mol, which shows that not only the enthalpy  $\Delta H_{298}$ , but also the free energy  $\Delta G$  is negative for the hydrolysis reaction (1) in the (ideal) gas phase at room temperature, with a value of  $-3.38$  kcal/mol.

To get an estimate of the dependence of the results on the geometry employed, we repeated the same calculations with geometries PV and PPV. A comparison is presented in Table III. It can be seen that the difference in energy is of  $\approx 1$  kcal/mol at the SCF level, and decreases to  $\approx 0.5$  kcal/mol when the MBPT(2) and ZPE contributions are included.

TABLE III

Comparison of the energies of the hydrolysis reaction (1) computed with (A) geometries PIII and PPIII, and (B) geometries PV and PPV, and either (a) the experimental geometry of water, or (b) that optimized with basis I.

Energies are given in kcal/mol.

	A		B	
	a	b	a	b
SCF, V	-7.67	-7.39	-8.71	-8.43
MBPT(2), V	2.96	2.25	4.51	3.80
ZPE/II	1.31	1.31	0.24	0.24
SUM	-3.40	-3.83	-3.96	-4.39

The trend shown in Table II as more extended basis sets are used points towards more negative values. Although even basis V is not saturated yet, the convergence of the results can be seen by taking the energy difference for successive columns in Table II: using the values for the eighth row (geometries PIII and PPIII), this difference is of  $-2.49$  kcal/mol,  $-1.11$  kcal/mol and  $-0.92$  kcal/mol, when going from column II to columns III, IV, and V, respectively. Obviously, it diminishes steadily (see Note added in proof).

#### Discussion and Conclusions

As far as the molecular structure is concerned, the results here presented are in good agreement with those previously reported [15,16,22]. The main conclusions that can be drawn about  $\text{H}_3\text{PO}_4$  are that it has a very well defined geometry, with the O atoms in a tetrahedral arrangement around the central P atom. The P=O bond length corresponds to a double bond, whereas the other three O atoms are single bonded to P. This structure is preserved in  $\text{H}_2\text{P}_2\text{O}_7$ , which closely resembles a pair of phosphoric acid molecules joined by the P-O-P bridge. This bridge has a relatively large bent force constant, and a significant bond length-bond angle correlation [15]. We also found that when considering rotations around the bridge, there are several energetically equivalent positions (within 2.0 kcal/mol, using basis IV). This kind of structure of two rigid moieties bound by a flexible 'ball-and-socket' joint explains the wide range of structures found in crystals and certainly becomes an important feature to be considered in envisaging the molecular mechanisms of the enzyme catalyzed reaction.

The main goal of this work was to give a reliable estimate of the heat of the pyrophosphate hydrolysis reaction in the gas phase. This is an *isodesmic* [24] reaction (i.e. the number of bonds of each type does not change), which allows its enthalpy to be computed at the SCF level with an accuracy of some  $\pm 5$  kcal/mol, depending on the basis set used. This uncertainty is intrinsic to the HF-SCF approximation, and is due to the neglect of electron correlation effects [24,25]. In addition, the estimates obtained for reaction (1) with minimal basis sets show a critical dependence on the geometry of the molecules considered. These facts lead us to conclude that the results computed at this level of the theory by Ewig and Van Wazer [16] cannot discriminate the exothermicity of the reaction. O'Keefe et al. [15] employed a substantially larger basis set and got a more negative heat of reaction (1),  $\Delta E = -5.50$  kcal/mol. In Table II it can be seen that we get very similar results with basis sets III and VII which are of the same quality as theirs. Another relevant feature in Table II is that the estimate of  $\Delta E^{\text{SCF}}$  shifts to more negative values when larger basis sets are used. In fact,

it becomes less than  $-7$  kcal/mol, so that even taking into account the aforementioned uncertainty of 5 kcal/mol, the heat of the reaction is still negative.

To get a better accuracy, in this work we included the electron correlation and the zero point energy contributions, and found them to be significant. Although both of them added positive corrections, their inclusion still leaves a negative value for the enthalpy and, since in our results the critical dependence of the predicted values on basis set and geometry has been eliminated, we can reliably conclude that the hydrolysis reaction (1) is definitely exothermic in the gas phase, with an estimated enthalpy of  $-3.83$  kcal/mol. This value may be subject to some small uncertainties caused for the following reasons: (a) even the largest basis set used is still not saturated; (b) the correlation energy was computed only up to the second order of the perturbation theory; and (c) the ZPE and thermal contributions are estimated with a harmonic approximation. These uncertainties can reasonably be expected not to modify substantially the above value, on behalf of the comprehensive and systematic survey of various basis sets and types of optimized geometries presented in Table II. Such an approach has been shown to give reliable results (see e.g. the revision by Sauer [26]). The above discussed trend of energy differences in successive columns in Table II allows us to propose that the overall uncertainty is of about 1 kcal/mol, and that we can assess an enthalpy of  $\Delta H = -3.8 \pm 1.0$  kcal/mol. Furthermore, it was found that the effect of a fully optimized geometry was of only 0.5 kcal/mol, and yielded a more negative value, hence supporting the conclusion of the exothermicity and the reliability of our estimated enthalpy.

An additional observation is that when using the pseudopotential calculation with basis VII, the  $\Delta E^{\text{SCF}}$  obtained for the hydrolysis reaction is comparable to that obtained in an all-electron computation with basis III, showing that it will be fruitful to use quite large basis sets with pseudopotentials to obtain better estimates for the heats of different chemical reactions in the gas phase.

Our result about the negative heat of reaction (1) strengthens the hypothesis that the intramolecular forces have a large contribution to the free energy of the phosphate anhydride bond hydrolysis, as was originally proposed by several authors [1-4]. However, in 1970, George et al. [8] showed that the differences in the hydration energies between reactants and products cannot be neglected and that their contribution to the  $\Delta G$  of the pyrophosphate hydrolysis in aqueous solution is highly relevant. Our estimate of the enthalpy of the hydrolysis of the pyrophosphate in the gas phase at 298° K,  $\Delta H = -3.8$  kcal/mol, is only one half of the value obtained from calorimetric measurements by George et al. [8] for the same reaction in aqueous

solution,  $\Delta H = -7.6$  kcal/mol, thus favouring the idea that the contribution of the hydration energies to the free energy of the hydrolysis is quite large. For the same reaction, the reported free energy is  $\Delta G = -9.5$  kcal/mol, so that the entropic term  $T\Delta S$  amounts to  $-1.9$  kcal/mol. These numbers suggest that the three effects, intramolecular interactions, solvation energies and entropy, are all relevant to the pyrophosphate hydrolysis.

Several attempts have been made to estimate the contribution of the hydration energies to the reactivity of the phosphate anhydrides: physical chemists have used molecular calculations and the reaction field theory (see e.g. Ref. 14), while biochemists have performed experiments in which water is substituted by aprotic organic solvents [9-13]. Unfortunately for the former approach, theoretical descriptions of liquid water are not yet well developed, though there have been important progresses recently (see e.g. Refs. 27 and 28). On the other hand, the results of the experiments show that the observed equilibrium constants,  $K_{\text{obs}}$ , for the hydrolysis of pyrophosphate in solution are significantly diminished when a mixture of an organic aprotic solvent (e.g. dimethyl sulfoxide), with water is used, instead of bulk water. From this result it is concluded that the free energy of the reaction gets close to zero and even becomes positive (in a mixture of 50% (w/v) poly(ethylene glycol) with water [13], at pH ranging from 7.0 to 8.0). The explanation advanced for this behaviour is in terms of a reduced water activity, and thus lower hydration energies and a large contribution of the entropy. Such a behaviour may help to explain how the same enzyme which catalyzes the hydrolysis of a high-energy phosphate compound can also catalyze the energetically unfavourable reverse process, i.e. its synthesis. The chemiosmotic hypothesis states that this synthesis is coupled to a proton gradient across a membrane [29], which supplies the required energy for the process to occur. Boyer's proposal [30] is that the synthesis takes place in an environment where it is either thermoneutral or exothermic, and that the energy derived from the proton gradient is used to release the products from the enzyme, whereas De Meis proposes [12] that the energy of the gradient is used to promote a conformational change in the enzyme, which in turn leads to a change in the binding environment of the catalytic site which is responsible for the change in the energy level of the phosphate compound. The desolvation hypothesis [6] proposes that the catalytic sites of enzymes are deprived of water and hence behave as a gas-phase-like environment. If this were true then the hydrolysis reaction should be at least thermoneutral in the gas phase. Such a conclusion does not agree with the present results: according to our estimated free energy, the hydrolysis is exothermic in the gas phase.

On the other hand, the concept of a lower water activity advanced by De Meis [12] does not necessarily require the reaction to be thermoneutral in the gas phase, provided that the entropic term  $T\Delta S$  is large enough to determine the value of  $\Delta G$ , even if  $\Delta H$  is negative. This assumption can help to explain the experimental data, but there are no calorimetric measurements proving this to be the case. For totally aqueous solutions [9], the values for  $\Delta G$ ,  $\Delta H$  and  $T\Delta S$  have been computed from the measurements of  $K_{obs}$  at different temperatures, and the results reported for the reaction occurring in mixtures with aprotic organic solvents [10–13] show a  $\Delta G$  also obtained from measurements of  $K_{obs}$ . The important point here is that a change in solvation would affect not only the entropy, but also the enthalpy, since there would be an effect on the interaction energies between the solutes and the solvent, apart from a simple reorganization. The comparison of our result with the microcalorimetric measurements of George et al. [8] support this statement. As it was mentioned before, there seem to be three almost equally relevant effects; the intramolecular interactions, the interactions with the solvent, and the change in entropy due to the reorganization of the solvent molecules around the reactants and the products. Therefore it is very relevant to get a reliable value of the enthalpy  $\Delta H$  of the reaction in the gas phase, to be able to determine the contribution of each effect.

In this work we have arrived at a quite reliable value for the enthalpy of reaction (1) in the gas phase. From the theoretical viewpoint this is a first step to be continued with the study of the interactions of liquid water with phosphates, which can give a better estimate of the solvation energies by means of a numerical simulation using the standard Monte Carlo or molecular dynamics methods. Of course this has to be done, but this first step is already providing valuable information about the intramolecular contributions to the enthalpy of the pyrophosphate hydrolysis.

An important caveat about the above discussion is that we are dealing only with electrically neutral species, while under physiological conditions the molecules are usually deprotonated, due to dissociation processes, so that a direct comparison of our results can only be made with the experimental data corresponding to the reaction involving only neutral molecular species [8]. The only reported SCF *ab initio* results concerning anions are those of Hayes et al. [14]. The heat of the reaction



that they obtain with a 4-31G basis set is of  $\Delta E = -70.98$  kcal/mol, but this value was obtained considering the orthophosphate molecules at an infinite separation, so that its validity is questionable. However, the

negative free energy obtained for this reaction is obvious to some extent, since the equation is highly unbalanced: the anion on the left hand side is doubly charged and each of the anions on the right hand side is only singly charged. As it was stated in the introduction, doubly charged anions are not stable in the gas phase. Even in aqueous solution, most of them are stable only in the presence of counterions. Therefore it is reasonable to expect that the hydrolysis of the charged species should be exothermic in the gas phase. Besides, as it was shown by Warshel et al. [7], in order to dehydrate the substrate the catalytic site must offer an energetically favourable environment, which cannot be gas-phase-like.

As a final comment, we want to stress that the presence of cations in the aqueous solution must be taken into account in the phosphate anhydride hydrolysis. In the case of the pyrophosphate it has been shown [31–34] that the enzymic catalysis of the hydrolysis does not occur in the absence of magnesium and that the true substrate of the enzyme is  $\text{MgP}_2\text{O}_7^{2-}$ . The role of magnesium in the free energy of the hydrolysis has not been considered yet.

It is clear that in order to understand the process of hydrolysis at a molecular level much work has still to be done on describing intramolecular rearrangements as well as several environmental effects, such as hydration, which ionic forms are present, which counterions, and certainly the catalytic sites of enzymes. In this respect, quantum calculations can be very helpful, because each molecular species involved can be studied separately, as opposed to an experimental approach where the data are necessarily thermodynamic, i.e., averaged over a large ensemble of interacting molecules, and hence cannot produce an accurate molecular description [35].

#### Acknowledgments

This study was supported by the National Science Foundation, U.S.-Consejo Nacional de Ciencia y Tecnología, México, Cooperative Science Program, under Grant No. INT-9016267. Two of us (HSM and IOB) thank J. Hernández for his help in the computational effort.

#### Note added in proof:

To further test the convergence of the results presented in Table II, we computed the heat of reaction (1) at the SCF level with two basis sets larger than basis V. These were a 6-311G (d,p) (triple zeta plus polarization functions on all atoms) and a 6-311G (2d,2p) (the same, but splitting the polarization functions), and the values obtained were  $-6.22$  kcal/mol and  $-9.01$  kcal/mol, respectively. These values extend

those presented in Table II and it can be seen that the trend towards more negative values for the reaction enthalpy is no longer valid, but instead an oscillatory behavior appears. This is to be expected when approaching the convergent value. This also allows us to estimate the discrepancy to the limit value to be less than  $\pm 1.5$  kcal/mol.

#### Appendix

This appendix presents the structural parameters of all the geometries studied for the molecules  $H_3PO_4$  and  $H_4P_2O_7$ . A comparison is made to test the convergence of the structural predictions. It is found that the structure of orthophosphate is quite well converged, as can be seen from the averages and standard deviations in Table AI. With the same criterion applied to Table

AII, it is concluded that the pyrophosphate preserves the structure of the orthophosphate, and that its most important degree of freedom is located in the P-O-P bridge. The geometries PV and PPV were fully optimized, so their parameters are all different, but due to the small change in relation to the structures PII and PIII, in Tables AI and AII some of the parameters shown have been averaged.

The SCF energies computed for all the geometries used are presented in Table AIII. They are related to the values computed with the corresponding basis set for the structures PII and PIII which were optimized in this work with basis II. The SCF energies computed with our largest basis set V for  $H_3PO_4$  and  $H_4P_2O_7$  are  $-642.036346$  a.u. and  $-1208.037309$  a.u., respectively.

TABLE AI

Comparison of the structural parameters of orthophosphate for the four geometries described in the text

The table is presented in the form of a 'Z-Matrix', which is a commonly used input for different programs, and shows the internal coordinates. The 'X' stands for a 'dummy' atom, whose only purpose is to define a symmetry. The last two columns correspond to the averages and standard deviations of the parameters, computed with the data in the first columns. Bond lengths are given in Å, and angles in degrees.

(A)

Z-Matrix used in the optimization

O1							
P2	O1	R1					
X	P2	1.0000	O1	90.00			
O3	P2	R2	O1	A1	X	0.00	
O4	P2	R2	O1	A1	X	120.00	
O5	P2	R2	O1	A1	X	-120.00	
H6	O3	R3	P2	A2	O1	D1	
H7	O4	R3	P2	A2	O1	D1	
H8	O5	R3	P2	A2	O1	D1	

(B)

Parameter	Geometries *					Average	S.D.
	PI	PII	PIII	PIV	PV		
I. Bond lengths (Å)							
R1	1.4940	1.4389	1.4526	1.4570	1.4526	1.4590	0.0185
R2	1.6020	1.5911	1.5711	1.5730	1.5710	1.5816	0.0127
R3	0.9419	0.9869	0.9644	0.9510	0.9644	0.9617	0.0152
II. Planar angles (deg)							
A1	114.60	117.41	115.47	115.50	115.46	115.69	0.93
A2	116.80	108.46	120.43	123.50	120.45	117.93	5.19
III. Dihedral angles (deg)							
D1	45.00	41.01	43.19		43.50	43.18	1.33

\* As described in the text.

TABLE II

Comparison of the structural parameters of pyrophosphate for the geometries described in the text

The table is presented in the form of a 'Z-Matrix', which is a commonly used input for different programs, and shows the internal coordinates. The last two columns correspond to the averages and standard deviations of the parameters, computed with the data in the first columns. Bond lengths are given in Å, and angles in degrees.

(A)

Z-Matrix used in the optimization

O1							
P2	O1	R1					
P3	O1	R1	P2	A1			
O4	P2	R2	O1	A2	P3	D1	
O5	P3	R3	O1	A3	P2	D2	
O6	P2	R4	O1	A4	P3	D3	
O7	P3	R5	O1	A5	P2	D4	
O8	P2	R6	O1	A6	P3	D5	
O9	P3	R7	O1	A7	P2	D6	
H10	O6	R8	P2	A8	O4	D7	
H11	O7	R8	P3	A9	O5	D8	
H12	O8	R8	P2	A10	O4	D9	
H13	O9	R8	P3	A11	O5	D10	

(B)

Parameter	Geometries *						Average	S.D.	
	PP1 (eclipsed)	PPII (staggered)		PPIII (staggered)	PPIV (staggered) (eclipsed)				PPV (staggered)
I. Bond lengths (Å)									
R1	1.5700	1.5948	1.5886	1.5809	1.5600	1.5680	1.5621	1.5749	0.0124
R2	1.4900	1.4370	1.4368	1.4471	1.4530	1.4510	1.4448	1.4514	0.0168
R3	1.4900	1.4375	1.4368	1.4471	1.4530	1.4510	1.4444	1.4514	0.0168
R4	1.5900	1.5904	1.5898	1.5697	1.5700	1.5710	1.5912	1.5817	0.0100
R5	1.6000	1.5915	1.5898	1.5697	1.5700	1.5710	1.5701	1.5803	0.0120
R6	1.6000	1.5878	1.5898	1.5697	1.5700	1.5710	1.5676	1.5794	0.0119
R7	1.5900	1.5866	1.5898	1.5697	1.5700	1.5710	1.5568	1.5763	0.0118
R8	0.9498	0.9869	0.9869	0.9654	0.9510	0.9510	0.9661	0.9653	0.0150
II. Planar angles (deg)									
A1	152.10	126.27	134.78	138.96	170.40	162.50	147.15	147.45	14.42
A2	115.49	117.59	113.75	115.23			116.68	115.75	1.31
A3	115.49	117.21	113.75	115.23			113.99	115.13	1.28
A4	102.74	100.30	102.29	101.80			98.80	101.19	1.45
A5	102.91	102.05	102.29	101.80			101.82	102.17	0.41
A6	102.91	99.06	102.29	101.80			105.21	102.25	1.98
A7	102.74	97.62	102.29	101.80			101.25	101.14	1.83
A8	118.84	108.43	108.24	120.36			120.54	115.28	5.70
A9	116.69	108.45	108.24	120.36			119.36	114.62	5.26
A10	116.69	108.70	108.24	120.36			121.04	115.01	5.54
A11	118.84	108.51	108.24	120.36			120.69	115.33	5.71
III. Dihedral angles (deg)									
D1	0.00	60.23	180.58	141.92			139.27		
D2	0.00	7.49	180.58	268.89			220.38		
D3	-126.78	-172.16	52.22	15.45			13.94		
D4	126.56	136.08	-51.79	141.29			96.33		
D5	126.56	-69.20	-51.79	-91.52			-91.54		
D6	-126.78	-121.32	52.22	34.73			-12.01		
D7	-56.86	34.10	36.95	18.71			41.81		
D8	-115.10	36.30	36.95	44.79			1.32		
D9	-115.10	31.49	36.95	18.71			28.42		
D10	-56.86	44.78	36.95	44.79			127.06		

\* As described in the text.

TABLE VIII

Energies computed with different basis sets for the different geometries of the  $H_3PO_4$  and  $H_4P_2O_7$  molecules

For each basis set, the values obtained for the PIII and PPIII geometries were used as a reference. All energies are given in kcal/mol.

Basis sets *	I	II	III	IV	V	VI	VII
Molecule: $H_3PO_4$							
PI	5.38	4.71		1.50		-1.97	5.26
PII	-8.46	7.09		1.64		13.73	12.09
PV		0.00			0.02		
Molecule: $H_4P_2O_7$							
PPI	9.12	12.93		8.27		-0.37	7.62
PPII eclipsed	-11.32	13.80		1.90		22.92	19.11
PPII staggered	-13.59	12.95		1.19		21.08	19.24
PPV		-2.30			1.09		

\* As described in the text.

## References

- Kalra, H.M. (1941) *Chem. Rev.* 28, 71-142.
- Hill, T.L. and Morales, M.H. (1951) *J. Am. Chem. Soc.* 73, 1656-1660.
- Pullman, A. and Pullman, B. (1963) *Quantum Biochemistry*, Interscience, New York.
- Boyd, D.B. and Lipscomb, W.N. (1969) *J. Theor. Biol.* 25, 403-420.
- Kraut, J. (1988) *Science* 242, 533-540.
- Dewar, M.J.S. and Storch, D.M. (1985) *Proc. Natl. Acad. Sci. USA* 82, 2225-2229.
- Warshel, A., Åqvist, J. and Creighton, S. (1989) *Proc. Natl. Acad. Sci. USA* 86, 5820-5824.
- Geuge, P., Witonsky, R.J., Trachtman, M., Wu, C., Dorwart, W., Richman, L., Richman, W., Shurayh, F. and Lentz, B. (1970) *Biochim. Biophys. Acta* 223, 1-15.
- De Meis, L. (1981) *J. Biol. Chem.* 259, 6090-6097.
- De Meis, L., Behrens, M.L., Petretski, J.H. and Politi, M.J. (1985) *Biochemistry* 24, 7783-7789.
- Woffenden, R. and Williams, R. (1985) *J. Am. Chem. Soc.* 107, 4345-4346.
- De Meis, L. (1989) *Biochim. Biophys. Acta* 973, 333-349.
- Romero, P.J. and De Meis, L. (1989) *J. Biol. Chem.* 264, 7869-7873.
- Hayes, M.D., Kenyon, L.G. and Kollman, P.A. (1978) *J. Am. Chem. Soc.* 100, 4331-4340.
- O'Keefe, M., Domengès, B. and Gibbs, G.V. (1985) *J. Phys. Chem.* 89, 2301-2309.
- Ewig, C.S. and Van Wazer, J.R. (1988) *J. Am. Chem. Soc.* 110, 79-86.
- Frisch, M.J., Binkley, J.S., Schlegel, H.B., Raghavachari, K., Melius, C.F., Martin, R.L., Stewart, J.J.P., Bobrowicz, F.W., Rohlfing, C.M., Kahn, L.R., Defrees, D.J., Seeger, R., Whiteside, R.A., Fox, D.J., Flender, E.M. and Pople, J.A. (1984) GAUSSIAN-86, Carnegie-Mellon QCPU, Pittsburgh, PA.
- Stewart, J.J.P. (1985) *QCPE Bull.* 5, 133-134.
- Frisch, M.J., Head-Gordon, M., Schlegel, H.B., Raghavachari, K., Binkley, J.S., González, C., Defrees, D.J., Fox, D.J., Whiteside, R.A., Seeger, R., Melius, C.F., Baker, J., Martin, R.L., Kahn, L.R., Stewart, J.J.P., Fluder, E.M., Topiol, S. and Pople, J.A. (1988) GAUSSIAN-88, Gaussian, Inc., Pittsburgh.
- Barthelat, J.C., Durand, Ph. and Serafini, A. (1977) *Mol. Phys.* 33, 159-170.
- Peterson, M. and Poirier, R.A. (1987) MONSTERGAUSS, University of Toronto, Ontario.
- Ewig, C.S. and Van Wazer, J.R. (1985) *J. Am. Chem. Soc.* 107, 1965-1971.
- Atkins, P.W. (1982) *Physical Chemistry*, 2nd Edn. Oxford University Press, Oxford.
- Hehre, W.J., Ditchfield, D., Radom, L. and Pople, J.A. (1970) *J. Am. Chem. Soc.* 92, 4796-4801.
- George, P., Trachtman, M., Brett, A.M. and Bock, C.W. (1977) *J. Chem. Soc. Perkin Trans. 2*, 1036-1047.
- Sauer, J. (1989) *Chem. Rev.* 89, 199-223.
- Sain-Martín, H., Medina-Llanos, C. and Ortega-Blake, I. (1990) *J. Chem. Phys.* 93, 6448-6452.
- Niesar, U., Corungiu, G., Clementi, E., Kneller, G.R. and Bhattacharya, D.K. (1990) *J. Phys. Chem.* 94, 7919-7958.
- Mitchell, P. (1968) *Chemiosmotic Coupling and Energy Transduction*, Glynn Research, Bodmin.
- Boyer, P.D., Chance, B., Ernster, L., Mitchell, P., Racker, E. and Slater, E.C. (1977) *Annu. Rev. Biochem.* 46, 955-1026.
- Celis, H. and Romero, J. (1977) *J. Bioenerget. Biomembr.* 19, 255-271.
- Knight, W.B., Fitts, S.W. and Dunaway-Mariano, D. (1981) *Biochemistry* 20, 4079-4086.
- Ting, S.J. and Dunaway-Mariano, D. (1984) *FEBS Lett.* 165, 251-256.
- Shorter, A.L., Harony, T.P., Scalzo-Brush, T., Knight, W.B., Dunaway-Mariano, D. and M. Sundaralingam (1987) *Biochemistry* 26, 2060-2066.
- Pratt, L.R. (1985) *Annu. Rev. Phys. Chem.* 36, 433-449.