



**UNIVERSIDAD DE EXTREMADURA**

Departamento de Química Física

---

---

**Cálculo de las derivadas de  
la energía libre en disolución  
empleando un método  
*Quantum Mechanics /  
Molecular Mechanics (QM/MM)* y  
la aproximación del campo medio**

---

**Ignacio Fernández Galván**

Badajoz, 2004

Edita: Universidad de Extremadura  
Servicio de Publicaciones  
c/ Caldereros 2, Planta 3ª  
Cáceres 10071  
Correo e.: [publicac@unex.es](mailto:publicac@unex.es)  
<http://www.pcid.es/public.htm>

ᑕᑭ ᑭᑕᑭ ᑭ ᑭᑕᑕᑭ ᑭᑕᑕᑕ  
ᑕᑭ ᑭᑕᑕᑕ ᑭ ᑕᑕᑕ ᑕᑕᑕ ᑕᑕᑕ



**DEPARTAMENTO DE  
QUÍMICA FÍSICA**

Campus Universitario  
Avda. de Elvas, s/n  
06071-BADAJOS  
Teléfono: 924-289399  
Fax: 924-275576  
e-mail: [angel@unex.es](mailto:angel@unex.es)

ÁNGEL LÓPEZ PIÑEIRO, CATEDRÁTICO DE UNIVERSIDAD Y DIRECTOR DEL DEPARTAMENTO DE QUÍMICA FÍSICA DE LA UNIVERSIDAD DE EXTREMADURA,

INFORMA: Que el licenciado D. Ignacio Fernández Galván ha venido realizando en el Departamento de Química Física, bajo la dirección de los Dres. Manuel Ángel Aguilar Espinosa y Francisco Javier Olivares del Valle, los trabajos de investigación que aparecen reflejados en la memoria titulada **«Cálculo de las derivadas de la energía libre en disolución empleando un método Quantum Mechanics / Molecular Mechanics (QM/MM) y la aproximación del campo medio»**, conducentes a la obtención del grado de Doctor Europeo en Ciencias Químicas.

Y para que conste, expido el presente informe en Badajoz, a 26 de octubre de 2004.


Fdo. Prof. Dr. Ángel López Piñeiro  
Director del Departamento



---

# RESUMEN

En las últimas décadas, la química computacional se ha convertido en una herramienta muy importante para el estudio de las propiedades de la materia. Los sistemas en fase condensada —especialmente en fase líquida— presentan, sin embargo, una serie de propiedades que hacen necesario el desarrollo de métodos específicos y la introducción de aproximaciones adicionales a las normalmente empleadas para los sistemas en fase gaseosa. Esta tesis se centra en uno de estos métodos, denominado ASEP/MD (*Averaged Solvent Electrostatic Potential from Molecular Dynamics*) y desarrollado por nuestro grupo de investigación.

El método ASEP/MD ha sido ampliado para permitir la optimización de estructuras moleculares en disolución, para lo cual es necesario disponer de las derivadas de la energía libre con respecto a las coordenadas nucleares. Como trabajo complementario, se ha puesto a punto un método para el cálculo de diferencias de energía libre y se ha verificado cuantitativamente la validez de la aproximación del campo medio, empleada por nuestro método. Asimismo, se ha elaborado un programa informático que facilita la puesta en práctica del desarrollo teórico.

Se han realizado también algunas aplicaciones prácticas a sistemas de interés. La validez de la aproximación del campo medio se ha comprobado con sistemas sencillos como alcoholes en estado líquido y disoluciones acuosas de compuestos carbonílicos. El buen funcionamiento del método de optimización de geometrías se ha verificado también con moléculas sencillas como el agua, el metanol o la formamida. Se ha efectuado un estudio más detallado sobre el equilibrio anomérico de un monosacárido, la D-xilopiranos, y sobre una reacción  $S_N2$ , la reacción de Menshutkin entre el  $NH_3$  y el  $CH_3Cl$ . Finalmente se ha estudiado la dinámica de la isomerización del triazeno ( $N_3H_3$ ) en agua mediante una aplicación combinada del método ASEP/MD y de un método QM/MM tradicional.





---

# AGRADECIMIENTOS

La elaboración de esta tesis doctoral no habría sido posible sin la colaboración y el apoyo de muchas personas, a las que quiero expresar mi agradecimiento.

Sin duda, gran parte del mérito que este trabajo pueda tener corresponde a mis directores de tesis, los Dres. Manuel Á. Aguilar Espinosa y Fco. Javier Olivares del Valle, por sus consejos, sugerencias, discusiones y críticas. Asimismo, ha sido de gran importancia la colaboración de Manuel F. Ruiz López, quien me acogió en su grupo durante mi estancia en Nancy.

No menos importantes han sido la ayuda y los ánimos prestados por mis compañeros de trabajo. Especialmente, quiero agradecer a M.<sup>a</sup> Luz Sánchez, M.<sup>a</sup> Elena Martín y Aurora Muñoz la paciencia y amistad que me han demostrado durante estos años, a José C. Corchado y Jorge A. Sansón su disponibilidad y colaboración, a M.<sup>a</sup> Dolores Feijoo su constante cooperación y, en general, a todos los miembros del Departamento de Química Física la hospitalidad y cercanía que han hecho tan agradables las horas de trabajo.

Por supuesto, quiero también agradecer a mi familia y amigos todas las muestras de cariño y el apoyo que me han prestado durante tantos años, y a todos los profesores que he tenido durante mi vida de estudiante, que me han ayudado a llegar hasta aquí.

Por último, si bien los aciertos de este trabajo son, en buena parte, atribuibles a otros, los fallos y errores, que los habrá, son únicamente de mi responsabilidad.



---

# ÍNDICE

ÍNDICE	13
ÍNDICE DE TABLAS	19
ÍNDICE DE FIGURAS	21
<b>1. INTRODUCCIÓN Y OBJETIVOS</b>	<b>23</b>
1.1. Métodos para el estudio de los efectos del disolvente .....	24
1.2. Objetivos .....	28
1.3. Sumario .....	29

**PRIMERA PARTE**  
**MÉTODOS MIXTOS**

---

<b>2. REPRESENTACIÓN CUÁNTICA DEL SOLUTO</b>	<b>33</b>
2.1. La aproximación de Born-Oppenheimer .....	33
2.2. Determinantes de Slater .....	35
2.3. El método Hartree-Fock .....	37
2.3.1. Combinación lineal de orbitales atómicos .....	39
2.3.2. El método del campo autocoherente.....	39

2.4.	Correlación electrónica . . . . .	40
2.4.1.	Métodos Møller-Plesset . . . . .	41
2.4.2.	Interacción de configuraciones . . . . .	42
2.4.3.	Métodos multiconfiguracionales . . . . .	43
2.4.4.	Métodos de grupos acoplados . . . . .	44
2.5.	La teoría del funcional de la densidad . . . . .	45
2.5.1.	La formulación de Kohn-Sham . . . . .	46
2.5.2.	El funcional de correlación-intercambio . . . . .	48
	<i>La aproximación de la densidad local . . . . .</i>	<i>48</i>
	<i>Funcionales corregidos con el gradiente . . . . .</i>	<i>49</i>
<b>3.</b>	<b>REPRESENTACIÓN CLÁSICA DEL DISOLVENTE</b>	<b>51</b>
3.1.	Métodos de disolvente continuo . . . . .	51
3.1.1.	Los primeros modelos . . . . .	52
3.1.2.	Métodos de campo de reacción autocohérente . . . . .	53
3.1.3.	El modelo del continuo polarizable . . . . .	54
3.2.	Modelos de disolvente estructurado . . . . .	55
3.2.1.	El modelo de referencia de lugares de interacción . . . . .	55
3.2.2.	El modelo de dipolos de Langevin . . . . .	56
3.3.	Métodos de mecánica molecular . . . . .	57
3.3.1.	Simulaciones . . . . .	59
	<i>El método de Monte Carlo . . . . .</i>	<i>61</i>
	<i>Dinámica molecular . . . . .</i>	<i>63</i>
3.3.2.	Métodos combinados de mecánica cuántica y molecular . . . . .	66
	<i>El método DFT/MM . . . . .</i>	<i>69</i>
<b>4.</b>	<b>EL MÉTODO ASEP/MD</b>	<b>73</b>
4.1.	Descripción del método . . . . .	74
4.2.	Métodos acoplados . . . . .	80
4.2.1.	Método parcialmente acoplado . . . . .	80
4.2.2.	Método totalmente acoplado . . . . .	82

---

4.3. Disolvente polarizable.....	83
4.4. La componente Stark.....	86
4.4.1. Evaluación aproximada de la componente Stark .....	87

**SEGUNDA PARTE**  
**AMPLIACIÓN DEL MÉTODO ASEP/MD**

---

<b>5. CÁLCULO DE LA ENERGÍA LIBRE</b>	<b>91</b>
5.1. Métodos para el cálculo de la energía libre .....	92
5.1.1. Perturbación termodinámica .....	92
5.1.2. Integración termodinámica.....	94
5.1.3. Desarrollo lento.....	95
5.2. Aplicación en el método ASEP/MD.....	96
<b>6. OPTIMIZACIÓN DE GEOMETRÍA</b>	<b>99</b>
6.1. Fundamento .....	100
6.2. Métodos de optimización .....	102
6.2.1. Métodos de primer orden .....	102
6.2.2. Métodos de segundo orden.....	103
<i>El método RFO</i> .....	105
6.3. Aplicación en el método ASEP/MD.....	108

**TERCERA PARTE**  
**APLICACIONES**

---

<b>7. CONDICIONES GENERALES</b>	<b>113</b>
<b>8. VALIDACIÓN DE LA APROXIMACIÓN DEL CAMPO MEDIO</b>	<b>117</b>
8.1. Detalles de los cálculos .....	117

8.2. Resultados y discusión .....	118
8.2.1. Influencia del número de configuraciones .....	119
8.2.2. Componente Stark.....	120
8.2.3. Gradiente de la energía libre.....	122
8.3. Conclusiones.....	125
<b>9. OPTIMIZACIÓN DE GEOMETRÍAS EN DISOLUCIÓN</b>	<b>127</b>
9.1. Detalles de los cálculos .....	127
9.2. Resultados y discusión .....	128
9.2.1. Variación de la energía libre .....	128
9.2.2. Formamida .....	129
9.2.3. Agua y metanol.....	133
9.3. Conclusiones.....	138
<b>10. EQUILIBRIO ANOMÉRICO EN LA D-XILOPIRANOSA</b>	<b>139</b>
10.1. Detalles de los cálculos .....	141
10.2. Resultados y discusión .....	143
10.2.1. Resultados energéticos .....	143
10.2.2. Resultados estructurales .....	147
10.3. Conclusiones.....	150
<b>11. BARRERA DE ACTIVACIÓN DE LA REACCIÓN DE MENSHTUKIN</b>	<b>151</b>
11.1. Detalles de los cálculos .....	152
11.2. Resultados y discusión .....	153
11.3. Conclusiones.....	157
<b>12. DESPLAZAMIENTO PROTÓNICO 1,3 EN EL TRIAZENO</b>	<b>161</b>
12.1. Método empleado .....	162
12.2. Detalles de los cálculos .....	163
12.2.1. Cálculos DFT/MM.....	164

12.3. Resultados y discusión .....	165
12.3.1. Triazeno en disolución acuosa .....	165
<i>Estructura del disolvente.</i> .....	168
12.3.2. Intermedios y estados de transición .....	171
12.3.3. Resultados energéticos .....	174
12.3.4. Trayectorias de reacción DFT/MM .....	178
12.4. Conclusiones .....	183
<b>13. CONCLUSIONES GENERALES</b>	<b>185</b>
13.1. Desarrollo futuro .....	186

## APÉNDICES

---

<b>A. AJUSTE DE LAS CARGAS AL POTENCIAL ELECTROSTÁTICO</b>	<b>191</b>
<b>B. DISOLVENTE POLARIZABLE</b>	<b>193</b>
B.1. Cálculo de dipolos inducidos .....	193
B.2. Energías de interacción .....	194
<b>C. ENERGÍA LIBRE EN LOS MÉTODOS QM/MM</b>	<b>197</b>
<b>D. ACTUALIZACIÓN DE LA HESSIANA</b>	<b>201</b>
<b>E. DERIVADAS DE LA ENERGÍA LIBRE</b>	<b>203</b>
<b>F. CÓDIGO DEL PROGRAMA ASEP/MD</b>	<b>207</b>
F.1. Principal.f90 .....	207
F.2. Ejecutar.f90 .....	223
F.3. SubrutinasMoldy.f90 .....	227
F.4. SubrutinasGaussian.f90 .....	236
F.5. SubrutinasHondo.f90 .....	242
F.6. Calculos.f90 .....	249
F.7. Optim.f90 .....	271

<b>BIBLIOGRAFÍA</b>	<b>285</b>
---------------------	------------

**ARTÍCULOS PUBLICADOS**

---

<b>J. PHYS. CHEM. B 106 (2002), 4813-4817</b>	<b>313</b>
<b>J. CHEM. PHYS. 118 (2003), 255-263</b>	<b>319</b>
<b>COMPUT. PHYS. COMMUN. 155 (2003), 244-259</b>	<b>329</b>
<b>THEOR. CHEM. ACC. 111 (2004), 196-203</b>	<b>345</b>
<b>J. COMPUT. CHEM. 25 (2004), 1227-1233</b>	<b>353</b>



---

# ÍNDICE DE TABLAS

Tabla 7.1.	Geometrías y cargas atómicas de las moléculas de disolvente ...	115
Tabla 7.2.	Geometrías experimentales de los compuestos carbonílicos . . . . .	115
Tabla 7.3.	Parámetros de Lennard-Jones utilizados en los cálculos.....	116
Tabla 8.1.	Momentos dipolares y polarizabilidades.....	119
Tabla 8.2.	Energías y momentos dipolares promedio.....	120
Tabla 8.3.	Energías de interacción y momentos dipolares . . . . .	121
Tabla 8.4.	Evaluación aproximada de la componente Stark.....	122
Tabla 8.6.	Gradientes promedio para la formamida.....	124
Tabla 9.7.	Evolución de la geometría de la formamida . . . . .	134
Tabla 9.8.	Cargas atómicas de la formamida . . . . .	134
Tabla 10.2.	Energías relativas de la D-xilopiranososa . . . . .	143
Tabla 10.5.	Energía libre para la D-xilopiranososa.....	145
Tabla 10.7.	Energías de interacción para la D-xilopiranososa . . . . .	146
Tabla 10.8.	Geometrías de la D-xilopiranososa . . . . .	148
Tabla 10.9.	Cargas atómicas de la D-xilopiranososa . . . . .	149
Tabla 11.1.	Energías y cargas en la reacción de Menshutkin . . . . .	154
Tabla 11.4.	Energía libre de activación de la reacción de Menshutkin.....	156
Tabla 12.2.	Geometrías, cargas y dipolo del triazeno.....	166

Tabla 12.7.	Geometrías y cargas del complejo triazeno-agua . . . . .	172
Tabla 12.9.	Geometrías y cargas del estado de transición unimolecular . . . . .	174
Tabla 12.10.	Geometrías y cargas del estado de transición bimolecular . . . . .	175
Tabla 12.12.	Energías libres para las reacciones del triazeno . . . . .	176
Tabla 12.13.	Energías libres de activación en vacío . . . . .	177

---

# ÍNDICE DE FIGURAS

Figura 1.1.	Métodos para el estudio de los efectos del disolvente . . . . .	25
Figura 4.1.	Esquema del método ASEP/MD . . . . .	76
Figura 4.2.	Representación del disolvente . . . . .	78
Figura 4.3.	Esquema del método ASEP/MD acoplado . . . . .	81
Figura 6.1.	Tipos de puntos estacionarios. . . . .	102
Figura 6.2.	Esquema del método ASEP/MD con optimización . . . . .	110
Figura 8.5.	Molécula de formamida . . . . .	123
Figura 9.1.	Diferencia de energía libre calculada por dos métodos . . . . .	129
Figura 9.2.	Evolución del gradiente de la formamida . . . . .	130
Figura 9.3.	Evolución de la energía libre de la formamida. . . . .	131
Figura 9.4.	Evolución del momento dipolar de la formamida . . . . .	132
Figura 9.5.	Evolución de la RDF oxígeno-oxígeno para la formamida. . . . .	132
Figura 9.6.	Geometrías de la formamida . . . . .	133
Figura 9.9.	Evolución del gradiente del agua. . . . .	135
Figura 9.10.	Evolución de la energía libre del agua . . . . .	135
Figura 9.11.	Geometrías del agua . . . . .	136
Figura 9.12.	Evolución del gradiente del metanol . . . . .	136
Figura 9.13.	Evolución de la energía libre del metanol . . . . .	137
Figura 9.14.	Geometrías del metanol . . . . .	137

Figura 10.1.	Estructuras de la D-xilopiranososa. . . . .	142
Figura 10.3.	Evolución del gradiente de la D-xilopiranososa . . . . .	144
Figura 10.4.	Evolución de la energía libre de la D-xilopiranososa. . . . .	144
Figura 10.6.	Cálculo de FEP para la D-xilopiranososa . . . . .	146
Figura 10.10.	RDF para el O anomérico de la D-xilopiranososa. . . . .	149
Figura 10.11.	RDF para el O <sup>8</sup> de la D-xilopiranososa. . . . .	150
Figura 11.2.	Geometrías del amoniaco y del cloruro de metilo . . . . .	154
Figura 11.3.	Geometrías del estado de transición. . . . .	155
Figura 11.5.	RDF del amoniaco y del cloruro de metilo . . . . .	158
Figura 11.6.	RDF del estado de transición . . . . .	158
Figura 12.1.	Mecanismos de desplazamiento protónico en el triazeno . . . . .	162
Figura 12.3.	Distribución de diedros en el triazeno con DFT/MM. . . . .	167
Figura 12.4.	RDF para el N <sup>1</sup> del triazeno. . . . .	169
Figura 12.5.	RDF para el N <sup>3</sup> del triazeno. . . . .	169
Figura 12.6.	RDF para el N <sup>3</sup> del triazeno con distintas geometrías. . . . .	170
Figura 12.8.	Geometrías del complejo triazeno-agua. . . . .	173
Figura 12.11.	Geometrías del estado de transición bimolecular . . . . .	175
Figura 12.14.	Tipos de trayectorias de reacción para el triazeno . . . . .	180
Figura 12.15.	Comportamientos del agua en la reacción con el triazeno. . . . .	181
Figura 12.16.	Trayectoria con persistencia del estado de transición . . . . .	182
Figura 12.17.	Persistencia del estado de transición frente a energía . . . . .	183

---

---

# 1

---

## INTRODUCCIÓN Y OBJETIVOS

La mecánica cuántica ha sido, indiscutiblemente, uno de los mayores logros científicos del siglo XX. A pesar de que sus fundamentos y predicciones resultan muchas veces contrarios al sentido común, una y otra vez se han visto confirmados los efectos más extraños y peregrinos. Hoy en día, la mecánica cuántica puede considerarse una de las teorías más exitosas de la historia de la ciencia, no obstante, Richard Feynman dijo que «puede decirse que nadie entiende la mecánica cuántica» y probablemente sigue siendo cierto.

El punto de vista de la mecánica cuántica ha revolucionado la manera de interpretar fenómenos espectroscópicos, eléctricos, químicos... ha afectado a la astronomía, la electrónica, la bioquímica, la farmacología... En especial, los conceptos que ahora se tienen de la naturaleza del enlace químico o del origen de las propiedades termodinámicas están tan íntimamente ligados a las ideas de la mecánica cuántica que resulta difícil encontrar algún campo de la química donde aquélla no tenga cabida. El gran desarrollo experimentado por la química cuántica ha permitido, al menos para moléculas aisladas, hacer predicciones de propiedades estructurales, termodinámicas, espectroscópicas y cinéticas.

Por otro lado, una gran parte de los procesos químicos de interés práctico tienen lugar en fase condensada y no en fase gaseosa. El papel que, en estos procesos, juega el entorno —el disolvente, la proteína envolvente, etc.— adquiere en muchas ocasiones una importancia crucial, que no puede dejarse de lado a la hora de explicar y predecir las propiedades y el comportamiento de los sistemas químicos. El estudio de este papel es básico para la comprensión de sistemas biológicos, bioquímicos y de otro tipo, con aplicaciones tecnológicas, médicas, etc. En procesos sometidos a

catálisis, por ejemplo, la interacción de la parte reactiva del sistema con su entorno puede determinar que el mecanismo de reacción sea uno u otro, modificando de manera importante tanto la velocidad de la reacción como la naturaleza y proporciones de los productos.

La conjunción de los dos aspectos mencionados —la mecánica cuántica como marco teórico, la química en fase condensada como objeto de estudio— no es sencilla. Todos los métodos químico-cuánticos, desde los más simples a los más complejos, se encuentran con problemas para abordar el estudio de sistemas en fase condensadas. Entre las dificultades existentes, sobre todo en el caso de sistemas líquidos, se pueden destacar:

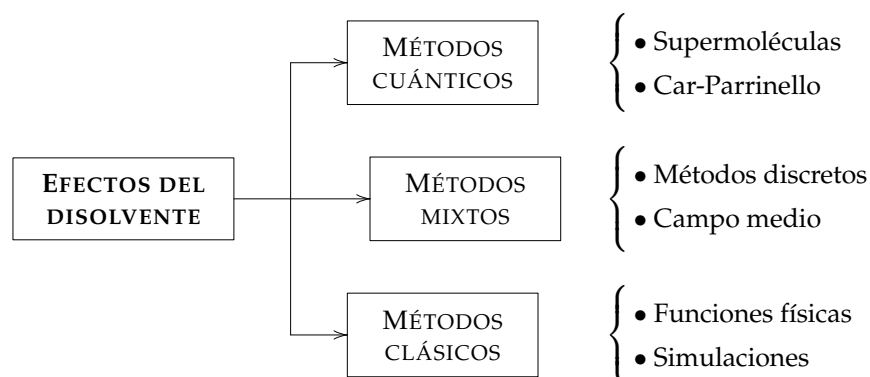
- El número de moléculas implicadas en estos sistemas es enorme y no puede reducirse arbitrariamente, llega un momento en que las propiedades del sistema reducido no pueden considerarse iguales a las del que se pretende estudiar. Incluso con el mínimo número de moléculas necesario para conservar las propiedades, los sistemas son a menudo inabordables.
- El carácter amorfo y desordenado de la mayoría de los líquidos hace que no sea posible aplicar simplificaciones por consideraciones de simetría, que en otro tipo de sistemas resultan muy prácticas.
- La mayor parte de las interacciones moleculares en disolución se caracterizan por una baja especificidad y direccionalidad. Esto implica que existe gran cantidad de configuraciones del sistema con energía similar que contribuyen apreciablemente a sus propiedades... y que es necesario considerar para obtener resultados estadísticamente significativos.

Es por ello importante el desarrollo de modelos y aproximaciones que, salvaguardando la esencia del problema, introduzcan simplificaciones que hagan factible el estudio de moléculas y procesos en disolución.

### **1.1. Métodos para el estudio de los efectos del disolvente**

Los avances tecnológicos que han tenido lugar, sobre todo en el último medio siglo, en el campo computacional han permitido una aplicación cada vez más rigurosa de la teoría cuántica al estudio de los sistemas químicos y de los efectos que el medio envolvente tiene sobre sus propiedades. Simultáneamente, la creciente capacidad de cálculo de las computadoras ha hecho que se puedan estudiar sistemas cada vez más complejos y se ha elevado el nivel de precisión y exactitud exigido a los cálculos teóricos. Esto ha llevado al desarrollo de un gran número

de modelos y aproximaciones para el estudio de estos sistemas. Este proceso continúa produciéndose, la potencia y capacidad de los sistemas informáticos sigue en aumento y persiste el desarrollo de nuevos métodos y modelos. Por lo tanto, cualquier clasificación detallada de éstos corre el riesgo de ser incompleta y quedar desfasada en unos pocos años. Sin embargo, puede establecerse una clasificación general de los modelos teóricos más empleados en el estudio de los efectos del disolvente, según el punto de vista desde el que se describen las distintas moléculas del sistema, esquematizada en la figura 1.1:



**Figura 1.1:** Clasificación de los métodos para el estudio de los efectos del disolvente.

- **Métodos cuánticos.** En este tipo de métodos, tanto el soluto como el disolvente se estudian desde un punto de vista cuántico y, por lo tanto, están descritos por su función de estado o función de onda. Los métodos cuánticos presentan la ventaja de los que proporcionan resultados potencialmente más precisos y fiables, pero su elevado coste computacional obliga muchas veces a hacer más simplificaciones de las deseables. Dentro de este grupo pueden destacarse:
  - *Método de supermoléculas.* Este método fue empleado por primera vez por Morokuma y Pedersen [1] y, posteriormente, por Alagona y col. [2] y Pullman y Pullman [3]. En este modelo se considera una sola molécula de soluto y un reducido número de moléculas de disolvente, generalmente sólo aquellas que, por su situación y orientación, se prevé que puedan ejercer una mayor influencia sobre las propiedades del soluto. Este procedimiento resulta adecuado para el estudio de agregados moleculares (*clusters*), siempre que se considere un número suficiente de moléculas y de configuraciones. Para líquidos y disoluciones, sin embargo, es del todo inapropiado, ya que desprecia el efecto de la mayor parte de las moléculas de disolvente y, en general, no tiene en cuenta los efectos termodinámicos asociados al gran número de configuraciones que son

accesibles al sistema en virtud de la agitación térmica, pues se obvia por completo su dinámica.

- *Método Car-Parrinello* [4]. Es la opción más compleja y costosa desde el punto de vista computacional, pero a cambio es la que proporciona una descripción más completa del sistema formado por el soluto y el disolvente. Este método se basa en la resolución simultánea de las ecuaciones clásicas del movimiento para los núcleos y la ecuación de Schrödinger para las distribuciones electrónicas de las moléculas que forman parte del sistema líquido, lo que conduce a la obtención de las propiedades dinámicas del sistema. La combinación de los aspectos dinámico y cuántico hace de este método una herramienta excepcional para el estudio de procesos reactivos en disolución. En la actualidad, sin embargo, el empleo del método Car-Parrinello está muy limitado debido a la gran complejidad de los cálculos que implica.
- **Métodos clásicos.** En este grupo se engloban todos aquellos modelos en los que tanto el soluto como el disolvente se representan de forma clásica [5]. En general, estos métodos no proporcionan información de tipo químico, ya que la ruptura y formación de enlaces es un problema estrictamente cuántico, pero sí pueden aportar una valiosa información estructural y termodinámica del sistema. Se pueden dividir en dos categorías:
  - *Métodos basados en la construcción de funciones físicas* [6]. Se incluyen los métodos basados en teorías perturbacionales y en la utilización de ecuaciones integrales para la descripción de las funciones de distribución de las moléculas de disolvente, así como otras aproximaciones basadas en el desarrollo del virial de la función de estado. En general, estos métodos dejan en un segundo plano los aspectos químicos del problema y recurren a importantes simplificaciones en la forma del potencial de interacción intermolecular. Pueden citarse las siguientes aproximaciones teóricas: PY (Percus-Yevick) [7], RISM (*Reference Interaction Site Model*) [8, 9], MSA (*Mean Spherical Approximation*) [10], HNC (*Hypernetted Chain*) [11-14], etc.
  - *Técnicas de simulación* [15]. En los métodos pertenecientes a este grupo, el soluto y el disolvente se representan como un conjunto de partículas interaccionantes sencillas. Por distintos métodos puede obtenerse un conjunto representativo de configuraciones de estas partículas y, a partir de éstas, las distintas propiedades del sistema. En este caso el problema cuántico se sustituye por uno clásico en el que la energía electrónica y las interacciones electrostáticas de los núcleos se modelan conjuntamente a través de un campo de fuerzas clásico sencillo. Según la metodología



empleada pueden, a su vez, clasificarse en métodos estocásticos (Monte Carlo [16]) o deterministas (dinámica molecular [17, 18]). Tradicionalmente estos modelos han asumido la aproximación de la aditividad por pares en el cálculo de las interacciones intermoleculares. Actualmente se sabe que las contribuciones no aditivas pueden ser significativas, por lo que comienza a extenderse el empleo de campos de fuerzas polarizables [19-27].

- **Métodos mixtos.** Este tercer grupo de métodos centra la atención en la descripción de la molécula de soluto, que se representa por su función de estado cuántica, y en cómo ésta se ve afectada por el disolvente, que se trata de manera clásica. Por lo general, los cálculos cuánticos se restringen a una porción del sistema —el soluto, la parte reactiva de una macromolécula, el cromóforo, etc.— mientras que el resto del sistema, considerado como entorno, se representa clásicamente y se introduce como una perturbación en el hamiltoniano molecular del soluto. Por esta razón, este tipo de métodos también se conocen como «métodos de hamiltoniano efectivo» y pueden dividirse en dos grupos principales:

- *Métodos discretos.* En estos métodos, el disolvente se representa de manera discreta, como un conjunto de moléculas con una determinada disposición espacial en torno a la molécula de soluto. Las moléculas de disolvente se consideran como entidades clásicas individuales y el efecto de cada una de ellas —normalmente sólo la interacción electrostática— se introduce en el hamiltoniano molecular del soluto [28-31]. Así pues, para una determinada configuración del disolvente puede efectuarse un cálculo cuántico del soluto, del que se obtienen sus propiedades moleculares. Las energías de interacción y las fuerzas sobre los núcleos, tanto de soluto como de disolvente, también pueden emplearse para obtener una muestra de configuraciones del sistema, mediante las técnicas de simulación de Monte Carlo o de dinámica molecular. Las propiedades termodinámicas se calculan como promedio de los valores obtenidos para cada una de las configuraciones del disolvente consideradas. Estos métodos también son conocidos como métodos combinados de mecánica cuántica y mecánica molecular (*Quantum Mechanics/Molecular Mechanics*, QM/MM).

Esta técnica considera los efectos de polarización del soluto y ha sido empleada en diversos estudios de reactividad en disolución, en otros relativos al efecto del disolvente en transiciones electrónicas [32, 33] y en la determinación de magnitudes energéticas y estructurales de líquidos puros y disoluciones. El elevado número de cálculos cuánticos que se necesita realizar, uno por cada configuración seleccionada del disolvente,

hace necesario introducir aproximaciones adicionales como limitar la calidad de estos cálculos (es frecuente el empleo de métodos semiempíricos y cálculos HF o DFT) o reducir el número de configuraciones del disolvente consideradas.

- *Métodos que emplean la aproximación del campo medio.* A diferencia de los métodos anteriores, éstos asumen que el promedio de las propiedades obtenidas en los distintos cálculos cuánticos —uno por cada configuración del disolvente— puede sustituirse por el resultado de un único cálculo en el que se introduce la «perturbación promedio» del disolvente. De esta manera se reduce de forma considerable el número de cálculos cuánticos a realizar. Se incluyen en este grupo los modelos que representan al disolvente como un continuo dieléctrico sin estructura interna: SCRF (*Self-Consistent Reaction Field*) [34, 35], PCM (*Polarisable Continuum Model*) [36]. También pertenecen a este grupo otros métodos en los que no se obvia la estructura de solvatación próxima al soluto. Esta estructura puede obtenerse por diferentes métodos, como medidas experimentales [37], cálculos mediante teorías integrales (RISM-SCF [38, 39]) o simulaciones (ASEP/MD [40-42]).

Una parte importante de los esfuerzos de la química computacional en los últimos años se ha dirigido a lograr un nivel de precisión y exactitud para los cálculos en fase condensada equivalente al que ya puede obtenerse para moléculas en fase gas. Para este fin, los métodos clásicos resultan inadecuados, ya que no permiten el estudio de procesos químicos, donde los efectos cuánticos son fundamentales. Los métodos cuánticos, por otra parte, resultan aún demasiado costosos computacionalmente para ser aplicados de manera cotidiana. Los métodos mixtos son los más prometedores en este sentido. De entre estos últimos, los métodos discretos o QM/MM requieren también una importante inversión computacional, lo que muchas veces obliga a emplear métodos semiempíricos para efectuar los cálculos cuánticos. Si se desea realizar cálculos *ab initio* en disolución en un lapso de tiempo razonable, la única opción es emplear la aproximación del campo medio. En la última década, nuestro grupo de investigación ha venido desarrollando un método mixto, denominado ASEP/MD (*Averaged Solvent Electrostatic Potential from Molecular Dynamics*), que emplea esta aproximación y que considera un disolvente estructurado.

## 1.2. Objetivos

El presente trabajo de investigación se ha dirigido principalmente a la mejora del método ASEP/MD, previamente desarrollado, para el estudio de la reactividad en disolución. Como ya se ha indicado, este método recurre a la aproximación del

campo medio, que puede introducir ciertos errores en los cálculos. Por otra parte, las versiones existentes del método no permitían la optimización de las estructuras moleculares en disolución, lo que obligaba a emplear datos experimentales u obtenidos por otros métodos. Para el estudio de la reactividad, además de la optimización de geometrías, resulta fundamental el cálculo de energías libres. Por todo ello, los objetivos de esta investigación fueron los siguientes:

1. Comprobar la validez de la aproximación del campo medio tal como se aplica en el método ASEP/MD. Proporcionar también, en la medida de lo posible, algún método para reducir o evaluar los posibles errores introducidos.
2. Desarrollar una técnica que permita el cálculo de energías libres en disolución, especialmente de diferencias de energía libre entre distintas estructuras.
3. Desarrollar un procedimiento para la optimización de geometrías en disolución, tanto estables como de transición, e incorporarlo al método ASEP/MD.
4. Elaborar un programa informático para la aplicación del método que incorpore las mejoras desarrolladas. El programa se pondrá a disposición de la comunidad científica.
5. Aplicar los métodos desarrollados a algunos casos de interés con el fin de comprobar su buen funcionamiento y validez práctica.

### 1.3. Sumario

La organización de la presente tesis doctoral es la que se expone a continuación. En primer lugar se presentan las bases teóricas de los métodos mixtos para el estudio de sistemas en fase líquida. En el capítulo 2 se da una visión general de los métodos cuánticos *ab initio* más empleados. Aunque en los métodos mixtos, sobre todo en los métodos QM/MM, es frecuente el empleo de métodos semiempíricos para el tratamiento cuántico del soluto, estos métodos resultan insuficientes para el estudio de fenómenos como la reactividad o las transiciones electrónicas; por lo tanto, esta investigación se ha centrado en el empleo de métodos *ab initio* para los cálculos cuánticos, y son éstos los únicos que se describen. El capítulo 3 proporciona una descripción de las principales técnicas utilizadas para representar al disolvente en el contexto de los métodos mixtos. En el capítulo 4 se exponen con mayor detalle los fundamentos y las distintas variantes del método desarrollado por nuestro grupo de investigación: ASEP/MD. También se analizan en este capítulo los errores que introduce el empleo de la aproximación del campo medio.

Posteriormente, se reflejan las bases y desarrollos teóricos de las mejoras introducidas en el método ASEP/MD durante el período de la investigación. El capítulo 5 se refiere al cálculo de diferencias de energía libre. En el capítulo 6 se revisan las técnicas de optimización de geometrías y su aplicación al método ASEP/MD.

A continuación se muestran y detallan las distintas aplicaciones de los métodos desarrollados a diferentes sistemas. En el capítulo 7 se exponen las condiciones generales de los cálculos realizados. En los capítulos 8 a 12 se dan más detalles y se analizan los resultados de cada una de las aplicaciones. Las conclusiones de este trabajo se presentan en el capítulo 13.

Finalmente, en los apéndices se muestran algunos aspectos formales concretos que se ha considerado más adecuado no incluir en el texto principal, incluido el código Fortran 90 del programa ASEP/MD elaborado, en el apéndice F. Se incluye también, tras las referencias bibliográficas, una copia de las diferentes publicaciones en revistas científicas internacionales a las que ha dado lugar este trabajo.

---

---

Primera parte

**MÉTODOS MIXTOS**

---

---

∞  
*Para todo problema existe una solución sencilla, ingeniosa y errónea.*  
∞

*Henry-Louis Mencken (1880-1956)*



---

---

## 2

---

# REPRESENTACIÓN CUÁNTICA DEL SOLUTO

A pesar del éxito de la mecánica cuántica, la aplicación rigurosa de sus principios a sistemas reales constituidos por más de dos partículas resulta sumamente compleja. Las ecuaciones que se obtienen suelen carecer de solución analítica y la resolución numérica tampoco es sencilla. Se hace necesario, por lo tanto, introducir algunas aproximaciones que, manteniendo la coherencia con el sentido físico del sistema tratado, permitan obtener resultados válidos en tiempos razonables. Existe una gran cantidad de métodos que permiten aplicar las ideas de la mecánica cuántica de manera más o menos precisa. Los métodos «semiempíricos» introducen gran cantidad de aproximaciones, despreciando el valor de ciertas integrales, sustituyendo otras por parámetros empíricos, etc. Los llamados métodos «*ab initio*» parten de principios teóricos e introducen aproximaciones matemáticas en ciertos conceptos o en la forma de la solución, pero no hacen uso de parámetros experimentales en los cálculos. Este capítulo está dedicado a este último conjunto de métodos. La tarea de escoger uno u otro de estos métodos recae sobre el investigador, quien, en función del sistema y las propiedades que pretende estudiar, debe seleccionar las aproximaciones más convenientes para obtener resultados significativos.

### 2.1. La aproximación de Born-Oppenheimer

La ecuación básica de la química cuántica es la llamada ecuación de Schrödinger, que, en su versión independiente del tiempo, puede escribirse de manera compacta como:

$$\hat{H}\Psi = E\Psi \quad (2.1)$$

donde  $\hat{H}$  es el operador hamiltoniano,  $E$  es la energía total y  $\Psi$  es la función de estado (también llamada «función de onda») del sistema. Esta función de estado tiene normalmente un número de variables independientes muy elevado, al menos tres por cada partícula que integra el sistema.

Se puede escribir el hamiltoniano total (despreciando los efectos relativistas) para un sistema de partículas cargadas (electrones y núcleos) como:

$$\hat{H} = - \sum_a \frac{\hbar^2}{2m_a} \nabla_a^2 + \frac{1}{4\pi\epsilon_0} \sum_{a < b} \frac{q_a q_b}{|\mathbf{r}_a - \mathbf{r}_b|} \quad (2.2)$$

donde los índices  $a$  y  $b$  designan a todas las partículas del sistema, independientemente de su naturaleza,  $\mathbf{r}_a$  son sus posiciones y  $q_a$  sus cargas.  $\nabla_a^2$  el operador laplaciano para la partícula  $a$ :

$$\nabla_a^2 = \nabla_a \cdot \nabla_a = \frac{\partial^2}{\partial x_a^2} + \frac{\partial^2}{\partial y_a^2} + \frac{\partial^2}{\partial z_a^2} \quad (2.3a)$$

$$\nabla_a = \begin{pmatrix} \partial/\partial x_a \\ \partial/\partial y_a \\ \partial/\partial z_a \end{pmatrix} \quad (2.3b)$$

Esta expresión puede simplificarse si se emplea un sistema de unidades en el que tanto  $\hbar^2/m_e$  como  $e^2/4\pi\epsilon_0$  sean iguales a 1, conocido como «unidades atómicas». Si distinguimos entre núcleos y electrones asignando los índices  $\{\alpha, \beta, \dots\}$  a los núcleos e  $\{i, j, \dots\}$  a los electrones, podemos reescribir el hamiltoniano (2.2) de la siguiente manera:

$$\hat{H} = \hat{T}_N + \hat{H}_e \quad (2.4a)$$

$$\hat{T}_N = - \frac{1}{2} m_e \sum_{\alpha} \frac{1}{m_{\alpha}} \nabla_{\alpha}^2 \quad (2.4b)$$

$$\begin{aligned} \hat{H}_e = & - \frac{1}{2} \sum_i \nabla_i^2 - \sum_i \sum_{\alpha} \frac{Z_{\alpha}}{|\mathbf{r}_{\alpha} - \mathbf{r}_i|} + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \\ & + \sum_{\alpha < \beta} \frac{Z_{\alpha} Z_{\beta}}{|\mathbf{r}_{\alpha} - \mathbf{r}_{\beta}|} \end{aligned} \quad (2.4c)$$

donde  $\hat{T}_N$  es el operador energía cinética nuclear, que no depende de las coordenadas electrónicas, y  $\hat{H}_e$  es el operador hamiltoniano electrónico, que depende de las coordenadas nucleares, pero no de las derivadas respecto a éstas, se considera una dependencia paramétrica. Por consiguiente, el problema electrónico puede resolverse considerando los núcleos momentáneamente fijos en determinadas posiciones [43]:

$$\hat{H}_e(\mathbf{r}, \mathbf{R}) \psi_e(\mathbf{r}, \mathbf{R}) = E_e(\mathbf{R}) \psi_e(\mathbf{r}, \mathbf{R}) \quad (2.5)$$



donde  $\mathbf{r}$  y  $\mathbf{R}$  designan, respectivamente, al conjunto de todas las coordenadas electrónicas y nucleares. Tanto la energía como la función de onda siguen siendo funciones de  $\mathbf{R}$ .

Se aproxima entonces la función de estado total mediante un producto,

$$\Psi_{\text{BO}}(\mathbf{r}, \mathbf{R}) = \psi_{\text{N}}(\mathbf{R}) \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) \quad (2.6)$$

siendo la función de onda nuclear  $\psi_{\text{N}}(\mathbf{R})$  una solución de la ecuación:

$$\left[ \hat{T}_{\text{N}}(\mathbf{R}) + E_{\text{e}}(\mathbf{R}) \right] \psi_{\text{N}}(\mathbf{R}) = E \psi_{\text{N}}(\mathbf{R}) \quad (2.7)$$

Desde este punto de vista,  $E_{\text{e}}(\mathbf{R})$  representa, además de la energía electrónica del sistema, la energía potencial que afecta a los movimientos de los núcleos.

Aplicando el hamiltoniano total a la función de onda Born-Oppenheimer y sustituyendo (2.4a) y (2.6) se obtiene:

$$\begin{aligned} \hat{H} \Psi_{\text{BO}} &= \left[ \hat{H}_{\text{e}} + \hat{T}_{\text{N}} \right] \psi_{\text{N}}(\mathbf{R}) \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) = \\ &= \psi_{\text{N}}(\mathbf{R}) \hat{H}_{\text{e}} \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) + \hat{T}_{\text{N}} [\psi_{\text{N}}(\mathbf{R}) \psi_{\text{e}}(\mathbf{r}, \mathbf{R})] = \\ &= \psi_{\text{N}}(\mathbf{R}) E_{\text{e}}(\mathbf{R}) \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) + \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) \hat{T}_{\text{N}} \psi_{\text{N}}(\mathbf{R}) - \\ &\quad - \frac{1}{2} m_{\text{e}} \sum_{\alpha} \frac{1}{m_{\alpha}} \left[ 2 \nabla_{\alpha} \psi_{\text{N}}(\mathbf{R}) \nabla_{\alpha} \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) + \psi_{\text{N}}(\mathbf{R}) \nabla_{\alpha}^2 \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) \right] \end{aligned} \quad (2.8)$$

La función de onda electrónica depende formalmente de las coordenadas nucleares y, por lo tanto, se ve afectada por el operador  $\hat{T}_{\text{N}}$ . Sin embargo, dada la enorme diferencia de masas entre los núcleos y los electrones, los primeros se mueven mucho más lentamente que los segundos. Esto hace que la función de onda electrónica varíe muy poco con los cambios en las coordenadas nucleares, con lo que los términos en los que  $\nabla_{\alpha}$  y  $\nabla_{\alpha}^2$  se aplican a  $\psi_{\text{e}}$  pueden despreciarse:

$$\hat{H} \Psi_{\text{BO}} \simeq \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) \left[ \hat{T}_{\text{N}} + E_{\text{e}}(\mathbf{R}) \right] \psi_{\text{N}}(\mathbf{R}) = \psi_{\text{e}}(\mathbf{r}, \mathbf{R}) E \psi_{\text{N}}(\mathbf{R}) = E \Psi_{\text{BO}} \quad (2.9)$$

Los errores que introduce la aproximación de Born-Oppenheimer son mínimos, y se considera que las soluciones de la parte electrónica (ec. 2.5) son exactas. Normalmente, en los problemas de estructura electrónica, uno se centra en la ecuación (2.5) y se olvida del resto, pero algunas aplicaciones —como la optimización de la estructura molecular— son resultados implícitos de esta aproximación.

## 2.2. Determinantes de Slater

Después de aplicar la aproximación de Born-Oppenheimer puede efectuarse de nuevo una separación de variables, expresando la función de onda electrónica total

mediante un producto de funciones de onda monoeléctricas [44-46]:

$$\psi_e(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \varphi_1(\mathbf{r}_1) \varphi_2(\mathbf{r}_2) \cdots \varphi_n(\mathbf{r}_n) \quad (2.10)$$

En la ecuación (2.10), las funciones  $\varphi_i$  describen un único electrón, son los llamados «orbitales»; además, en esta ecuación se ha eliminado la dependencia con  $\mathbf{R}$ , se supone que la ecuación se resuelve para una determinada configuración nuclear.

Esta aproximación presenta el problema de que no cumple con el Principio de Exclusión de Pauli, según el cual la función de onda polieletrónica debe ser antisimétrica con respecto al intercambio de coordenadas de dos electrones cualesquiera, es decir:

$$\forall i, j \neq i \quad \hat{P}_{ij} \psi_e = -\psi_e \quad (2.11)$$

donde  $\hat{P}_{ij}$  es operador permutación que intercambia las coordenadas de las partículas  $i$  y  $j$ . La manera más común de adaptar la función de onda para que cumpla con el Principio de Exclusión es utilizar un «determinante de Slater» [47]:

$$\psi_e = \frac{1}{\sqrt{n!}} \begin{vmatrix} \varphi_1(\mathbf{r}_1) & \varphi_1(\mathbf{r}_2) & \cdots & \varphi_1(\mathbf{r}_n) \\ \varphi_2(\mathbf{r}_1) & \varphi_2(\mathbf{r}_2) & \cdots & \varphi_2(\mathbf{r}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_n(\mathbf{r}_1) & \varphi_n(\mathbf{r}_2) & \cdots & \varphi_n(\mathbf{r}_n) \end{vmatrix} = |\varphi_1 \varphi_2 \cdots \varphi_n\rangle \quad (2.12)$$

donde el término  $1/\sqrt{n!}$  asegura que, si los  $\varphi_i$  están normalizados,  $\psi_e$  también lo estará.

El hamiltoniano electrónico (ec. 2.4c) puede escribirse de una manera más compacta como:

$$\hat{H}_e = \hat{h}_0 + \sum_i \hat{h}_i + \sum_{i < j} \hat{g}_{ij} \quad (2.13)$$

$$\hat{h}_0 = \sum_{\alpha < \beta} \sum \frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} \quad \hat{h}_i = -\frac{1}{2} \nabla_i^2 - \sum_\alpha \frac{Z_\alpha}{r_{i\alpha}} \quad \hat{g}_{ij} = \frac{1}{r_{ij}}$$

El término  $\hat{h}_0$  es una constante para posiciones nucleares fijas; los distintos  $\hat{h}_i$  dependen de las coordenadas de un solo electrón; el operador bielectrónico  $\hat{g}_{ij}$  describe la interacción entre dos electrones. Es de destacar que no existe interacción física directa entre las partículas elementales que implique a más de dos partículas, en esta escala fundamental las interacciones son aditivas; esto no se cumple cuando se consideran sistemas más complejos —como átomos o moléculas— de forma colectiva.

La energía correspondiente a una determinada función de onda puede calcularse como el valor esperado del hamiltoniano, que para una función normalizada

viene dado por:

$$E_e = \langle \psi_e | \hat{H}_e | \psi_e \rangle = \hat{h}_0 + \langle \psi_e | \sum_i \hat{h}_i | \psi_e \rangle + \langle \psi_e | \sum_{i < j} \hat{g}_{ij} | \psi_e \rangle \quad (2.14)$$

Esta ecuación puede expresarse en función de los orbitales  $\varphi_i$ . A partir de la forma de la función  $\psi_e$  (un determinante de Slater) y de las propiedades de los operadores  $\hat{h}_i$  y  $\hat{g}_{ij}$  (operadores mono y bielectrónico), puede comprobarse [48] que la mayoría de los términos se anulan, y queda:

$$\begin{aligned} E_e &= \hat{h}_0 + \sum_i \langle \varphi_i | \hat{h} | \varphi_i \rangle + \sum_{i < j} (\langle \varphi_i \varphi_j | \hat{g} | \varphi_i \varphi_j \rangle - \langle \varphi_i \varphi_j | \hat{g} | \varphi_j \varphi_i \rangle) = \\ &= \hat{h}_0 + \sum_i \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_i \sum_j \langle ij | ij \rangle \end{aligned} \quad (2.15)$$

Otra forma de escribir esta ecuación se obtiene empleando los operadores de Coulomb,  $\hat{J}$ , e intercambio,  $\hat{K}$ , que se definen como:

$$\begin{aligned} \hat{J}_j f(\mathbf{r}_1) &= \left[ \int \frac{\varphi_j^*(\mathbf{r}_2) \varphi_j(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 \right] f(\mathbf{r}_1) & \hat{J} &= \sum_j \hat{J}_j \\ \hat{K}_j f(\mathbf{r}_1) &= \left[ \int \frac{\varphi_j^*(\mathbf{r}_2) f(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 \right] \varphi_j(\mathbf{r}_1) & \hat{K} &= \sum_j \hat{K}_j \end{aligned} \quad (2.16)$$

de manera que la contribución bielectrónica —el último término de la ecuación (2.15)— se puede escribir de la forma:

$$\frac{1}{2} \sum_i \sum_j \langle ij | ij \rangle = \frac{1}{2} \sum_i \langle i | \hat{J} - \hat{K} | i \rangle \quad (2.17)$$

La ecuación (2.15) nos da el valor de la energía electrónica total de una función de onda con forma de determinante de Slater, dentro de los límites de la aproximación de Born-Oppenheimer y del hamiltoniano no relativista.

## 2.3. El método Hartree-Fock

Una vez que se tiene una expresión para calcular la energía en función de los distintos orbitales  $\varphi_i$ , el siguiente paso es calcular la forma de estos orbitales, que por el momento son desconocidos. Para ello se aplica el principio variacional [49] a la expresión (2.15). Al variar los orbitales en una cantidad infinitesimal,  $\delta\varphi_i$ , se produce un cambio en la función de onda  $\delta\psi_e$  y un cambio en la energía que se

puede calcular:

$$\begin{aligned}
E[\psi_e + \delta\psi_e] &= \hat{h}_0 + \sum_i \langle \varphi_i + \delta\varphi_i | \hat{h} | \varphi_i + \delta\varphi_i \rangle + \\
&+ \frac{1}{2} \sum_i \sum_j \langle (\varphi_i + \delta\varphi_i)(\varphi_j + \delta\varphi_j) | \hat{g} | (\varphi_i + \delta\varphi_i)(\varphi_j + \delta\varphi_j) \rangle - \\
&- \frac{1}{2} \sum_i \sum_j \langle (\varphi_i + \delta\varphi_i)(\varphi_j + \delta\varphi_j) | \hat{g} | (\varphi_j + \delta\varphi_j)(\varphi_i + \delta\varphi_i) \rangle
\end{aligned} \tag{2.18}$$

Con esta expresión, la variación de primer orden de la energía,  $\delta^{(1)}E[\psi_e]$ , se obtiene como:

$$\begin{aligned}
\delta^{(1)}E[\psi_e] &= \sum_i \langle \delta\varphi_i | \hat{h} | \varphi_i \rangle + \sum_i \sum_j [\langle \delta\varphi_i \varphi_j | \hat{g} | \varphi_i \varphi_j \rangle - \langle \delta\varphi_i \varphi_j | \hat{g} | \varphi_j \varphi_i \rangle] + \text{c.c.} = \\
&= \sum_i \langle \delta\varphi_i | \hat{h} | \varphi_i \rangle + \sum_i \langle \delta\varphi_i | \hat{J} - \hat{K} | \varphi_i \rangle + \text{c.c.} = \\
&= \sum_i \langle \delta\varphi_i | \hat{F} | \varphi_i \rangle + \text{c.c.}
\end{aligned} \tag{2.19}$$

donde c.c. sustituye a la expresión conjugada compleja de todo lo anterior. El operador  $\hat{F}$  es el llamado «operador de Fock», que puede expresarse como combinación de operadores mono y bielectrónicos:

$$\hat{F} = \hat{h} + \hat{J} - \hat{K} \tag{2.20}$$

Así,  $E[\psi_e]$  será mínima cuando la variación  $\delta^{(1)}E$  se anule, lo que implica que, para cualquier variación permitida en los orbitales,  $\delta\varphi_i$ , debe cumplirse:

$$\langle \delta\varphi_i | \hat{F} | \varphi_i \rangle = 0 \tag{2.21}$$

considerando como variaciones permitidas las que conservan la ortonormalidad de los orbitales  $\varphi_i$ .

Expresando  $\delta\varphi_i$  en función de los distintos  $\varphi_k$  y haciendo una transformación adecuada de orbitales ( $\varphi_i \rightarrow \tilde{\varphi}_i$ ) se llega a las «ecuaciones Hartree-Fock canónicas»:

$$\hat{F} \tilde{\varphi}_i = \epsilon_i \tilde{\varphi}_i \tag{2.22}$$

Los orbitales que se obtienen son los «orbitales Hartree-Fock canónicos» y los valores propios  $\epsilon_i$  se conocen como «energías orbitales». Conviene destacar que la energía total no es igual a la suma de las energías orbitales, de hecho:

$$E_e = \sum_i \epsilon_i - \frac{1}{2} \sum_i \langle \tilde{\varphi}_i | \hat{J} - \hat{K} | \tilde{\varphi}_i \rangle \tag{2.23}$$

### 2.3.1. Combinación lineal de orbitales atómicos

Aunque la ecuación (2.22) proporciona, en principio, una manera de obtener los orbitales moleculares  $\tilde{\varphi}_i$ , sigue siendo prácticamente imposible de resolver sin más datos sobre la forma funcional de dichos orbitales. La situación mejora si se expresan los orbitales moleculares como combinación lineal de funciones de base («orbitales atómicos»)  $\chi_p(\mathbf{r})$  [50, 51]:

$$\tilde{\varphi}_i(\mathbf{r}) = \sum_{p=1}^m C_{pi} \chi_p(\mathbf{r}) \quad (2.24)$$

Las funciones  $\chi_p$  pueden escogerse de manera más o menos arbitraria, aunque los resultados finales dependerán del conjunto concreto que se tome. Existen diferentes conjuntos de funciones de base cuya utilidad para diversos sistemas químicos ha sido ampliamente comprobada.

Sustituyendo (2.24) en (2.22) se obtiene la expresión:

$$\hat{F} \sum_p C_{pi} \chi_p = \epsilon_i \sum_p C_{pi} \chi_p \quad (2.25)$$

conocida como la ecuación de Roothaan-Hall. Esta ecuación puede escribirse en forma matricial:

$$F C = S C \epsilon \quad (2.26)$$

$$F_{ij} = \langle \chi_i | \hat{F} | \chi_j \rangle \quad S_{ij} = \langle \chi_i | \chi_j \rangle \quad \epsilon_{ij} = \delta_{ij} \epsilon_i$$

La matriz de los coeficientes,  $C$ , es la representación de los orbitales moleculares en la base de orbitales atómicos  $\chi_i$ . De la misma manera, la matriz  $F$  representa al operador de Fock.  $S$  es la matriz de solapamiento; si las funciones de base son ortonormales,  $S = I$ . En el caso de que las funciones de base no sean ortonormales, siempre pueden transformarse en un conjunto que sí lo sea. En una base de estas características, la ecuación (2.26) queda:

$$F' C' = C' \epsilon \quad (2.27)$$

que es una ecuación de valores propios matricial, para cuya resolución existen diversos métodos.

### 2.3.2. El método del campo autoconistente

Para resolver la ecuación (2.27) basta, en principio, con conocer la matriz  $F'$ ; sin embargo, esta matriz se obtiene a partir del operador de Fock, que depende de los orbitales (ya que los operadores  $\hat{J}$  y  $\hat{K}$  están definidos en función de los

orbitales del sistema, según la ecuación 2.16). Se da, por tanto, la paradoja de que para poder plantear las ecuaciones hay que conocer de antemano la solución del sistema. Para resolver este problema se parte de unos orbitales arbitrarios, a partir de los cuales se define un operador de Fock aproximado y la matriz  $F$ , que se diagonaliza para obtener un nuevo conjunto de orbitales,  $C$ . Con estos nuevos orbitales puede construirse un nuevo operador de Fock y continuar el proceso. Tras un cierto número de iteraciones suele ocurrir que los orbitales ya no cambian, en este momento se cumple la ecuación (2.22) y podemos considerar que se ha obtenido la solución.

Éste es el llamado método del «campo autocoherente» (*Self-Consistent Field*, SCF), puesto que los campos de Coulomb y de intercambio de los orbitales, en forma de los operadores  $\hat{J}$  y  $\hat{K}$ , definen un operador de Fock cuyas funciones propias son esos mismos orbitales.

## 2.4. Correlación electrónica

La principal aproximación que introduce el método Hartree-Fock es considerar que cada electrón se mueve en el campo promedio que genera el resto de los electrones. En realidad, el movimiento de los electrones depende de las posiciones instantáneas de todos los demás electrones, es decir, está «correlacionado». El método Hartree-Fock no tiene en cuenta esta correlación y, por lo tanto, produce resultados que no son los exactos (aun sin tener en cuenta los errores numéricos y los efectos relativistas); el error que se comete en la energía recibe el nombre de «energía de correlación»:

$$E_e^{\text{corr}} = E_e^{\text{exacta}} - E_e^{\text{HF}} \quad (2.28)$$

La energía de correlación no es demasiado importante cuando se estudian sistemas en estado fundamental y con geometría de equilibrio, pero no puede despreciarse en el caso de que se estudien fenómenos en los que tenga lugar una redistribución electrónica importante (potencial de ionización, energías de excitación, etc.), reacciones químicas con ruptura o formación de enlaces o sistemas donde la componente de dispersión sea significativa.

Para estimar la energía de correlación existe una amplia variedad de métodos que difieren en el grado de aproximación utilizado y en el coste computacional que implican. De entre todos éstos, podemos destacar el método perturbativo de Møller-Plesset (MP), los métodos de «interacción de configuraciones» (*Configuration Interaction*, CI), los métodos «multiconfiguracionales» (*Multiconfigurational SCF*, MCSCF) y los métodos de «grupos acoplados» (*Coupled Cluster*, CC).

### 2.4.1. Métodos Møller-Plesset

Los métodos Møller-Plesset (MP) [52] están basados en la teoría de perturbaciones de Rayleigh-Schrödinger, empleando la siguiente partición del hamiltoniano:

$$\begin{aligned}\hat{H}_e &= \hat{F} + \hat{V} \\ \hat{F} &= \hat{h} + \hat{J} - \hat{K} \quad \hat{V} = \sum_{i < j} \sum \frac{1}{r_{ij}} - (\hat{J} - \hat{K})\end{aligned}\quad (2.29)$$

es decir, se representa el hamiltoniano electrónico mediante el operador de Fock más una perturbación, que corresponde a la diferencia entre el operador bielectrónico exacto y el contenido en el operador de Fock. La energía y la función de onda se expanden en series según el orden de la perturbación:

$$\begin{aligned}\psi_e &= \psi_e^{(0)} + \psi_e^{(1)} + \psi_e^{(2)} + \dots \\ E_e &= E_e^{(0)} + E_e^{(1)} + E_e^{(2)} + \dots\end{aligned}\quad (2.30)$$

Las series se truncan en cierto orden,  $n$ , dando origen a los métodos denominados MP $n$ .

Con la partición de la ecuación (2.29) se tiene que:

- La energía de orden cero,  $E_e^{(0)}$ , es simplemente la suma de las energías orbitales,  $\epsilon_i$ .
- La energía de orden uno,  $E_e^{(0)} + E_e^{(1)}$ , es la energía Hartree-Fock ( $E_e^{\text{HF}}$ ).
- La suma del resto de los términos,  $E_e^{(2)} + \dots + E_e^{(\infty)}$ , es la energía de correlación ( $E_e^{\text{corr}}$ ).
- La función de onda de orden cero,  $\psi_e^{(0)}$ , es la función de onda Hartree-Fock.
- A partir de la función de onda de orden  $n$  puede calcularse la corrección a la energía de orden hasta  $2n + 1$ .

La versión más sencilla de estos métodos es la que calcula las energías de segundo orden (MP2). En este caso, la energía viene dada por:

$$\begin{aligned}E_e^{\text{MP2}} &= E_e^{\text{HF}} - \frac{1}{4} \sum_r \sum_s \sum_i \sum_j (a_{ij}^{rs})^2 (\epsilon_r + \epsilon_s - \epsilon_i - \epsilon_j) \\ a_{ij}^{rs} &= -a_{ij}^{sr} = -a_{ji}^{rs} = -a_{rs}^{ij} = \frac{\langle rs || ij \rangle}{\epsilon_r + \epsilon_s - \epsilon_i - \epsilon_j}\end{aligned}\quad (2.31)$$

donde los subíndices  $\{i, j\}$  denotan a los orbitales ocupados y los subíndices  $\{r, s\}$  a los orbitales virtuales. Por su parte, la corrección de orden uno a la función de

onda puede calcularse como:

$$\psi_e^{(1)} = \sum_{r < s} \sum_{i < j} a_{ij}^{rs} \psi_{ij}^{rs} \quad (2.32)$$

donde  $\psi_{ij}^{rs}$  representa al determinante de Slater resultante de sustituir, en  $\psi_e^{\text{HF}}$ , los orbitales  $\varphi_i$  y  $\varphi_j$  por  $\varphi_r$  y  $\varphi_s$ , respectivamente. El método MP2 es uno de los más empleados para el cálculo de las energías de correlación.

Existen fórmulas equivalentes para calcular las correcciones de órdenes superiores: MP3, MP4, etc., pero en general son computacionalmente muy exigentes y no proporcionan mejores resultados que otros métodos.

### 2.4.2. Interacción de configuraciones

En los métodos de interacción de configuraciones [53, 54] se representa la función de onda total no mediante un único determinante de Slater —como en el método Hartree-Fock—, sino mediante una combinación lineal de determinantes:

$$\psi_e^{\text{CI}} = c_0 \psi_e^{\text{HF}} + \sum_i \sum_r c_i^r \psi_i^r + \sum_{r < s} \sum_{i < j} c_{ij}^{rs} \psi_{ij}^{rs} + \dots \quad (2.33)$$

donde  $\psi_i^r$  son los determinantes resultantes de monoexcitaciones (excitaciones de un único electrón) con respecto al determinante de referencia,  $\psi_e^{\text{HF}}$ ;  $\psi_{ij}^{rs}$  son los resultantes de biexcitaciones, etc. Los coeficientes,  $c_i^r$ ,  $c_{ij}^{rs}$ , etc., se optimizan mediante un cálculo variacional de manera que el valor de  $E_e$  sea mínimo.

Si se incluyen todas las configuraciones posibles con los  $N$  electrones del sistema, se trata de un cálculo de interacción de configuraciones completo (*full CI*), que proporcionará la función de onda y propiedades exactas dentro de las aproximaciones introducidas (base no completa, hamiltoniano no relativista...). En la práctica, sin embargo, es necesario truncar la expresión (2.33). Generalmente esto se hace en las biexcitaciones, se incluyen, por tanto, en la función de onda todas las configuraciones electrónicas resultantes de excitaciones de uno o dos electrones con respecto al determinante Hartree-Fock (además de éste mismo); este método recibe el nombre de CISD (*CI with Single and Double excitations*). Existen otros métodos que incluyen mayor número de configuraciones en la función de onda: CISDT (excitaciones triples), CISDTQ (excitaciones cuádruples), etc.

El empleo de métodos CI truncados, como los señalados, presenta el problema de que se pierde la uniformidad del método con el tamaño del sistema (se dice que el método no es *size-consistent*). Esto quiere decir que la «calidad» del método depende del tamaño del sistema estudiado. Por ejemplo, en un sistema de  $N$  componentes idénticos y aislados entre sí la energía total debe  $N$  veces la energía de uno solo de los componentes. Para un método CISD, sin embargo, esto



no se cumple, ya que para describir el sistema completo con el mismo detalle que cada componente por separado sería necesario incluir configuraciones resultantes de excitaciones de  $2N$  electrones, con lo que el método dejaría de ser CISD. Este problema puede evitarse empleando la teoría de grupos acoplados, como se verá más adelante.

Es conveniente destacar que, a pesar de que se optimizan los coeficientes que multiplican a las diferentes configuraciones, los coeficientes que definen los orbitales moleculares se obtienen a partir de un cálculo Hartree-Fock y se mantienen fijos.

### 2.4.3. Métodos multiconfiguracionales

En muchas ocasiones la estructura electrónica de la molécula estudiada es tal que no puede describirse adecuadamente mediante un único determinante de Slater, ni siquiera la función de onda CI (al menos con un número asequible de determinantes) soluciona el problema, puesto que los orbitales moleculares siguen siendo los obtenidos en un cálculo Hartree-Fock y han sido optimizados para una función de onda de un solo determinante.

En estos casos es necesario emplear un nuevo tipo de métodos, los métodos multiconfiguracionales [55]. En éstos, la función de onda se representa mediante una combinación lineal de determinantes de Slater:

$$\psi_e^{\text{MC}} = \sum_k c_k |k\rangle \quad |k\rangle = |\varphi_1^k \varphi_2^k \cdots \varphi_n^k\rangle \quad (2.34)$$

donde cada determinante  $|k\rangle$  corresponde a una configuración (conjunto de orbitales  $\varphi_i^k$ ). La principal diferencia con los métodos CI es que en este caso se optimizan tanto los coeficientes que multiplican a los distintos determinantes,  $c_k$ , como los que definen a los orbitales moleculares  $\varphi_i$  en función de los orbitales atómicos  $\chi_p$ ,  $C_{pi}$ . De esta manera los orbitales que se obtienen son los óptimos para una función de onda de la forma especificada en la ecuación (2.34).

Para obtener resultados aceptables es necesario escoger adecuadamente las configuraciones que mayor influencia pueden tener para formar la función de onda, lo que no siempre es sencillo. Una posibilidad es utilizar el método del «espacio activo completo» (*Complete Active Space SCF*, CASSCF) [56]. En este método se dividen los orbitales moleculares en tres grupos: orbitales inactivos, orbitales activos y orbitales externos. Los orbitales inactivos están completamente ocupados en todas las configuraciones consideradas, los orbitales externos son siempre orbitales virtuales, es la ocupación de los orbitales activos lo que diferencia unas configuraciones de otras. En un cálculo CASSCF( $n,m$ ) se incluyen en la función de onda todas las configuraciones posibles (conservando el espín total) de  $n$  electrones en los  $m$  orbitales del

espacio activo. La elección de los orbitales que formarán parte del espacio activo dependerá del tipo de molécula, propiedad o fenómeno que se desee estudiar; por ejemplo, en el caso de moléculas conjugadas el espacio activo debe incluir al menos los orbitales  $\pi$  de valencia y los pares libres.

La función de onda MCSCF, una vez obtenida, puede mejorarse del mismo modo que se mejora la función Hartree-Fock. Si a la función CASSCF se le aplica un tratamiento perturbativo similar al MP2, se obtiene el método llamado CASPT2 (*Complete Active Space 2nd order Perturbation Theory*) [57, 58]. También puede efectuarse un cálculo de interacción de configuraciones tomando la función de onda MCSCF como referencia, lo que se denomina MRCI (*Multireference CI*).

#### 2.4.4. Métodos de grupos acoplados

Como ya se ha indicado, uno de los problemas de los métodos CI es que no son uniformes con el tamaño del sistema. Los métodos de grupos acoplados permiten evitar este problema. La función de onda CI (ec. 2.33) puede expresarse también como:

$$\begin{aligned} \psi_e^{\text{CI}} &= \left(1 + \hat{C}_1 + \hat{C}_2 + \dots\right) \psi_e^0 = \left(1 + \hat{C}\right) \psi_e^0 \\ \hat{C}_1 \psi_e^0 &= \sum_i \sum_r c_i^r \psi_i^r & \hat{C}_2 \psi_e^0 &= \sum_{i < j} \sum_{r < s} c_{ij}^{rs} \psi_{ij}^{rs} \end{aligned} \quad (2.35)$$

donde  $\psi_e^0$  es la función de referencia, generalmente la función Hartree-Fock,  $\psi_e^{\text{HF}}$ ;  $\hat{C}_1$  es un operador que genera, con sus coeficientes correspondientes, todas las configuraciones resultantes de la excitación de un solo electrón,  $\hat{C}_2$  las de dos electrones, etc. El operador  $\hat{C}$  genera todas las configuraciones posibles, excepto la de referencia.

En los métodos de grupos acoplados [59, 60] se parte de una función análoga, pero con forma exponencial:

$$\psi_e^{\text{CC}} = \exp(\hat{D}) \psi_e^0 \quad \hat{D} = \hat{D}_1 + \hat{D}_2 + \dots \quad (2.36)$$

expresión que puede expandirse en una serie:

$$\begin{aligned} \psi_e^{\text{CC}} &= \left(1 + \hat{D}_1 + \frac{1}{2!} \hat{D}_1^2 + \hat{D}_2 + \frac{1}{3!} \hat{D}_1^3 + \hat{D}_1 \hat{D}_2 + \hat{D}_3 + \dots\right) \psi_e^0 \\ \hat{D}_1 \psi_e^0 &= \sum_i \sum_r d_i^r \psi_i^r & \hat{D}_2 \psi_e^0 &= \sum_{i < j} \sum_{r < s} d_{ij}^{rs} \psi_{ij}^{rs} \end{aligned} \quad (2.37)$$

Se comprueba que aparecen, además de los términos lineales,  $\hat{D}_1$ ,  $\hat{D}_2$ , etc., términos de mayor orden que contribuyen a niveles de excitación superiores,  $\hat{D}_1^2$ ,  $\hat{D}_1 \hat{D}_2$ , etc. Al desarrollar los distintos términos puede obtenerse la relación que existe entre

los correspondientes coeficientes en las expansiones (2.35) y (2.37):

$$c_i^r = d_i^r \quad c_{ij}^{rs} = d_{ij}^{rs} + d_i^r d_j^s - d_i^s d_j^r \quad (2.38)$$

La función de onda  $\psi_e^{\text{CC}}$  es equivalente a la de un método CI completo, aunque su forma es más complicada que la de  $\psi_e^{\text{CI}}$ . Cuando se emplean expresiones truncadas para  $\hat{D}$ , se mantiene la uniformidad con el tamaño del sistema debido a que, aunque se trunque el operador  $\hat{D}$  en un determinado orden, en la función de onda continúan apareciendo términos de orden superior —denominados «grupos inco-nexos» (*disconnected cluster*)— cuyos coeficientes,  $d_x^y$ , son producto de los términos inferiores.

Las ecuaciones se plantean normalmente de la siguiente manera:

$$\begin{aligned} \hat{H} \exp(\hat{D})\psi_e^0 &= E_e \exp(\hat{D})\psi_e^0 \\ \exp(-\hat{D})\hat{H} \exp(\hat{D})\psi_e^0 &= E_e \psi_e^0 \end{aligned} \quad (2.39)$$

y se aprovecha la expansión de Campbell-Baker-Hausdorff [61-63]:

$$\begin{aligned} \exp(-\hat{D})\hat{H} \exp(\hat{D}) &= \hat{H} + [\hat{H}, \hat{D}] + \frac{1}{2!} [[\hat{H}, \hat{D}], \hat{D}] + \\ &+ \frac{1}{3!} [[[ \hat{H}, \hat{D} ], \hat{D} ], \hat{D}] + \frac{1}{4!} [[[[ \hat{H}, \hat{D} ], \hat{D} ], \hat{D} ], \hat{D}] \end{aligned} \quad (2.40)$$

donde  $[\hat{H}, \hat{D}]$  es el conmutador  $\hat{H}\hat{D} - \hat{D}\hat{H}$ . Puesto que el operador hamiltoniano contiene sólo componentes mono y bielectrónicas, esta expansión no es infinita, se trata de una expresión exacta que contiene únicamente cinco términos, lo que permite plantear ecuaciones de como máximo cuarto grado en las amplitudes (coeficientes  $d_x^y$ ). No obstante, la extensión del operador  $\hat{D}$  es excesiva y es necesario truncarla. Según el orden de las excitaciones que se consideren se generan distintos métodos, así se tienen, por ejemplo, los métodos CCD, CCSD, CCSD(T), donde (T) indica que se realiza un cálculo aproximado de las excitaciones triples, etc.

## 2.5. La teoría del funcional de la densidad

Todos los métodos mencionados hasta ahora se basan en aproximaciones a la función de onda electrónica,  $\psi_e$ , que es una función que depende de  $3n$  coordenadas espaciales ( $n$  es el número de electrones). Sin embargo, toda la información que proporciona la función de onda también puede proporcionarla la densidad electrónica,  $\rho(\mathbf{r})$ , que es función de solamente 3 coordenadas espaciales y, además, tiene un sentido físico claro y es observable. Por lo tanto, todas las propiedades —entre ellas la energía— del sistema pueden determinarse a partir de la densidad electrónica. Ésta es la base de la «Teoría del funcional de la densidad» (*Density Functional Theory*, DFT) [64].

Hohenberg y Kohn [65] desarrollaron los dos teoremas sobre los que se apoya toda la metodología DFT:

**Teorema 1.** *La densidad electrónica  $\rho(\mathbf{r})$  determina el potencial externo.*

Esto es, en la forma de la densidad electrónica están implícitas las posiciones y cargas de los núcleos, así como cualquier potencial exterior al sistema. Además, la integración de  $\rho(\mathbf{r})$  en todo el espacio debe proporcionar el número de electrones, por lo que el sistema está completamente determinado por la densidad electrónica.

**Teorema 2.** *La energía del estado fundamental puede obtenerse variacionalmente: la densidad que minimiza la energía total es la densidad exacta del estado fundamental.*

Es decir, cualquier densidad electrónica de prueba,  $\tilde{\rho}(\mathbf{r})$ , dará una energía mayor que la energía exacta del estado fundamental. Mediante un proceso de minimización similar al desarrollado en el método Hartree-Fock se puede obtener la densidad electrónica exacta y, por lo tanto, la energía exacta del estado fundamental.

La principal dificultad en el desarrollo de esta teoría es la obtención del funcional que relaciona a la energía con la densidad electrónica. En primera instancia este funcional puede expresarse como:

$$E[\rho] = \int v(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} + T[\rho] + V_{ee}[\rho] \quad (2.41)$$

donde  $v(\mathbf{r})$  es el potencial externo a los electrones (incluyendo los núcleos y cualquier perturbación exterior),  $T[\rho]$  es la energía cinética de los electrones y  $V_{ee}[\rho]$  es la energía de interacción electrón-electrón. Son estos dos últimos términos los que presentan los mayores problemas.

### 2.5.1. La formulación de Kohn-Sham

Kohn y Sham [66] introdujeron la idea de considerar un sistema de referencia consistente en  $n$  electrones no interaccionantes. Se supone que los electrones están definidos por orbitales monoelectrónicos ortonormales  $\phi_i(\mathbf{r})$  y la función de onda es un determinante de Slater; en tal caso la densidad electrónica viene dada por:

$$\rho(\mathbf{r}) = \sum_i^n |\phi_i(\mathbf{r})|^2 \quad (2.42)$$

y la energía cinética, por:

$$T_s[\rho] = -\frac{1}{2} \sum_i^n \langle \phi_i(\mathbf{r}) | \nabla^2 | \phi_i(\mathbf{r}) \rangle \quad (2.43)$$

esta energía cinética es la correspondiente al sistema de referencia, pero puede ser una buena aproximación a la energía cinética real,  $T[\rho]$ .

Los orbitales, por su parte, se obtienen al resolver ecuaciones de la forma:

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right] \phi_i = \epsilon_i \phi_i \quad (2.44)$$

donde  $v_s(\mathbf{r})$  es el potencial externo que se introduce en el sistema de referencia, que no tiene por qué ser igual al del sistema real,  $v(\mathbf{r})$ . Los orbitales  $\phi_i(\mathbf{r})$  no tienen, en principio, ningún sentido físico, son sólo un instrumento con el que obtener la densidad electrónica  $\rho(\mathbf{r})$

Se intenta entonces encontrar una expresión para  $v_s(\mathbf{r})$  que permita obtener, para el sistema de referencia, una densidad electrónica igual a la del sistema real. Para ello se introduce la siguiente partición de la energía:

$$\begin{aligned} T[\rho] + V_{ee}[\rho] &= T_s[\rho] + J[\rho] + E_{xc}[\rho] & (2.45) \\ J[\rho] &= \frac{1}{2} \iint \frac{1}{r_{12}} \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ E_{xc}[\rho] &= (T[\rho] - T_s[\rho]) + (V_{ee}[\rho] - J[\rho]) \end{aligned}$$

$J[\rho]$  es la energía de interacción electrostática de los electrones y  $E_{xc}[\rho]$  es la llamada energía de «correlación-intercambio» e incluye todos los efectos no contemplados en  $T_s[\rho]$  ni en  $J[\rho]$ .

Se puede demostrar que, si en la ecuación (2.44) se introduce:

$$\begin{aligned} v_s(\mathbf{r}) &= v(\mathbf{r}) + \frac{\delta J[\rho]}{\delta \rho} + \frac{\delta E_{xc}[\rho]}{\delta \rho} = v(\mathbf{r}) + v_J(\mathbf{r}) + v_{xc}(\mathbf{r}) & (2.46) \\ v_J(\mathbf{r}) &= \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r}' \end{aligned}$$

entonces se cumple el objetivo planteado y la densidad obtenida mediante (2.42) es idéntica a la densidad exacta del sistema real. Es decir, resolviendo las ecuaciones monoeléctricas:

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + v_J(\mathbf{r}) + v_{xc}(\mathbf{r})\right] \phi_i = \epsilon_i \phi_i \quad (2.47)$$

se obtienen unos orbitales  $\phi_i$  (expresados como combinación lineal de funciones de base) que proporcionan la densidad electrónica buscada. En esta ecuación, tanto  $v_J(\mathbf{r})$  como  $v_{xc}(\mathbf{r})$  dependen de  $\rho(\mathbf{r})$ , por lo tanto el sistema ha de resolverse iterativamente, igual que ocurre en el método Hartree-Fock. De hecho, las ecuaciones resultantes son idénticas a las ecuaciones Hartree-Fock si en éstas se sustituye el término de intercambio por  $v_{xc}(\mathbf{r})$ .

Hasta este punto, la Teoría del funcional de la densidad es exacta, así como

el método Kohn-Sham, sin embargo, la forma del potencial de correlación-intercambio,  $v_{xc}(\mathbf{r})$ , no es conocida y probablemente no llegue a serlo nunca, por lo que es necesario emplear potenciales aproximados y muchas veces con parámetros ajustables. Esto convierte a los métodos DFT en semiempíricos, a pesar de estar basados en principios exactos.

### 2.5.2. El funcional de correlación-intercambio

Como se ha mencionado, el principal problema —aunque también la fuente de variedad— de los métodos DFT es el potencial de correlación-intercambio. No hay manera actualmente de obtener la expresión exacta, pero sí pueden obtenerse aproximaciones basándose en las soluciones de sistemas sencillos e introduciendo parámetros cuyo valor, a falta de otra posibilidad, se calcula por ajuste a datos experimentales u obtenidos con otros métodos.

#### – La aproximación de la densidad local

Un sistema sencillo para el que existe una solución prácticamente exacta es el gas de electrones uniforme. Se define como un número grande,  $N \rightarrow \infty$ , de electrones en un cubo de volumen  $V = l^3 \rightarrow \infty$ , por el que está distribuida una carga positiva uniforme suficiente para hacer neutro al sistema, y de forma que la densidad  $\rho = N/V$  se mantiene constante. Los resultados que se obtienen son:

$$E_{xc}^{\text{LDA}}[\rho] = \int \varepsilon_{xc}^{\text{LDA}}[\rho] \rho(\mathbf{r}) d\mathbf{r} \quad \varepsilon_{xc}^{\text{LDA}}[\rho] = \varepsilon_x^{\text{LDA}}[\rho] + \varepsilon_c^{\text{LDA}}[\rho] \quad (2.48)$$

$$\varepsilon_x^{\text{LDA}}[\rho] = -\frac{3}{4} \left( \frac{3\rho(\mathbf{r})}{\pi} \right)^{1/3} \quad (2.49)$$

donde  $\varepsilon_x^{\text{LDA}}[\rho]$  es la energía de intercambio por electrón, que da lugar al funcional de intercambio de Slater-Dirac [67, 68]. La componente de correlación es más complicada y se ha calculado por métodos numéricos [69]. Vosko, Wilk y Nusair [70] propusieron un funcional de la siguiente forma:

$$\varepsilon_c^{\text{VWN}}[\rho] = A \left[ \ln \frac{x^2}{X(x)} + \frac{2b}{Q} P - \frac{b x_0}{X(x_0)} \left( \ln \frac{(x-x_0)^2}{X(x)} + \frac{2(b+2x_0)}{Q} P \right) \right] \quad (2.50)$$

$$x = \left( \frac{3}{4\pi\rho(\mathbf{r})} \right)^{1/6} \quad X(t) = t^2 + bt + c$$

$$Q = (4c - b^2)^{1/2} \quad P = \arctan \frac{Q}{2x + b}$$

con los parámetros:  $A=0,031\,090\,7$ ,  $x_0=-0,104\,98$ ,  $b=3,727\,44$ ,  $c=12,9352$ . La combinación de estos dos funcionales es lo que se denomina «aproximación de la densidad local» (*Local Density Approximation*, LDA), ya que los funcionales dependen

únicamente del valor de la densidad en cada punto, son funcionales locales. Esta aproximación resulta adecuada para sistemas en los que la densidad electrónica varía poco en el espacio, pero por lo general, para sistemas químicos, es insuficiente y no proporciona mejores resultados que el método Hartree-Fock.

### – Funcionales corregidos con el gradiente

Una posibilidad para mejorar los funcionales LDA es incluir los efectos del gradiente de la densidad electrónica, dando lugar a las «aproximaciones del gradiente» (*Generalised Gradient Approximations*, GGA). Existe una amplia variedad de funcionales, entre los que podemos destacar:

- Funcional de intercambio de Becke (88) [71]:

$$E_x^{\text{B88}}[\rho] = E_x^{\text{LDA}}[\rho] - \beta \int \rho^{4/3} \frac{x^2}{1 + 6\beta x \operatorname{arcsinh} x} \mathrm{d}\mathbf{r} \quad (2.51)$$

$$x = \frac{|\nabla\rho|}{\rho^{4/3}} \quad \beta = 0,0042$$

- Funcional de correlación de Lee-Yang-Parr [72]:

$$E_c^{\text{LYP}}[\rho] = -a \int \frac{\rho}{1 + d\rho^{-1/3}} \mathrm{d}\mathbf{r} - \quad (2.52)$$

$$-ab \int \omega \rho^2 \left[ \frac{3}{10} (3\pi^2)^{2/3} \rho^{8/3} - \left( \frac{1}{24} + \frac{7}{72} \delta \right) |\nabla\rho|^2 \right] \mathrm{d}\mathbf{r}$$

$$\omega = \frac{\exp(-c\rho^{-1/3})}{1 + d\rho^{-1/3}} \rho^{-11/3} \quad \delta = c\rho^{-1/3} + \frac{d\rho^{-1/3}}{1 + d\rho^{-1/3}}$$

$$a = 0,04918 \quad b = 0,132 \quad c = 0,2533 \quad d = 0,349$$

- Funcional híbrido de tres parámetros de Becke [73]:

$$E_{\text{xc}}^{\text{B3}} = E_{\text{xc}}^{\text{LDA}} + a_0 (E_x^{\text{HF}} - E_x^{\text{LDA}}) + a_x (E_x^{\text{B88}} - E_x^{\text{LDA}}) + a_c (E_c - E_c^{\text{VWN}}) =$$

$$= a_0 E_x^{\text{HF}} + a_x E_x^{\text{B88}} + (1 - a_0 - a_x) E_x^{\text{LDA}} + a_c E_c + (1 - a_c) E_c^{\text{VWN}} \quad (2.53)$$

El funcional  $E_x^{\text{HF}}$  se conoce como «intercambio exacto» y es el calculado de la misma manera que en el método Hartree-Fock, pero con los orbitales Kohn-Sham; la introducción de esta contribución es lo que caracteriza a los funcionales híbridos.  $E_c$  puede ser cualquier funcional de correlación, así, si se utiliza el funcional  $E_c^{\text{LYP}}$ , el funcional completo se denomina B3LYP.

En general, el coste computacional de los métodos DFT es del orden del método Hartree-Fock. La calidad de los resultados, sin embargo, es comparable a la del

método MP2, e incluso superior para propiedades como la geometría de estructuras estables y de transición. Esto hace de estos métodos una opción muy atractiva y eficiente, al menos para el estudio de estados electrónicos fundamentales.



---

---

# 3

---

## REPRESENTACIÓN CLÁSICA DEL DISOLVENTE

La principal característica de los métodos mixtos para el estudio de los efectos del disolvente es el empleo de métodos cuánticos para el tratamiento de la molécula del soluto y de técnicas de la mecánica clásica para el disolvente. Como ya se ha mencionado, esto permite obtener una descripción precisa y detallada de la estructura electrónica del soluto, que es el centro de atención, mientras el efecto del disolvente se introduce de forma más simplificada y sencilla. De esta manera se reduce el coste computacional respecto a los métodos totalmente cuánticos, a la vez que se mantiene una consideración adecuada de los factores más importantes.

La representación clásica del disolvente implica que no se consideran los efectos cuánticos de sus interacciones, o bien que estos efectos se introducen de una manera «clásica», mediante potenciales empíricos o propiedades promedio. En este capítulo se describen algunas de las principales maneras de introducir esta representación clásica en los cálculos teóricos, siempre dentro del contexto de los métodos mixtos.

### 3.1. Métodos de disolvente continuo

Este tipo de métodos considera al disolvente como un medio continuo sin estructura interna, cuya respuesta frente al soluto —su polarización— puede representarse por medio de una constante macroscópica, tal como la permisividad eléctrica o constante dieléctrica,  $\epsilon$  [74]. Esta simplificación facilita mucho los cálculos, pero generalmente resulta excesiva, sobre todo cuando las interacciones específicas entre las moléculas de disolvente de la primera capa de solvatación y el soluto

juegan un papel importante en el sistema estudiado, como ocurre cuando existen enlaces de hidrógeno.

### 3.1.1. Los primeros modelos

El primer modelo de continuo polarizable es debido a Born [75], quien en 1920 propuso un método para el cálculo de las energías libres de solvatación de iones. En este modelo, un soluto esférico de radio  $R$  y con una carga  $q$  uniformemente distribuida interacciona con un disolvente continuo de constante dieléctrica  $\epsilon$ . La contribución electrostática a la energía libre de solvatación viene dada por la diferencia de trabajo necesario para cargar dicho ion en el vacío,  $W^0$ , y en el disolvente,  $W$ :

$$\begin{aligned} W^0 &= \frac{q^2}{2R} & W &= \frac{q^2}{2\epsilon R} \\ \Delta G^{\text{elec}} &= W - W^0 = \frac{1 - \epsilon}{2\epsilon} \frac{q^2}{R} \end{aligned} \quad (3.1)$$

A pesar de su sencillez, el modelo de Born sigue siendo de utilidad para estimar las energías de solvatación de iones. Rashin y Honig [76] han reproducido de manera precisa las energías libres de hidratación de aniones y cationes empleando el modelo de Born con radios  $R$  adecuadamente escogidos.

Posteriormente, Bell y Onsager [77, 78] ampliaron el modelo de Born para hacerlo aplicable al caso de moléculas neutras. En el modelo de Onsager se considera al soluto como un dipolo puntual con polarizabilidad  $\alpha$  [79] y con momento dipolar  $\mu_0$ , el centro de una cavidad esférica de radio  $R$  y constante dieléctrica 1. En el exterior de esta cavidad se encuentra el disolvente continuo de constante dieléctrica  $\epsilon$ . Al introducirse la molécula en el disolvente dos efectos tienen lugar: por una parte, el disolvente genera un campo eléctrico,  $E_R$ , como reacción al dipolo que está en la cavidad; por otra parte, el soluto se polariza debido a este «campo de reacción» y su momento dipolar pasa a ser igual a  $\mu$ . En última instancia estos dos efectos se hallan en equilibrio mutuo, de manera que:

$$\begin{aligned} E_R &= f \mu & f &= \frac{2(\epsilon - 1)}{(2\epsilon + 1)R^3} \\ \mu &= \mu_0 + \alpha E_R = \frac{\mu_0}{1 - \alpha f} \end{aligned} \quad (3.2)$$

Para el cálculo de la componente electrostática de la energía libre de solvatación hay que tener en cuenta la energía de interacción entre el dipolo y el campo de reacción (incluyendo la creación de éste),  $W_R$ , pero también la energía de distorsión

necesaria para polarizar el soluto,  $W_{\text{dist}}$ :

$$\begin{aligned} W_{\text{R}} &= -\frac{1}{2}\mu E_{\text{R}} & W_{\text{dist}} &= \frac{1}{2}(\mu - \mu_0)E_{\text{R}} \\ \Delta G^{\text{elec}} &= W_{\text{R}} + W_{\text{dist}} = -\frac{1}{2}\mu_0 E_{\text{R}} = \frac{f\mu_0^2}{2(\alpha f - 1)} \end{aligned} \quad (3.3)$$

Kirkwood [80] generalizó este modelo introduciendo multipolos de orden superior y desarrollando una teoría general de interacciones para distribuciones arbitrarias de carga en cavidades esféricas. Suponiendo un conjunto de  $N$  cargas puntuales en el interior de una cavidad esférica centrada en el origen de coordenadas, Kirkwood llega a la expresión:

$$\begin{aligned} \Delta G^{\text{elec}} &= \frac{1}{2} \sum_{i=1}^N q_i V_{\text{R}}(\mathbf{r}_i) = -\frac{1}{2} \sum_{l=0}^{\infty} f_l \sum_{i=1}^N \sum_{j=1}^N q_i q_j r_i^l r_j^l P_l(\cos \theta_{ij}) \\ f_l &= \frac{(l+1)(\epsilon-1)}{(l+1)\epsilon+l} \frac{1}{R^{2l+1}} \end{aligned} \quad (3.4)$$

donde  $q_i$  es el valor de la carga situada en el punto  $\mathbf{r}_i$ ,  $V_{\text{R}}(\mathbf{r}_i)$  es el potencial de reacción en dicho punto,  $\theta_{ij}$  es el ángulo formado por los vectores  $\mathbf{r}_i$  y  $\mathbf{r}_j$  y  $P_l(x)$  son los polinomios de Legendre. Cada término en  $l$  representa la contribución del momento multipolar de orden  $l$ . A partir de la ecuación (3.4) pueden obtenerse las ecuaciones de Born y Onsager (para un dipolo no polarizable) si se considera una única carga en el centro de la cavidad o dos cargas idénticas de signo opuesto en posiciones simétricas respecto al centro, respectivamente.

En 1938, Kirkwood y Westheimer [81, 82] aplicaron este método al cálculo de constantes de disociación de ácidos orgánicos empleando tanto cavidades esféricas como elipsoidales.

### 3.1.2. Métodos de campo de reacción autocoherente

La expresión multipolar de Kirkwood ha servido como base para el desarrollo de una serie de modelos mejor adaptados a su uso conjunto con la descripción cuántica del soluto. En estos métodos se supone un soluto complejo, tratado mediante alguno de los métodos cuánticos disponibles, situado en una cavidad creada en un medio dieléctrico y continuo. Las primeras versiones se limitaban a cavidades esféricas, pero muchos de los métodos han sido ampliados para considerar cavidades elipsoidales o con formas irregulares.

En general, estos métodos se conocen como métodos de «campo de reacción autocoherente» (*Self-Consistent Reaction Field*, SCRFF) y se basan en introducir el efecto del campo de reacción en el hamiltoniano del soluto. Se tienen distintos métodos

según la manera de definir la cavidad en la que se introduce el soluto, la representación de la distribución de carga del soluto, la forma de calcular la contribución de la dispersión a la energía de interacción, etc. Los primeros métodos, desarrollados por Rinaldi y Rivail [34, 83] y Tapia y col. [35] empleaban una cavidad esférica y aproximaban el soluto mediante una serie de multipolos centrados en el origen. Posteriormente se ampliaron para el uso de cavidades elipsoidales [84] y adaptadas a la forma molecular del soluto [85]. También se han desarrollado otros métodos en los que el campo de reacción se calcula representando al soluto mediante un conjunto de cargas o multipolos asociados a los distintos grupos químicos de la molécula [86].

En los métodos SCRF entran en juego, además de la energía de interacción electrostática y la energía de polarización del soluto, otros términos como la interacción de Van der Waals o la energía necesaria para crear la cavidad en el medio dieléctrico. Normalmente, estos términos se suponen proporcionales al área de la cavidad, ya sea con una única constante de proporcionalidad o con parámetros para cada tipo de átomo ajustados a valores experimentales.

### 3.1.3. El modelo del continuo polarizable

Quizá el más conocido de los métodos de disolvente continuo sea el llamado «modelo del continuo polarizable» (*Polarisable Continuum Model*, PCM) [36, 87-89]. En este modelo la cavidad en la que se introduce el soluto se construye mediante una serie de esferas solapadas, centradas en los núcleos del soluto y con radios proporcionales a los correspondientes radios de Van der Waals. La presencia del soluto en la cavidad induce una polarización en el medio circundante, que se traduce en una densidad de carga aparente,  $\sigma$ , en la superficie de la cavidad. A efectos prácticos, esta superficie se divide en elementos de área  $S_i$  y la densidad de carga superficial se sustituye por unas cargas puntuales,  $q_i$ , de manera que:

$$q_i = \sigma(\mathbf{r}_i)S_i \quad (3.5)$$

siendo  $\mathbf{r}_i$  el punto central del elemento de superficie  $i$ , donde se sitúa la carga  $q_i$ .

Con esta forma de representar la reacción del continuo dieléctrico frente al soluto se evita el empleo de expansiones multipolares, ya que los valores de las cargas pueden calcularse fácilmente de la siguiente manera:

$$q_i = \frac{1 - \epsilon}{4\pi} \mathbf{E}(\mathbf{r}_i) \cdot \mathbf{S}_i \quad (3.6)$$

donde  $\mathbf{S}_i$  es el vector que representa al elemento de superficie  $i$  (su módulo,  $S_i$ , es perpendicular a la superficie de la cavidad en  $\mathbf{r}_i$  y está dirigido hacia el dieléctrico) y  $\mathbf{E}(\mathbf{r}_i)$  es el campo eléctrico existente en la posición  $\mathbf{r}_i$ . Este campo eléctrico puede

dividirse en tres contribuciones:

$$E(\mathbf{r}_i) = E_{\text{sol}}(\mathbf{r}_i) + \sum_{j \neq i} \frac{q_j}{r_{ji}^3} \mathbf{r}_{ji} + 2\pi q_i (1 + \eta_i) \frac{S_i}{S_i^2} \quad (3.7)$$

El primer término corresponde al campo eléctrico generado por el soluto, que puede calcularse a partir de su densidad electrónica, obtenida por métodos cuánticos. El segundo término es el campo eléctrico generado por el resto de las cargas ficticias  $q_j$ . El último término representa el efecto de la densidad de carga del elemento  $i$  sobre el mismo elemento de superficie  $i$ ,  $\eta_i$  es un factor de corrección para tener en cuenta el hecho de que la superficie de la cavidad no es plana.

A partir de una determinada densidad electrónica del soluto puede obtenerse un conjunto de cargas  $q_i$ , ya sea por métodos iterativos o mediante la resolución de una ecuación matricial. El efecto de estas cargas se introduce en el hamiltoniano molecular del soluto para obtener una nueva densidad electrónica y el proceso se repite hasta alcanzar la convergencia.

Con un fundamento similar al del modelo PCM existen algunas variaciones según la forma de definir las cargas en la superficie de la cavidad. Por ejemplo, el «modelo de apantallamiento tipo conductor» (*Conductor-like Screening Model*, COSMO) [90, 91] impone la restricción de que el potencial electrostático ha de ser nulo en la superficie y resulta menos sensible a la forma y tamaño escogidos para la cavidad. El «formalismo de ecuaciones integrales» (*Integral Equation Formalism*, IEF) [92] permite el tratamiento unificado de líquidos anisótropos, disoluciones iónicas, etc.

## 3.2. Modelos de disolvente estructurado

En una situación intermedia entre los métodos del continuo y la mecánica molecular, que se verá más tarde, se encuentran estos modelos, que introducen una cierta estructura —aunque bastante simplificada— dentro del disolvente. Dentro de este conjunto podemos incluir el «modelo de referencia de lugares de interacción» (*Reference Interaction Site Model*, RISM) [8], que describe, mediante ecuaciones integrales, la estructura de fluidos compuestos por moléculas de forma irregular. También pertenece a este grupo el modelo de «dipolos de Langevin».

### 3.2.1. El modelo de referencia de lugares de interacción

Existe una serie de modelos que intentan obtener, a partir de las propiedades moleculares de los líquidos —forma, tamaño e interacciones entre sus moléculas—, otras propiedades estructurales y termodinámicas como pueden ser las funciones de correlación (ver pág. 61), la energía libre, etc. Uno de los modelos de este tipo

más utilizados es el llamado RISM, en el que se supone que las moléculas del líquido están formadas por «lugares de interacción» —identificados normalmente con sus átomos— y que mantienen una estructura rígida. El objetivo principal es determinar la función de correlación,  $g(r)$  entre lugares de interacción pertenecientes a moléculas diferentes. Puesto que cada lugar de interacción viene acompañado por el resto de la molécula, la función  $g(r)$  dependerá no sólo de los dos lugares implicados, sino de la molécula completa.

Las ecuaciones a resolver en este tipo de métodos no son sencillas [6] e implican integrales de convolución (integrales de la forma  $f * g = \int f(\tau)g(t - \tau) d\tau$ ), transformadas de Fourier ( $F(k) = \int f(x)e^{-2\pi i k x} dx$ ), diferentes funciones de correlación, etc. La ecuación principal de estos métodos es la relación de Ornstein-Zernike [93] que, en un sistema homogéneo e isótropo toma la forma:

$$h(r) = c(r) + \rho \int c(|r - r'|)h(r') dr' \quad (3.8)$$

donde  $h(r) = g(r) - 1$ ,  $c(r)$  es la función de correlación directa y  $\rho$  es la densidad numérica.

El método RISM ha sido adaptado para su utilización en un método mixto, denominado RISM-SCF [38, 39, 94]. En este método, la estructura del disolvente obtenida mediante RISM se introduce en el hamiltoniano molecular del soluto, de manera que los cálculos cuánticos se realizan en presencia del disolvente. Además, la estructura electrónica obtenida para el soluto se emplea para efectuar un nuevo cálculo RISM para el sistema soluto-disolvente, lo que permite considerar su influencia mutua. Este procedimiento es similar al empleado en el método ASEP/MD (ver capítulo 4, pág. 73), la principal diferencia es la utilización de ecuaciones integrales (en el caso de RISM-SCF) o de técnicas de simulación (en el caso de ASEP/MD) para la descripción de la estructura del disolvente.

### 3.2.2. El modelo de dipolos de Langevin

El modelo de dipolos de Langevin, desarrollado principalmente por Warshel y col. [95-97] representa al disolvente no como un medio continuo, sino como una red de dipolos puntuales cuya polarización debe remedar la polarización promedio de las moléculas de disolvente en la correspondiente zona del espacio. Esto se consigue mediante una relación de tipo Langevin:

$$\mu_i = \mu_0 \left( \coth \frac{C\mu_0 E_i}{k_B T} - \frac{k_B T}{C\mu_0 E_i} \right) \frac{E_i}{E_i} + \alpha E_i \quad (3.9)$$

$\mu_0$  y  $C$  son parámetros cuyo valor varía de un disolvente a otro;  $\mu_0$  representa el momento dipolar permanente de las moléculas de disolvente y  $C$ , su resistencia

frente a la orientación —debida, por ejemplo, a la existencia de enlaces de hidrógeno intermoleculares.  $\alpha$  es la polarizabilidad isotropa de las moléculas de disolvente y  $E_i$  es el campo eléctrico en el lugar ocupado por el dipolo  $i$  generado por el resto del sistema.

Puesto que el valor de  $\mu_i$  depende de  $E_i$  y  $E_i$  depende a su vez de  $\mu_i$ , el método debe aplicarse de manera iterativa hasta alcanzar la convergencia. Una vez obtenidos los dipolos del disolvente, su efecto se introduce en el hamiltoniano del soluto de la manera usual en los métodos de hamiltoniano efectivo.

Este método sólo ha sido bien parametrizado para el agua como disolvente, en tal caso los dipolos se disponen en una red cúbica con un tamaño de celda de 0,3 nm, valor condicionado por la densidad del agua. En una primera versión del método, en la que no se incluían los dipolos inducidos,  $\alpha E_i$ , se tomó  $\mu_0$  como 1,7 D (el valor experimental es 1,8 D). Cuando se añadió el término de los dipolos inducidos se redujo el valor de  $\mu_0$  a 1,5 D, contrarrestando la sobrestimación de la polarización debida al empleo de una polarizabilidad isotropa.

La principal aplicación del método de dipolos de Langevin ha sido en combinación con el método de «enlace de valencia empírico» (*Empirical Valence Bond*, EVB), desarrollado también por Warshel [98, 99], para el tratamiento cuántico del soluto. Este método se basa en la «teoría de enlace de valencia» [100], según la cual la estructura y las propiedades de una molécula pueden representarse por una combinación de estructuras ideales en las que los distintos fragmentos están unidos por enlaces covalentes apolares o enlaces iónicos. Por ejemplo, una molécula diatómica XY podría representarse como una combinación de la estructura completamente covalente, X–Y, y de las iónicas, X<sup>+</sup>Y<sup>-</sup> y X<sup>-</sup>Y<sup>+</sup>. En el método EVB, se introducen parámetros empíricos (experimentales) como las energías de disociación, las frecuencias vibracionales o los potenciales de ionización de las moléculas o fragmentos implicados. Cuando se estudian las moléculas en un disolvente, se asume que sólo las estructuras ideales con carácter iónico se ven afectadas; este efecto se calcula empleando el modelo de dipolos de Langevin. También puede emplearse un modelo similar para representar el entorno cuando éste es una proteína en lugar de un disolvente, lo que ha permitido el estudio de algunas reacciones enzimáticas [101].

### 3.3. Métodos de mecánica molecular

En los métodos de «mecánica molecular» (*Molecular Mechanics*, MM), las moléculas se representan explícitamente y sus interacciones se modelan por medios clásicos, es decir, se basan en la aplicación de potenciales analíticos sencillos aplicados a las leyes de la mecánica clásica. Básicamente, los átomos se tratan como masas

puntuales y la energía potencial del sistema se divide en varias contribuciones:

$$E = E_r + E_\theta + E_\omega + E_q + E_{\text{vdw}} + E_{\text{acop}} \quad (3.10)$$

Algunas fórmulas sencillas para los distintos términos son las siguientes:

- $E_r$  es la energía de tensión de los enlaces. Se considera cada enlace como un resorte con una distancia de equilibrio  $r^0$  y una constante de elasticidad  $k^r$ , cuyos valores dependen del tipo de enlace de que se trate. La energía de cada enlace es proporcional a la desviación de su longitud respecto a la distancia de equilibrio:

$$E_r = \sum_i^{\text{enl.}} k_i^r (r_i - r_i^0)^2 \quad (3.11)$$

- $E_\theta$  es la energía de flexión. Dos enlaces con un átomo en común forman un ángulo, que tiene un valor de equilibrio  $\theta^0$  y una constante de rigidez  $k^\theta$ . De manera similar a la tensión de los enlaces, la energía de flexión viene dada por:

$$E_\theta = \sum_i^{\text{áng.}} k_i^\theta (\theta_i - \theta_i^0)^2 \quad (3.12)$$

- $E_\phi$  es la energía de torsión. Cada ángulo define un plano, los planos definidos por dos ángulos que comparten un enlace forman un ángulo de torsión  $\phi$ . La energía correspondiente a estas torsiones tiene la forma ( $n$  determina la simetría de cada torsión):

$$E_\phi = \sum_i^{\text{tor.}} k_i^\phi [1 + \cos(n_i \phi_i)] \quad (3.13)$$

- $E_q$  es la energía de interacción electrostática. A cada átomo se le asigna una carga  $q$  y la interacción total se obtiene como:

$$E_q = \sum_i^{\text{átom.}} \sum_j \frac{q_i q_j}{r_{ij}} \quad (3.14)$$

- $E_{\text{vdw}}$  es la energía de Van der Waals, que incluye la repulsión electrónica y las interacciones de dispersión entre los distintos átomos. Generalmente se modela mediante un potencial de Lennard-Jones, para cada pareja de átomos existen dos parámetros,  $A$  y  $B$ , la energía total es:

$$E_{\text{vdw}} = \sum_i^{\text{átom.}} \sum_j \left( \frac{B_{ij}}{r_{ij}^{12}} - \frac{A_{ij}}{r_{ij}^6} \right) \quad (3.15)$$



- $E_{\text{acop}}$  es la energía de acoplamiento entre los términos anteriores, que viene dada por términos cruzados que implican distintas distancias o ángulos. La introducción de estos términos cruzados —y de la consiguiente energía de acoplamiento— es muchas veces necesaria para obtener buenos resultados en las simulaciones. Algunos de los tipos de términos cruzados más empleados son los siguientes:

- Términos tensión-tensión:

$$E_{\text{rr}}^{ij} = \frac{1}{2} k_{ij}^{\text{rr}} \left[ (r_i - r_i^0) (r_j - r_j^0) \right] \quad (3.16)$$

- Términos tensión-flexión:

$$E_{\text{r}\theta}^{ij} = \frac{1}{2} k_{ij}^{\text{r}\theta} \left[ (r_i - r_i^0) + (r_j - r_j^0) \right] (\theta_{ij} - \theta_{ij}^0) \quad (3.17)$$

- Términos tensión-torsión:

$$E_{\text{r}\phi}^{ij} = k_{ij}^{\text{r}\phi} (r_i - r_i^0) [1 + \cos (n_j \phi_j)] \quad (3.18)$$

Naturalmente, estas fórmulas pueden complicarse introduciendo términos anarmónicos, otros términos de acoplamiento, términos adicionales que representen, por ejemplo, los enlaces de hidrógeno, etc. La forma concreta de los potenciales empleados, junto con el conjunto de valores asignados a los distintos parámetros, recibe el nombre de «campo de fuerzas». Se han desarrollado diferentes campos de fuerzas ajustados para el estudio de una gran variedad de propiedades, procesos y sistemas, algunos de los nombres de estos campos de fuerzas son: AMBER [102], CHARMM [103], OPLS [104-106], MM2 [107] o MMFF94 [108].

### 3.3.1. Simulaciones

El empleo de los métodos de mecánica molecular en el estudio de sistemas en fase líquida se basa en la generación de un conjunto de configuraciones —disposiciones espaciales de todas las moléculas del sistema— y el análisis estadístico de sus propiedades.

En la mecánica estadística, el valor promedio o valor esperado de una propiedad  $\mathcal{X}$  puede expresarse como:

$$\langle \mathcal{X} \rangle = \iint \mathcal{X}(\mathbf{p}^N, \mathbf{r}^N) \rho(\mathbf{p}^N, \mathbf{r}^N) d\mathbf{p}^N d\mathbf{r}^N \quad (3.19)$$

donde  $\rho(\mathbf{p}^N, \mathbf{r}^N)$  es la densidad de probabilidad del colectivo, es decir, la probabilidad de encontrar una configuración en la que las partículas tienen momentos  $\mathbf{p}^N$  y

posiciones  $\mathbf{r}^N$ . Bajo condiciones de número de partículas, volumen y temperatura constantes, la densidad de probabilidad es la conocida distribución de Boltzmann:

$$\rho(\mathbf{p}^N, \mathbf{r}^N) = \frac{e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N)}}{\iint e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N)} d\mathbf{p}^N d\mathbf{r}^N} \quad (3.20)$$

donde  $\beta = 1/k_B T$  y  $E(\mathbf{p}^N, \mathbf{r}^N)$  es la energía de cada configuración.

Según la «hipótesis ergódica», uno de los axiomas fundamentales en la mecánica estadística, el promedio colectivo de una propiedad (el promedio de los valores para todas las configuraciones posibles) es igual a su promedio temporal (el promedio para las configuraciones que, a lo largo del tiempo, toma un sistema). Existen, por lo tanto, dos métodos fundamentales para generar las distintas configuraciones del sistema: el método de «Monte Carlo» (MC) y la «dinámica molecular» (DM).

En la práctica, la simulación completa de sistemas reales es inviable, ya que el número de partículas presentes es enorme, del orden de  $10^{23}$ . Lo normal es que se represente únicamente un subsistema compuesto por un número más manejable de partículas, generalmente entre 100 y 1000. Para evitar artefactos debidos a la elevada superficie relativa de un sistema tan reducido, se supone que éste corresponde a una celdilla unitaria —normalmente cúbica— que, por repetición en todas las direcciones, ocupa todo el espacio. Con esta técnica, conocida como «condición de límite periódico», sólo es necesario simular una celdilla, el resto del espacio está ocupado por imágenes idénticas, que no es necesario considerar explícitamente.

Para el cálculo de las interacciones en el sistema, se tienen en cuenta las interacciones de cada partícula con todas las otras partículas y con sus imágenes periódicas situadas en celdas vecinas. Es frecuente, sin embargo, considerar las interacciones únicamente cuando la distancia entre las partículas implicadas es menor que una distancia dada. Este truncamiento de las interacciones de largo alcance —especialmente las electrostáticas— puede introducir errores importantes.

Existen métodos para tratar estas interacciones de largo alcance, entre los que se puede destacar la suma de Ewald [15, 109, 110], expresando el potencial electrostático mediante una suma de términos en el espacio real y en el espacio recíproco (relacionado con el espacio real a través de una transformada de Fourier, aprovechando la periodicidad del sistema):

$$E_q = \sum_{\mathbf{n}=0}^{\infty} \sum_{i=1}^N \sum_{j>i}^N q_i q_j \frac{\operatorname{erfc}(\alpha |\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} + \frac{4\pi}{L^3} \sum_{\mathbf{k} \neq 0}^{\infty} \frac{e^{-k^2/4\alpha^2}}{k^2} \left( \sum_{i=1}^N q_i e^{-i\mathbf{k} \cdot \mathbf{r}_i} \right)^2 - \frac{\alpha}{\sqrt{\pi}} \sum_{i=1}^N q_i^2 \quad (3.21)$$

donde:

- erfc( $x$ ) función de error complementaria,  $\text{erfc}(x) = 2/\sqrt{\pi} \int_x^\infty e^{-t^2} dt$ .
- $\mathbf{n}$  vector de red en el espacio real,  $\mathbf{n} = (n_x L, n_y L, n_z L)$ .
- $\mathbf{k}$  vector de red en el espacio recíproco,  $\mathbf{k} = 2\pi \mathbf{n}/L^2$ .
- $\alpha$  parámetro de relación entre los espacios real y recíproco.
- $N$  número de cargas en el sistema.
- $L$  lado de la celdilla unidad, supuesta cúbica.

Por lo general, se intenta truncar el primer sumando en  $\mathbf{n} = 0$ , escogiendo un valor adecuado para  $\alpha$  y un número suficiente de términos (vectores  $\mathbf{k}$ ) en el segundo sumando.

Una de las propiedades más interesantes que pueden extraerse de una simulación de mecánica molecular es la «función de distribución radial» (*radial distribution function*, RDF) o función de correlación de pares,  $g(r)$ . Esta función representa la probabilidad de encontrar un par de partículas de determinado tipo con una separación  $r$ , comparada con la probabilidad en una distribución de gas ideal, su expresión general es:

$$g(r) = \frac{L^3}{N^2} \left\langle \sum_{j \neq i} \delta(r - r_{ij}) \right\rangle \quad (3.22)$$

donde  $\delta(x)$  es la función delta de Dirac. La RDF ofrece una información muy valiosa sobre la estructura del sistema y, cuando todas las interacciones se consideran aditivas por pares, es suficiente para calcular sus propiedades termodinámicas. Además, empleando técnicas de difracción de rayos X y neutrones, puede determinarse la RDF experimentalmente.

Relacionados con la RDF están los números de coordinación, que permiten estimar el número de partículas que se encuentran a una determinada distancia de una partícula dada. Los números de coordinación pueden obtenerse por integración de la densidad local.

### – El método de Monte Carlo

La técnica conocida como método de Monte Carlo, o muestreo estocástico, consiste en la simulación de un experimento en una muestra aleatoria, con la finalidad de determinar las propiedades probabilísticas de una población. El ingrediente básico del método es una serie de números aleatorios, a partir de los cuales —y mediante reglas que varían según el problema— se obtienen unos resultados de cuyo análisis pueden extraerse soluciones a problemas complejos. Se considera que el origen del método (o, al menos, de su nombre) es el artículo publicado en 1949 por Metropolis y Ulam [111].

Una aplicación inmediata de este método al estudio de propiedades termodinámicas podría basarse en una versión discreta de la ecuación (3.19):

$$\langle \mathcal{X} \rangle \simeq \sum_{i=1}^M \mathcal{X}(\mathbf{r}_i^N) w(\mathbf{r}_i^N) \quad w(\mathbf{r}^N) = \frac{e^{-\beta E(\mathbf{r}^N)}}{\sum_{i=1}^M e^{-\beta E(\mathbf{r}_i^N)}} \quad (3.23)$$

donde  $E(\mathbf{r}^N)$  es la energía potencial del sistema y se ha considerado que  $\mathcal{X}$  no depende de las velocidades de las partículas. Se generan  $M$  configuraciones aleatoriamente, se calcula el valor de la propiedad  $\mathcal{X}$  para cada configuración, así como el de  $w$  (el factor de probabilidad), y se obtiene  $\langle \mathcal{X} \rangle$  como resultado. El problema de este procedimiento es que el número de configuraciones necesarias para que la muestra sea representativa del sistema es demasiado grande. Las configuraciones que más influencia tienen en el promedio son las de menor energía, ya que las de energía más elevada tienen un factor  $w$  muy pequeño; sin embargo, un muestreo aleatorio y uniforme del espacio de configuraciones generará una proporción muy grande de configuraciones de alta energía, pues son las más probables (en ausencia de restricciones).

Una posible solución es el llamado algoritmo de Metropolis [16], mediante el cual se consigue una muestra de configuraciones en la que la probabilidad de cada configuración es igual al factor  $w$ , es decir, se muestrean preferentemente las configuraciones que mayor influencia tienen en el promedio termodinámico. De esta manera, el valor esperado de  $\mathcal{X}$  puede calcularse con un promedio sencillo:

$$\langle \mathcal{X} \rangle \simeq \frac{1}{M} \sum_{i=1}^M \mathcal{X}(\mathbf{r}_i^N) \quad (3.24)$$

El algoritmo es el siguiente:

1. Se toma una configuración aleatoria de las partículas del sistema,  $\mathbf{r}_i^N$ .
2. Se toma una configuración de prueba,  $\mathbf{r}_p^N$ , obtenida aleatoriamente a partir de  $\mathbf{r}_i^N$  y, por lo general, similar a ésta.
3. Se calcula  $P_{ip}$  como el cociente  $w_p/w_i$ :

$$P_{ip} = \frac{w_p}{w_i} = e^{-\beta[E(\mathbf{r}_p^N) - E(\mathbf{r}_i^N)]} \quad (3.25)$$

4. Se toma un número aleatorio entre 0 y 1. Si este número es menor que  $P_{ip}$ , se hace  $\mathbf{r}_{i+1}^N = \mathbf{r}_p^N$ ; en caso contrario,  $\mathbf{r}_{i+1}^N = \mathbf{r}_i^N$ .
5. Se vuelve al paso 2, sustituyendo  $i$  por  $i + 1$ .

Este proceso se repite hasta que se han obtenido  $M$  configuraciones. El punto clave del algoritmo es calcular la relación entre las probabilidades de dos configuraciones en lugar de la probabilidad de cada configuración.

El método empleado para obtener la configuración de prueba es también aleatorio. Generalmente se escoge una partícula del sistema (molécula, grupo funcional, átomo...) y se altera ligeramente su posición. Si la nueva configuración es más estable (menos energética) que la anterior, se acepta; en caso contrario la probabilidad de aceptar esta nueva configuración es mayor cuanto menor sea la diferencia de energía.

El número de configuraciones necesarias en la muestra es aún elevado, pero mucho menor que en el caso de utilizar la fórmula (3.23). Una vez que se cuenta con un conjunto de configuraciones representativas del sistema, se pueden extraer las diferentes propiedades termodinámicas de su análisis. La principal limitación del método MC es que sólo ofrece una imagen estática del sistema en estado de equilibrio, no proporciona información sobre su evolución temporal, aunque existen modificaciones para tratar este aspecto [112].

### – Dinámica molecular

El método de dinámica molecular es una simulación determinista (no aleatoria) de los movimientos de todas las partículas de un sistema. Se basa en la idea, expresada por Laplace en 1820, de que, si se conociese con absoluta precisión la posición y la velocidad de todos los componentes de un sistema así como las fuerzas que actúan sobre ellos, podría deducirse su situación en cualquier momento pasado o futuro. La dinámica molecular no aspira a tales pretensiones, pero sí asume que a partir de un determinado conjunto de posiciones y velocidades se puede determinar la evolución temporal del sistema con una precisión suficiente para el estudio de propiedades termodinámicas y estructurales.

Este método fue introducido por Alder y Wainwright [17, 113] para el estudio de interacciones de esferas rígidas; sus trabajos fueron muy importantes en la comprensión de los líquidos sencillos. Rahman y Stillinger [18, 114] llevaron a cabo las primeras simulaciones de sistemas realistas. Hoy en día las técnicas de dinámica molecular son ampliamente usadas en una gran variedad de campos científicos.

El proceso comienza definiendo las posiciones,  $\mathbf{r}^N$ , y las velocidades o momentos lineales,  $\mathbf{p}^N$ , de todas las partículas del sistema. Para esta configuración inicial puede calcularse la energía cinética y la energía potencial, según una fórmula del tipo (3.10). También puede calcularse la fuerza sobre cada una de las partículas:

$$\mathbf{F}_i(\mathbf{r}^N) = -\nabla_i E(\mathbf{r}^N) \quad (3.26)$$

Dada la forma de  $E$  dentro del marco de la mecánica molecular, el cálculo de  $F_i$  suele ser sencillo y se reduce a la suma de diferentes contribuciones. Además, según la Segunda Ley de Newton, la fuerza que actúa sobre cada partícula induce una aceleración directamente proporcional, por lo tanto:

$$-\frac{\partial E}{\partial \mathbf{r}_i} = m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = m_i \ddot{\mathbf{r}}_i = \frac{\partial \mathbf{p}_i}{\partial t} = \dot{\mathbf{p}}_i \quad (3.27)$$

con lo que se tiene una relación entre la configuración actual y su evolución en el tiempo.

La integración analítica de la ecuación (3.27) es generalmente imposible y se recurre a una resolución numérica basada en desarrollos en serie de Taylor. Para ello se toma un intervalo temporal pequeño,  $\Delta t$ , que suele ser del orden de fs, y se calculan las posiciones  $\mathbf{r}$  y velocidades  $\mathbf{v}$  de las partículas del sistema en diferentes instantes separados por dicho intervalo.

Un método sencillo, propuesto por Verlet [115] es el siguiente:

$$\begin{aligned} \mathbf{r}(t + \Delta t) &= 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \ddot{\mathbf{r}}(t)\Delta t^2 \\ \mathbf{v}(t) &= \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t)}{2\Delta t} \end{aligned} \quad (3.28)$$

en el que no se hace uso de las velocidades para calcular las posiciones de las partículas en cada instante, sino que es a la inversa. El algoritmo es compacto y rápido, pero puede conducir a errores numéricos considerables.

Se han propuesto algunas modificaciones del algoritmo de Verlet que resultan más robustas, una de ellas es el algoritmo de la «pídola» (*leapfrog*) [116]:

$$\begin{aligned} \mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \mathbf{v}(t + \frac{1}{2}\Delta t)\Delta t \\ \mathbf{v}(t + \frac{1}{2}\Delta t) &= \mathbf{v}(t - \frac{1}{2}\Delta t) + \ddot{\mathbf{r}}(t)\Delta t \end{aligned} \quad (3.29)$$

cuya principal característica es que calcula las posiciones en los instantes  $t$ ,  $t + \Delta t$ ,  $t + 2\Delta t$ , ... y las velocidades en los instantes  $t + \frac{1}{2}\Delta t$ ,  $t + \frac{3}{2}\Delta t$ ,  $t + \frac{5}{2}\Delta t$ , ...

Otro método que minimiza los errores numéricos en cálculos realizados con precisión finita es el propuesto por Swope y col. [117]:

$$\begin{aligned} \mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\ddot{\mathbf{r}}(t)\Delta t^2 \\ \mathbf{v}(t + \frac{1}{2}\Delta t) &= \mathbf{v}(t) + \frac{1}{2}\ddot{\mathbf{r}}(t)\Delta t \\ \mathbf{v}(t + \Delta t) &= \mathbf{v}(t + \frac{1}{2}\Delta t) + \frac{1}{2}\ddot{\mathbf{r}}(t + \Delta t)\Delta t \end{aligned} \quad (3.30)$$

que se conoce como algoritmo de Verlet de la velocidad (*velocity Verlet*).

Siguiendo este procedimiento se obtienen configuraciones pertenecientes a un colectivo microcanónico, es decir, se mantienen constantes el número de partículas,

el volumen del sistema y su energía total (suma de las energías cinética y potencial),  $NVE$ . En el inicio de una simulación el sistema raramente se encuentra en un estado de equilibrio, sino que tiene una energía potencial relativamente elevada. Con el transcurso de la simulación disminuye la energía potencial del sistema y aumenta la energía cinética y, por lo tanto, la temperatura. Cuando el objetivo es obtener configuraciones de un colectivo canónico,  $NVT$ , en el que la temperatura es constante, se hace necesario el empleo de algún método para controlarla.

El método más sencillo consiste simplemente en renormalizar las velocidades para que la temperatura del sistema sea la requerida. Cada cierto número de pasos de simulación se sustituyen las velocidades  $v$  por otras  $v'$  tales que:

$$v'_i = \sqrt{\frac{3Nk_B T}{2E_k}} v_i \quad E_k = \frac{1}{2} \sum_i^N m_i v_i^2 \quad (3.31)$$

sin embargo, esto introduce perturbaciones artificiales en la simulación y en las propiedades termodinámicas derivadas de ella. Si esta renormalización se realizara en cada paso de la dinámica, se anularía la fluctuación en la energía cinética, lo cual es irreal.

Un método más sofisticado es el llamado «termostato de Nosé-Hoover» [118, 119], en el que el sistema se acopla con un baño térmico, que se introduce a través de una masa ficticia,  $Q$ . En lugar de  $F_i = m_i \ddot{r}_i$  se obtiene:

$$F_i = m_i \ddot{r}_i + \zeta v_i \quad \zeta = \frac{3N}{Q} \left( \sum_i^N m_i v_i^2 - k_B T \right) \quad (3.32)$$

Este método también introduce fluctuaciones sin sentido físico en la simulación, pero éstas pueden minimizarse con una correcta elección de  $Q$  y un tiempo de simulación suficientemente extenso.

En el caso de que se consideren moléculas rígidas, en las que las posiciones relativas de los distintos núcleos son fijas, la expresión (3.10) se simplifica y sólo es necesario tener en cuenta los términos  $E_q$  y  $E_{vdw}$  entre átomos de diferentes moléculas, ya que las interacciones intramoleculares se mantienen constantes y no juegan, por lo tanto, ningún papel en la dinámica del sistema. Además, en este caso no es necesario tratar de manera independiente a cada átomo, sino que las ecuaciones de movimiento pueden describir los movimientos de traslación y rotación de las moléculas completas según las ecuaciones de Newton-Euler:

$$M_i \ddot{R}_i = F_i \quad (3.33a)$$

$$J_i \dot{\omega}_i - \omega_i \times (J_i \omega_i) = N_i \quad (3.33b)$$

donde:

- $\mathbf{R}_i$  es la posición del centro de masas de la molécula  $i$ ,  $\mathbf{R}_i = \frac{1}{M_i} \sum_{\alpha} m_{i\alpha} \mathbf{r}_{i\alpha}$ .  
 $M_i$  es la masa total de la molécula  $i$ ,  $M_i = \sum_{\alpha} m_{i\alpha}$ .  
 $\mathbf{F}_i$  es la fuerza total en el centro de masas de la molécula  $i$ ,  $\mathbf{F}_i = \sum_{\alpha\beta} \mathbf{f}_{i\alpha\beta}$ .  
 $\mathbf{f}_{i\alpha\beta}$  es la fuerza existente entre el átomo  $\alpha$  de la molécula  $i$  y el átomo  $\beta$  de la molécula  $j$ .  
 $\mathbf{J}_i$  es el tensor momento de inercia de la molécula  $i$ ,  $\mathbf{J}_i = \sum_{\alpha} m_{i\alpha} (s_{i\alpha}^2 \mathbf{I} - \mathbf{s}_{i\alpha} \mathbf{s}_{i\alpha}^T)$ .  
 $\mathbf{s}_{i\alpha}$  es el vector posición del átomo  $\alpha$  respecto al centro de masas de la molécula  $i$ ,  $\mathbf{s}_{i\alpha} = \mathbf{r}_{i\alpha} - \mathbf{R}_i$ .  
 $\boldsymbol{\omega}_i$  es la velocidad angular de la molécula  $i$ .  
 $\mathbf{N}_i$  es el momento de la fuerza total sobre la molécula  $i$ ,  $\mathbf{N}_i = \sum_{\alpha} \mathbf{s}_{i\alpha} \times \mathbf{f}_{i\alpha}$ .

La simulación de un sistema descrito por estas ecuaciones puede realizarse de manera elegante representando las orientaciones moleculares por medio de «cuaterniones» [120, 121]. Los cuaterniones son vectores de cuatro componentes,  $\mathbf{q} = (q_0, q_1, q_2, q_3)$ , cuyas operaciones siguen unas reglas algebraicas particulares. Una orientación determinada por los ángulos de Euler  $\theta$ ,  $\phi$  y  $\psi$  puede definirse mediante un cuaternión unitario  $\mathbf{q}$  tal que:

$$\begin{aligned}
 q_0 &= \cos \frac{\phi + \psi}{2} \cos \frac{\theta}{2} & q_1 &= \sin \frac{\phi - \psi}{2} \sin \frac{\theta}{2} \\
 q_2 &= \cos \frac{\phi - \psi}{2} \sin \frac{\theta}{2} & q_3 &= \sin \frac{\phi + \psi}{2} \cos \frac{\theta}{2}
 \end{aligned} \tag{3.34}$$

y sus derivadas temporales pueden darse en función de la velocidad y aceleración angular [122]:

$$\dot{\mathbf{q}} = \frac{1}{2} \mathbf{q} \cdot (0, \omega_x, \omega_y, \omega_z) \tag{3.35}$$

$$\ddot{\mathbf{q}} = \frac{1}{2} \mathbf{q} \cdot (-2(\dot{q}_0^2 + \dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2), \dot{\omega}_x, \dot{\omega}_y, \dot{\omega}_z) \tag{3.36}$$

donde los productos siguen la mencionada álgebra de cuaterniones.

Este tratamiento tiene la ventaja de que elimina las singularidades que tienen lugar cuando se emplean ángulos de Euler y su estabilidad numérica es mucho mayor.

Los métodos de dinámica molecular permiten el estudio no sólo de las propiedades termodinámicas de un sistema, sino también de las propiedades cinéticas o dependientes del tiempo, como coeficientes de difusión, viscosidades o conductividades térmicas.

### 3.3.2. Métodos combinados de mecánica cuántica y molecular

La combinación de las técnicas de mecánica molecular con formulaciones mecano cuánticas da lugar a los denominados «métodos QM/MM» (*Quantum Mechanics/*



*Molecular Mechanics*). En éstos, un pequeño número de átomos son tratados mediante métodos cuánticos, mientras que el resto del sistema se rige por las ecuaciones de la mecánica molecular. Los primeros métodos QM/MM fueron desarrollados por Warshel y col. [95, 99, 123, 124] para el estudio de reacciones enzimáticas haciendo uso del método EVB. Desde entonces, numerosos grupos han desarrollado sus propios métodos QM/MM.

Los primeros modelos se limitaban a la determinación de energías y caminos de reacción de moléculas en interacción con un medio representado por cargas fijas, empleando métodos de minimización de energía y sin tener en cuenta el espacio de configuraciones del sistema. Field y col. [125] llevaron a cabo simulaciones en las que se combinaba el campo de fuerzas de mecánica molecular CHARMM y el método cuántico semiempírico AM1, lo que supuso el inicio de los métodos QM/MM tal como se conocen hoy en día.

En estos métodos, el sistema se divide en dos porciones, la parte que merece más atención se trata cuánticamente y en su hamiltoniano se introduce, a modo de perturbación, la influencia del resto del sistema, que se resuelve mediante la mecánica molecular. El hamiltoniano total queda como:

$$\hat{H} = \hat{H}_{\text{QM}} + \hat{H}_{\text{QM/MM}} + \hat{H}_{\text{MM}} \quad (3.37)$$

donde  $\hat{H}_{\text{QM}}$  es el hamiltoniano del subsistema cuántico, tal como se indica en la ecuación (2.4c);  $\hat{H}_{\text{MM}}$  es la energía MM del subsistema clásico, generalmente expresada como en la ecuación (3.10). El término de interacción entre ambos subsistemas,  $\hat{H}_{\text{QM/MM}}$ , puede expresarse como una suma de tres componentes:

$$\hat{H}_{\text{QM/MM}} = \hat{H}_{\text{QM/MM}}^{\text{elec}} + \hat{H}_{\text{QM/MM}}^{\text{pol}} + \hat{H}_{\text{QM/MM}}^{\text{vdw}} \quad (3.38)$$

donde  $\hat{H}_{\text{QM/MM}}^{\text{elec}}$  es la interacción electrostática,  $\hat{H}_{\text{QM/MM}}^{\text{pol}}$  es la contribución de la polarización del subsistema clásico, presente únicamente en el caso de que se considere explícitamente su polarización electrónica, y  $\hat{H}_{\text{QM/MM}}^{\text{vdw}}$  es el término de Van der Waals, que es una manera aproximada de considerar las repulsiones electrónicas e interacciones de dispersión.

Puesto que el término  $\hat{H}_{\text{MM}}$  no depende de las coordenadas de la parte cuántica, contribuye a la energía con un valor constante para una configuración determinada,  $E_{\text{MM}}$ , y puede separarse del hamiltoniano total. Se tiene entonces la siguiente ecuación de Schrödinger:

$$\left( \hat{H}_{\text{QM}} + \hat{H}_{\text{QM/MM}} \right) |\Psi\rangle = E |\Psi\rangle \quad (3.39)$$

donde  $E$  es la energía potencial de la parte cuántica, incluyendo la interacción con la parte clásica. La suma de  $E$  y  $E_{\text{MM}}$  es la energía total del sistema QM/MM.

Este método general permite hacer cálculos detallados de una parte de un sistema más complejo, pero sin despreciar del todo el resto del sistema. El subsistema cuántico, la porción del sistema que se estudia en detalle suele ser la zona de una molécula donde tiene lugar una reacción química, el cromóforo o grupo activo en una transición electrónica, el sustrato en una reacción enzimática o el soluto en el caso de disoluciones. La división entre los subsistemas cuántico y clásico puede atravesar enlaces covalentes, es decir, puede ocurrir que existan enlaces entre los átomos descritos por la mecánica cuántica y los descritos por la mecánica molecular. En este caso, es necesario tratar explícitamente este tipo de enlaces introduciendo algún tipo de término adicional, para lo cual existen distintas técnicas [126]. Para disoluciones, sin embargo, resulta natural establecer esta división entre el soluto y el disolvente, de manera que no existen enlaces covalentes entre las porciones cuántica y clásica, por lo que no es frecuente encontrar este problema.

La aplicación de los métodos QM/MM a líquidos y disoluciones y, en general, a todo sistema con importantes fluctuaciones térmicas, implica obtener un conjunto de configuraciones representativas del sistema y resolver la ecuación (3.39) para cada una de ellas. Para obtener este conjunto de configuraciones pueden emplearse las técnicas de simulación de Monte Carlo (pág. 61) o de dinámica molecular (pág. 63). En el primer caso sólo es necesario calcular la energía de cada configuración, por muestreo estocástico se va obteniendo el conjunto representativo. En el caso de la dinámica molecular, se requiere también el cálculo de las fuerzas sobre los distintos núcleos, obtenidas como derivadas de la energía total respecto a sus coordenadas.

Dependiendo de la forma de representar el disolvente, el término de interacción,  $\hat{H}_{\text{QM/MM}}$  puede tomar distintas expresiones. Si las moléculas de disolvente se representan mediante cargas puntuales  $q_s$ , el término  $\hat{H}_{\text{QM/MM}}^{\text{elec}}$  vendrá dado por:

$$\hat{H}_{\text{QM/MM}}^{\text{elec}} = - \sum_i \sum_s \frac{q_s}{r_{is}} + \sum_\alpha \sum_s \frac{q_s Z_\alpha}{r_{s\alpha}} \quad (3.40)$$

donde  $i$  son los electrones del soluto,  $s$  son las cargas del disolvente y  $\alpha$  son los núcleos del soluto. Si además se incluyen dipolos  $\mu_t^0$ :

$$\hat{H}_{\text{QM/MM}}^{\text{elec}} = - \sum_i \sum_s \frac{q_s}{r_{is}} + \sum_\alpha \sum_s \frac{q_s Z_\alpha}{r_{s\alpha}} - \sum_t \mu_t^0 E_t^{\text{QM}} \quad (3.41)$$

donde  $E_t^{\text{QM}}$  es el campo eléctrico generado por el soluto en la posición  $t$ .

El término  $\hat{H}_{\text{QM/MM}}^{\text{pol}}$  está presente sólo cuando los dipolos del disolvente se consideran variables, en cuyo caso se generan dipolos inducidos cuyo valor, en una aproximación lineal, viene dado por  $\mu_t^{\text{ind}} = \alpha_t E_t$ , donde  $\alpha$  es el tensor polarizabilidad y  $E_t$  es el campo eléctrico total en la posición  $t$ . La contribución de estos dipolos

inducidos al hamiltoniano del soluto es:

$$\hat{H}_{\text{QM/MM}}^{\text{pol}} = - \sum_t \mu_t^{\text{ind}} E_t^{\text{QM}} \quad (3.42)$$

y hay que considerar también la energía necesaria para inducir los dipolos.

Por último, el término  $\hat{H}_{\text{QM/MM}}^{\text{vdw}}$  se representa generalmente por potenciales Lennard-Jones y, al no depender de las coordenadas electrónicas del soluto, contribuye con un valor constante para cada configuración del sistema, y no es necesario incluirlo en el cálculo cuántico.

Este tipo de métodos presenta la ventaja de proporcionar una adecuada descripción de la molécula de soluto y de sus propiedades a la vez que se tiene en cuenta la estructura microscópica del disolvente. Además, si se emplean métodos de dinámica molecular para obtener las configuraciones, permiten el estudio de procesos que se desarrollan en el tiempo, como reacciones químicas. Sin embargo, el coste computacional sigue siendo muy elevado, pues se requiere efectuar un cálculo cuántico para cada una de las configuraciones del sistema que vayan a considerarse en el promedio final.

Para reducir el coste de estos cálculos se han propuesto diferentes alternativas. Una de ellas consiste en disminuir la exactitud de los cálculos cuánticos, empleando formulaciones como los métodos semiempíricos, la teoría de enlace de valencia [98, 127] o la teoría del funcional de la densidad [128-131]

Otra posibilidad es reducir el número de cálculos cuánticos a realizar, lo que se puede conseguir realizando más de un paso en la simulación QM/MM manteniendo fija la descripción cuántica del soluto [131-133].

### – El método DFT/MM

Una de las posibilidades para hacer los cálculos QM/MM más asequibles es utilizar técnicas DFT para los cálculos cuánticos. Esto es lo que se hace en el método DFT/MM desarrollado en Nancy [131, 134-136]. En este método se divide la energía total del sistema como es habitual:

$$E(\mathbf{r}_\alpha, \mathbf{r}_s) = E_{\text{QM}}(\mathbf{r}_\alpha) + E_{\text{QM/MM}}(\mathbf{r}_\alpha, \mathbf{r}_s) + E_{\text{MM}}(\mathbf{r}_\alpha) \quad (3.43)$$

donde  $\mathbf{r}_\alpha$  y  $\mathbf{r}_s$  son las coordenadas de los núcleos de la porción cuántica y clásica, respectivamente.

Tanto  $E_{\text{QM}}$  como  $E_{\text{QM/MM}}$  dependen no sólo de las coordenadas de la porción cuántica, sino de su densidad electrónica,  $\rho(\mathbf{r})$ . Esta densidad electrónica se obtiene

por resolución de las ecuaciones de Kohn-Sham (ec. 2.47):

$$\begin{aligned} \epsilon_i \phi_i(\mathbf{r}) &= \left[ -\frac{1}{2} \nabla^2 - \left( \sum_{\alpha} \frac{Z_{\alpha}}{r_{i\alpha}} + \sum_s \frac{q_s}{r_{is}} \right) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}}{\delta \rho} \right] \phi_i(\mathbf{r}) \\ \rho(\mathbf{r}) &= \sum_i |\phi_i(\mathbf{r})|^2 \end{aligned} \quad (3.44)$$

donde  $i$  y son los electrones de la parte cuántica y  $q_s$  son las cargas atómicas del disolvente.

A continuación, una vez hallada la densidad electrónica para una determinada configuración de soluto y disolvente, las componentes energéticas se calculan como:

$$\begin{aligned} E_{QM}(\mathbf{r}_{\alpha}) &= -\frac{1}{2} \int \sum_i \phi_i(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) d\mathbf{r} + \int \sum_{\alpha} \frac{Z_{\alpha}}{|\mathbf{r}_{\alpha} - \mathbf{r}|} \rho(\mathbf{r}) d\mathbf{r} + \\ &+ \frac{1}{2} \iint \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[\rho] + \sum_{\alpha < \beta} \sum \frac{Z_{\alpha} Z_{\beta}}{r_{\alpha\beta}} \end{aligned} \quad (3.45a)$$

$$\begin{aligned} E_{QM/MM}(\mathbf{r}_{\alpha}, \mathbf{r}_s) &= \int \sum_s \frac{q_s}{|\mathbf{r}_s - \mathbf{r}|} \rho(\mathbf{r}) d\mathbf{r} + \sum_{\alpha} \sum_s \frac{Z_{\alpha} q_s}{r_{\alpha s}} + \\ &+ \sum_{\alpha} \sum_s 4\epsilon_{\alpha s} \left[ \left( \frac{\sigma_{\alpha s}}{r_{\alpha s}} \right)^{12} - \left( \frac{\sigma_{\alpha s}}{r_{\alpha s}} \right)^6 \right] \end{aligned} \quad (3.45b)$$

$$E_{MM}(\mathbf{r}_s) = \sum_{s < t} \sum \frac{q_s q_t}{r_{st}} + \sum_{s < t} 4\epsilon_{st} \left[ \left( \frac{\sigma_{st}}{r_{st}} \right)^{12} - \left( \frac{\sigma_{st}}{r_{st}} \right)^6 \right] \quad (3.45c)$$

Las fuerzas que actúan sobre los núcleos clásicos y cuánticos se calculan analíticamente como las derivadas de la energía con respecto a sus posiciones respectivas, es decir:

$$\mathbf{F}_{\alpha} = -\frac{\partial E(\mathbf{r}_{\alpha}, \mathbf{r}_s)}{\partial \mathbf{r}_{\alpha}} \quad \mathbf{F}_s = -\frac{\partial E(\mathbf{r}_{\alpha}, \mathbf{r}_s)}{\partial \mathbf{r}_s} \quad (3.46)$$

Una vez que se tienen las energías y las fuerzas, pueden aplicarse las técnicas de simulación de Monte Carlo o de dinámica molecular. En concreto, para las simulaciones dinámicas, el método es el siguiente:

1. Para una configuración determinada de soluto y disolvente, se resuelven las ecuaciones de Kohn-Sham (ec. 3.44).
2. Se calculan las energías totales (ec. 3.45) y las fuerzas (ec. 3.46).
3. Se actualizan las propiedades promedio y se almacenan los orbitales moleculares para utilizarlos como orbitales iniciales en el siguiente cálculo cuántico.
4. Se resuelven las ecuaciones de movimiento para las moléculas clásicas (consideradas rígidas) y para los átomos cuánticos.

Este proceso se repite tantas veces como pasos de dinámica se quieran realizar, de manera que normalmente es necesario efectuar un número de cálculos bastante elevado, del orden de decenas de miles o más.



---

---

# 4

---

## EL MÉTODO ASEP/MD

Como se ha visto anteriormente, los métodos más eficientes para el estudio de moléculas en disolución son los métodos mixtos, en los que la molécula de interés —el soluto— se trata con algún tipo de método cuántico mientras que el resto —el disolvente— se representa de una manera más simplificada, mediante potenciales clásicos. Dentro de estos métodos existen dos grandes grupos. Por un lado, los métodos discretos proporcionan una información detallada y valiosa sobre la estructura del disolvente. Por otro lado, los métodos que hacen uso del modelo del continuo, requieren un menor esfuerzo computacional, con lo que permiten mejorar la descripción que se hace del soluto.

Durante los últimos años se ha venido desarrollando en nuestro grupo de investigación un nuevo método mixto que intenta aunar las virtudes de ambos grupos. Este método ha sido denominado ASEP/MD (*Averaged Solvent Electrostatic Potential from Molecular Dynamics*) o «potencial electrostático promedio del disolvente obtenido a partir de datos de dinámica molecular» [40-42, 137]. En él, y al igual que ocurre en los modelos de continuo, el efecto del disolvente se introduce de una forma promediada, a través del potencial electrostático que genera sobre la molécula de soluto; sin embargo, este potencial promedio se obtiene de simulaciones de dinámica molecular y no en función de magnitudes macroscópicas, como pueda ser la constante dieléctrica.

El método ASEP/MD se basa, como los métodos QM/MM, en la combinación de técnicas de mecánica cuántica y mecánica molecular; pero, a diferencia de éstos, la perturbación que se introduce en el hamiltoniano del soluto representa el efecto promedio ejercido por el disolvente, con lo que se reduce drásticamente el número

de cálculos cuánticos a realizar.

En este método se emplea la separación del hamiltoniano usual en los métodos QM/MM (ecuación 3.37). En éstos, el término  $\hat{H}_{\text{QM/MM}}^{\text{elec}}$  también puede expresarse como:

$$\hat{H}_{\text{QM/MM}}^{\text{elec}} = \sum_s q_s \hat{V}_s^{\text{sol}} = \int \hat{\rho}(\mathbf{r}) V^{\text{dis}}(\mathbf{r}) d\mathbf{r} \quad (4.1)$$

donde  $\hat{V}_s^{\text{sol}}$  es el potencial generado por el soluto en la posición ocupada por la carga  $q_s$ ,  $\hat{\rho}(\mathbf{r})$  es el operador densidad de carga del soluto y  $V^{\text{dis}}(\mathbf{r})$  es el potencial generado por el disolvente en el punto  $\mathbf{r}$ . En el método ASEP/MD, en lugar de realizar un cálculo cuántico por cada configuración del disolvente, se introduce el potencial electrostático promedio del disolvente,  $V_m(\mathbf{r})$ , que no es más que el promedio termodinámico de los potenciales creados por el disolvente en cada una de las configuraciones consideradas:

$$V_m(\mathbf{r}) = \langle V^{\text{dis}}(\mathbf{r}) \rangle \quad (4.2)$$

Por lo tanto, el hamiltoniano del soluto perturbado por el efecto promediado del disolvente es:

$$\hat{H} = \hat{H}^0 + \int \hat{\rho}(\mathbf{r}) V_m(\mathbf{r}) d\mathbf{r} + \hat{H}_{\text{QM/MM}}^{\text{vdw}} + \hat{H}_{\text{MM}} \quad (4.3)$$

donde  $\hat{H}^0$  es el hamiltoniano en vacío del soluto, equivalente a  $\hat{H}_{\text{QM}}$  en la ecuación (3.37). Los términos  $\hat{H}_{\text{MM}}$  y  $\hat{H}_{\text{QM/MM}}^{\text{vdw}}$  como ya se ha mencionado, contribuyen con valores constantes a la energía del sistema y pueden añadirse posteriormente.

Los métodos de continuo, que también hacen uso de la aproximación del campo medio, obtienen el potencial  $V_m(\mathbf{r})$  de distintas maneras. En algunos casos se utiliza un desarrollo multipolar para describir el potencial electrostático de interacción entre el soluto y el disolvente, desarrollo que debe truncarse en los primeros términos para que pueda ser manejable, lo que introduce ciertos errores y ocasiona problemas de convergencia. La principal característica del método ASEP/MD es que, como indica su nombre, calcula el potencial promedio a partir de los resultados de simulaciones de dinámica molecular. El empleo de técnicas de dinámica molecular para la representación del disolvente permite obtener una descripción adecuada tanto de su estructura como de sus propiedades termodinámicas y cinéticas, mientras que la aplicación de la aproximación del campo medio hace posible la utilización de métodos cuánticos de alta calidad para el tratamiento del soluto.

## 4.1. Descripción del método

A continuación se presentan con más detalle las distintas etapas que componen el método ASEP/MD. Las condiciones iniciales que deben especificarse antes de

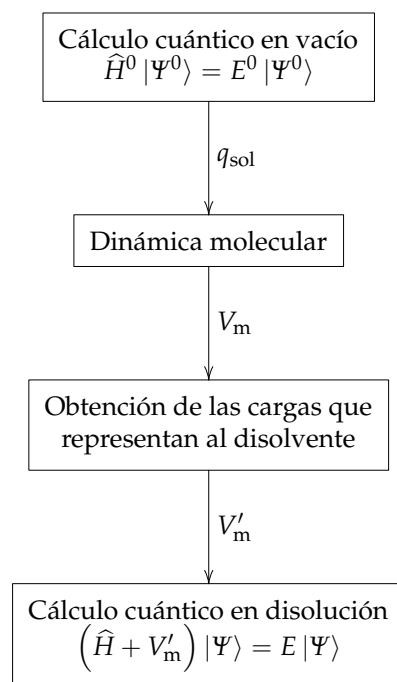


comenzar el cálculo incluyen el método cuántico que se empleará para el soluto —adecuado a su naturaleza y a las propiedades que deseen estudiarse—, así como el potencial clásico —tipo de potencial y parámetros— con que interaccionarán las distintas moléculas en las simulaciones de dinámica molecular.

1. **Cálculo en vacío.** El primer paso es realizar un cálculo cuántico de la molécula de soluto en vacío empleando el método elegido. Este cálculo sirve como punto de partida para el método ASEP/MD y también como referencia para comparar los resultados obtenidos en disolución.
2. **Cargas del soluto.** A partir del cálculo en vacío se obtiene un conjunto de cargas atómicas para el soluto. Para ello se establece la condición de que estas cargas, situadas generalmente en los núcleos de los átomos del soluto, deben reproducir lo más exactamente posible el potencial electrostático generado por el soluto en el espacio circundante. En algunos casos puede ser necesario añadir cargas adicionales sobre enlaces o pares de electrones [138].
3. **Dinámica molecular.** Con las cargas atómicas del soluto y el campo de fuerzas clásico definido inicialmente se lleva a cabo una simulación completa de dinámica molecular, en la que la molécula de soluto está rodeada de un número determinado de moléculas de disolvente; se realiza con geometría y cargas fijas tanto para el soluto como para el disolvente. De esta simulación, una vez que se ha alcanzado el equilibrio dinámico, se extrae una serie de configuraciones de las moléculas del sistema.
4. **Potencial promedio del disolvente.** Para cada una de las configuraciones del paso 3, se calcula el potencial electrostático que las moléculas de disolvente generan en la zona ocupada por el soluto. Primero se define un conjunto de puntos dentro del volumen de Van der Waals de la molécula del soluto y a continuación se calcula el potencial en cada uno de estos puntos. Posteriormente se calcula el promedio de estas configuraciones: en cada punto se suman los potenciales y se divide entre el número de configuraciones consideradas.
5. **Representación del potencial promedio del disolvente por un conjunto de cargas discretas.** Una vez que se tiene el potencial electrostático promedio del disolvente en forma numérica (el valor del potencial en una serie de puntos), se obtiene un conjunto de cargas que lo reproducen. Se definen las posiciones de las cargas alrededor del soluto y se calculan sus valores de manera que el potencial que generan reproduzca el calculado en el paso 4.
6. **Cálculo en disolución.** Finalmente, se realiza un nuevo cálculo cuántico del soluto, esta vez incluyendo el conjunto de cargas externas obtenido en el paso 5. En este cálculo, por lo tanto, está presente la influencia del disolvente y proporciona el resultado final.

Este proceso se muestra esquemáticamente en la figura 4.1, donde  $q_{\text{sol}}$  designa a las cargas atómicas del soluto,  $V_m$  es el potencial electrostático promedio del disolvente y  $V'_m$  es el potencial generado por las cargas ajustadas obtenidas en el paso 5:

$$V'_m(\mathbf{r}) = \sum_i \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|} \quad (4.4)$$



**Figura 4.1:** Esquema del método ASEP/MD.

De esta manera, existen dos descripciones complementarias —aunque no simultáneas— del sistema. Durante la simulación de dinámica molecular, todas las moléculas, tanto el soluto como el disolvente, se representan mediante cargas puntuales que interactúan por medio de campos de fuerza clásicos y cuyas posiciones relativas evolucionan en el tiempo. En el cálculo cuántico en disolución, el soluto se representa mediante su función de onda cuántica, mientras que el disolvente es un conjunto de cargas fijas que ejercen un potencial electrostático sobre el soluto.

Los pasos más característicos del método son los puntos 4 y 5: el cálculo del potencial electrostático promedio del disolvente y del conjunto de cargas lo reproducen. En principio se podría intentar una aproximación directa para calcular el potencial promedio del disolvente. Esto es, se podría incluir en  $\hat{H}_{\text{QM}}^{\text{elec}}$  todas las cargas del disolvente de cada una de las configuraciones seleccionadas y con un

valor para cada carga  $q/N$ . De esta forma el potencial promedio se expresaría como:

$$V_m(\mathbf{r}) = \langle V^{\text{dis}}(\mathbf{r}) \rangle = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^s \frac{q_{ji}}{|\mathbf{r}_{ji} - \mathbf{r}|} \quad (4.5)$$

donde  $N$  es el número total de configuraciones consideradas y  $s$ , el número de cargas en cada configuración. Sin embargo, este procedimiento llevaría a un número excesivo de cargas. Por ejemplo, si  $N = 2000$  y si tenemos 200 moléculas de agua, el total de cargas sería  $2000 \times 200 \times 3 = 1\,200\,000$ . Es por ello que optamos por la estrategia que se indica a continuación.

**Selección de los puntos donde se calculará el potencial.** A la hora de calcular el potencial electrostático que el disolvente ejerce sobre el soluto, interesa hacerlo en la zona donde su densidad electrónica es mayor, ya que es aquí donde la influencia sobre la función de onda será más importante. Esta zona queda delimitada por una superficie imaginaria en cuyo interior se dispone una serie de puntos. Será en estos puntos donde se calcule el potencial electrostático promedio.

Para definir la «zona ocupada por el soluto» se emplean los radios de Van der Waals de los átomos del soluto. Se consideran pertenecientes a esta zona todos los puntos cuya distancia a alguno de los núcleos del soluto es menor que el correspondiente radio de Van der Waals, multiplicado por un determinado factor  $f$ . Esto resulta en una unión de esferas centradas en los núcleos del soluto y con radios proporcionales a los radios de Van der Waals.

Dentro de este espacio se genera una red tridimensional dividiendo la máxima extensión de la zona ocupada por el soluto en cada una de las tres direcciones del espacio ( $x$ ,  $y$  y  $z$ ) en  $n$  partes. En los nodos de esta red se sitúan los puntos donde se calculará el potencial electrostático del disolvente. El número de puntos resultante dependerá de la forma de la molécula de soluto.

En trabajos anteriores [40] se estudió la influencia de los valores tomados para  $f$  y  $n$  en los resultados obtenidos, las conclusiones indican que valores de 0,7 para  $f$  y 12 para  $n$  representan un compromiso aceptable entre calidad de los resultados y coste computacional.

**Cálculo del potencial electrostático promedio.** Para cada configuración del disolvente obtenida en la simulación de dinámica molecular,  $X$ , se calcula el potencial electrostático que las moléculas de disolvente generan sobre los distintos puntos de la zona ocupada por el soluto,  $r_j$  (figura 4.2-a). Cada molécula de disolvente está representada por un número de cargas puntuales, el potencial que generan las

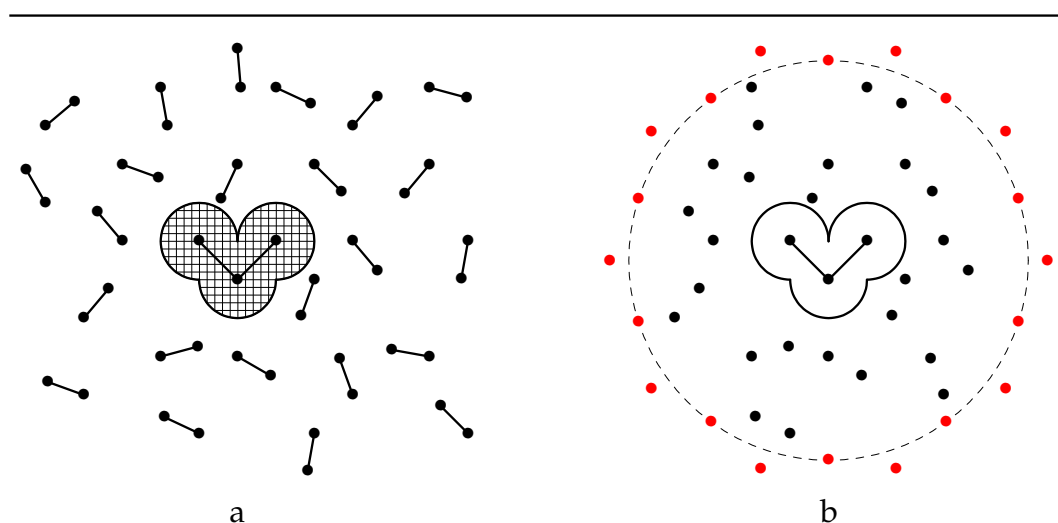
cargas de todas las moléculas,  $q_s$ , se calcula como:

$$V(\mathbf{r}_j, X) = \sum_s \frac{q_s}{r_{sj}} \quad (4.6)$$

donde  $r_{sj}$  es la distancia entre la carga  $q_s$  y el punto  $\mathbf{r}_j$ .

Una vez se tienen  $N$  valores de potencial (uno por cada configuración) en cada uno de los puntos  $\mathbf{r}_j$ , se calcula el potencial promedio sin más que calcular la media aritmética en cada punto:

$$V_m(\mathbf{r}_j) = \frac{1}{N} \sum_X V(\mathbf{r}_j, X) \quad (4.7)$$



**Figura 4.2:** Representación del disolvente en el método ASEP/MD. a) El potencial electrostático generado por el disolvente se calcula en una red de puntos situados en la zona del espacio ocupada por el soluto. b) En el cálculo cuántico se introducen dos conjuntos de cargas: las cargas interiores (en negro) pertenecen a las moléculas más cercanas al soluto, el resto del disolvente se representa mediante las cargas exteriores (en rojo).

**Determinación de las posiciones de las cargas.** El potencial calculado hasta ahora no tiene una forma adecuada para introducirlo en los cálculos cuánticos, ya que se halla en forma numérica y es necesario contar con una forma analítica que permita conocer el potencial en cualquier punto del espacio. Con este fin se define un conjunto de cargas,  $q_i$ , de manera que su potencial electrostático en los puntos  $\mathbf{r}_j$ ,  $V'_m(\mathbf{r}_j)$ , reproduzca lo más exactamente posible el potencial promedio  $V_m(\mathbf{r}_j)$ .

Estas cargas se dividen en dos grupos que denominamos «cargas interiores» y «cargas exteriores» (figura 4.2-b). Se establece una cierta distancia o radio de corte,

$r_c$ , a partir del centro de masas de la molécula de soluto, de manera que abarque al menos la primera capa de solvatación. Para cada configuración del disolvente considerada, se incluyen en el grupo de cargas interiores las cargas pertenecientes a todas las moléculas de disolvente situadas a una distancia menor o igual a  $r_c$  del centro de masas de la molécula de soluto.

A una distancia mayor que  $r_c$  se distribuyen las cargas exteriores, cuya función será representar la porción del potencial del disolvente no incluida en las cargas interiores. Estas cargas se sitúan en los vértices de un icosaedro truncado —son, por lo tanto, 60 cargas— y se reparten, de manera aproximadamente alternada, en dos capas, que se encontrarán a distancias ligeramente diferentes para evitar posibles artefactos.

**Determinación de los valores de las cargas del disolvente.** A las cargas interiores se les asigna un valor igual a la carga empleada durante la dinámica molecular dividida entre el número de configuraciones consideradas,  $N$ . Además para reducir el número de cargas y evitar la presencia de cargas excesivamente próximas entre sí, se agrupan todas las cargas cuya separación sea menor que un valor dado,  $r_p$ . De esta manera, la distancia entre dos cargas interiores cualesquiera nunca será menor que  $r_p$ .

Se calcula el potencial generado por las cargas interiores en los puntos  $\mathbf{r}_j$  de la zona ocupada por el soluto,  $V_m^{\text{int}}(\mathbf{r}_j)$ . Al restar este potencial del potencial promedio del disolvente se obtiene un potencial residual:

$$\begin{aligned} V_m^{\text{res}}(\mathbf{r}_j) &= V_m(\mathbf{r}_j) - V_m^{\text{int}}(\mathbf{r}_j) \\ V_m^{\text{int}}(\mathbf{r}_j) &= \sum_i \frac{q_i^{\text{int}}}{r_{ij}} \end{aligned} \quad (4.8)$$

Finalmente, los valores de las cargas exteriores se ajustan para que reproduzcan este potencial residual. El ajuste se realiza como se detalla en el apéndice A (pág. 191). Con este procedimiento se consigue que el potencial generado por el conjunto de todas las cargas interiores y exteriores,  $V_m'(\mathbf{r}_j)$ , reproduzca adecuadamente el potencial promedio  $V_m(\mathbf{r}_j)$ .

**Determinación de las cargas del soluto empleadas durante la dinámica molecular.** Las cargas atómicas del soluto se definen de una manera similar (paso 2 en la pág. 75). Las cargas se sitúan generalmente en los núcleos de los átomos del soluto y los puntos donde se calcula el potencial empleado en el ajuste se encuentran en su entorno. Existen distintos métodos para el ajuste de las cargas al potencial que se diferencian, principalmente, en la disposición de los puntos donde se calcula este potencial alrededor del soluto. Dos de los métodos más empleados

son los conocidos como «cargas a partir del potencial electrostático» (*Charges from Electrostatic Potential*, CHELP) [139] y «cargas a partir del potencial electrostático en una red» (*Charges from Electrostatic Potential Grid*, CHELPG) [140].

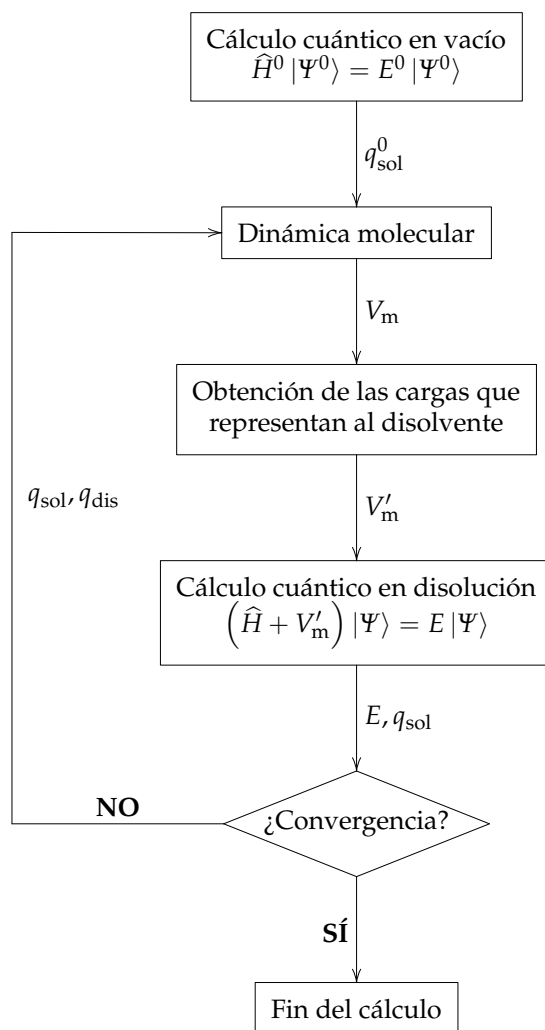
## 4.2. Métodos acoplados

El procedimiento expuesto en la sección anterior y esquematizado en la figura 4.1 es el denominado «método no acoplado». Como se ha visto, este método permite realizar un cálculo cuántico de la molécula de soluto en presencia del potencial electrostático promedio del disolvente. Sin embargo, el efecto del disolvente se obtiene a partir de una simulación de dinámica molecular en la que la molécula de soluto está representada por sus cargas atómicas calculadas en vacío, es decir, soluto y disolvente no están realmente en equilibrio: el soluto se ve afectado por el disolvente, pero el disolvente no se ve afectado por el nuevo soluto.

### 4.2.1. Método parcialmente acoplado

Para obviar el inconveniente anterior se han desarrollado los «métodos acoplados» [42], en los que se realiza un proceso coherente de polarización mutua entre soluto y disolvente. Esto se consigue introduciendo las cargas atómicas del soluto calculadas en disolución en una nueva simulación de dinámica molecular; el proceso se repite hasta que las variaciones en las cargas del soluto y en la energía de interacción soluto-disolvente son mínimas. Éste es el «método parcialmente acoplado», cuyo esquema se muestra en la figura 4.3. Los pasos más importantes son:

1. **Dinámica molecular inicial.** Con las cargas atómicas obtenidas a partir del cálculo cuántico en vacío,  $q_{\text{sol}}^0$ , se lleva a cabo una simulación de dinámica molecular. Se obtiene la estructura del disolvente en equilibrio con la distribución electrónica del soluto en vacío.
2. **Cálculo cuántico en disolución.** Una vez obtenido el potencial electrostático promedio del disolvente a partir de los resultados de la simulación anterior, se introduce en un nuevo cálculo cuántico del soluto. Se obtiene la distribución electrónica del soluto en equilibrio con la estructura del disolvente.
3. **Nueva dinámica molecular.** La distribución electrónica del soluto —representada por un conjunto de cargas puntuales— se utiliza para efectuar una nueva simulación de dinámica molecular. Esta vez, la estructura del disolvente estará en equilibrio con la distribución electrónica del soluto en disolución.



**Figura 4.3:** Esquema del método ASEP/MD acoplado. En el método parcialmente acoplado la distribución de cargas de las moléculas de disolvente,  $q_{\text{dis}}$ , permanece constante; en el método totalmente acoplado,  $q_{\text{dis}}$  se iguala a  $q_{\text{sol}}$  en cada iteración.

4. **Repetición del proceso.** Se vuelve a realizar el paso 2. El ciclo se repite hasta que la distribución electrónica del soluto se mantenga prácticamente constante, entonces, tanto ésta como la estructura del disolvente estarán en equilibrio mutuo y pueden analizarse los resultados.

#### 4.2.2. Método totalmente acoplado

En cada repetición del proceso anterior se modifica la distribución electrónica de la molécula de soluto; sin embargo, la de las moléculas de disolvente permanece constante. Esto resulta adecuado para el tratamiento de disoluciones, en las que soluto y disolvente tienen diferente naturaleza y, por lo tanto, pueden representarse con modelos distintos. En el caso de líquidos puros, por otra parte, la distinción entre la molécula de soluto y las de disolvente resulta artificiosa ya que, en principio, es imposible diferenciar una molécula de las demás. Por supuesto, la distribución electrónica de una molécula depende de su entorno, y éste es diferente, en un instante dado, para cada molécula del líquido; no obstante, dado que se está empleando la aproximación del campo medio, podemos suponer que, en promedio, el entorno de todas las moléculas será equivalente y, en consecuencia, también sus distribuciones electrónicas.

En el «método totalmente acoplado», también representado en la figura 4.3, se aprovecha esta equivalencia de todas las moléculas dentro de un líquido puro. En cada repetición del proceso iterativo, no sólo se reintroduce la distribución de cargas de la molécula de soluto en una nueva dinámica molecular, sino que, además, se emplea esta misma distribución de cargas para representar las moléculas de disolvente. De esta manera, en cada simulación de dinámica molecular todas las moléculas son completamente equivalentes y, aunque los cálculos cuánticos se centran en una de ellas, sus resultados se aplican a todas las moléculas del líquido por igual. En la práctica esto supone que durante cada iteración ASEP/MD todas las moléculas del líquido se polarizan, algo que no ocurre en la versión parcialmente acoplada, donde sólo se permite que el soluto se polarice.

La energía de interacción entre la molécula de soluto y las de disolvente puede expresarse como:

$$E^{\text{int}} = \langle \Psi | \hat{H}_{\text{QM/MM}} | \Psi \rangle = E^{\text{elec}} + E^{\text{vdw}} \quad (4.9)$$

$$E^{\text{elec}} = \sum_i q_i V_i^{\text{sol}} \quad E^{\text{vdw}} = \frac{1}{N} \sum_X E_X^{\text{vdw}}$$

donde  $E^{\text{elec}}$  no es más que la energía de las cargas que representan al disolvente en el cálculo cuántico,  $q_i$ , en el potencial electrostático generado por el soluto,  $V^{\text{sol}}$ .  $E^{\text{vdw}}$  es el promedio de las energías de interacción de Van der Waals entre soluto



y disolvente —calculadas de manera clásica— para las  $N$  configuraciones consideradas en el cálculo. Dado que todas las moléculas del sistema son equivalentes, la energía de interacción total vendrá dada por la expresión:

$$\frac{1}{2}M E^{\text{int}} = \frac{1}{2}M (E^{\text{elec}} + E^{\text{vdw}}) \quad (4.10)$$

siendo  $M$  el número total de moléculas presentes en el sistema, de forma que la energía de interacción promedio por molécula es  $\frac{1}{2}(E^{\text{elec}} + E^{\text{vdw}})$ .

Por otra parte, la distribución electrónica del soluto en disolución varía con respecto a la que presenta en vacío, este proceso consume una determinada energía que es la energía de distorsión o de polarización del soluto:

$$E^{\text{dist}} = \langle \Psi | H^0 | \Psi \rangle - \langle \Psi^0 | H^0 | \Psi^0 \rangle \quad (4.11)$$

De igual manera, esta energía será idéntica para todas las moléculas del sistema.

La magnitud susceptible de comparación con los datos experimentales es la energía de vaporización o energía potencial promedio de cada molécula, que puede calcularse fácilmente como:

$$E^{\text{vap}} = \frac{1}{2} (E^{\text{elec}} + E^{\text{vdw}}) + E^{\text{dist}} \quad (4.12)$$

### 4.3. Disolvente polarizable

Existen casos en los que es necesario tratar explícitamente la polarización electrónica de las moléculas de disolvente. Por una parte, mediante esta polarización se permite que la respuesta de las moléculas más cercanas al soluto sea diferente de la de las moléculas más alejadas. Por otra parte, se representa de manera adecuada la distinta velocidad con la que reaccionan los núcleos y electrones del disolvente frente a las perturbaciones que se puedan producir en el soluto.

En el método ASEP/MD, el modelo de disolvente polarizable se aplica principalmente en el estudio de estados excitados y transiciones electrónicas de la molécula de soluto [141-143]. Cuando ocurre una transición electrónica en el soluto, su distribución electrónica sufre un cambio brusco, así como el potencial electrostático que el soluto genera en el disolvente. Este cambio puede considerarse prácticamente instantáneo y, mientras que los electrones de las moléculas de disolvente pueden adaptarse de manera prácticamente inmediata, los núcleos reaccionan con más lentitud y mantienen sus posiciones por más tiempo. Por lo tanto, la estructura del disolvente en torno a la molécula de soluto permanece fija al pasar el soluto de un estado electrónico a otro, no así el campo electrostático que aquél genera en

éste, pues los momentos dipolares inducidos en las moléculas de disolvente serán diferentes en ambos estados electrónicos.

Para calcular la polarización del disolvente se parte de un cálculo previo en el que se considera el disolvente no polarizable. Como resultado de este cálculo previo se obtiene un conjunto de configuraciones del disolvente alrededor del soluto. Entonces se calcula nuevamente el potencial electrostático que el disolvente genera sobre el soluto, esta vez representando cada molécula mediante un conjunto de cargas permanentes y una polarizabilidad electrónica, que dará lugar a un momento dipolar inducido. Con más detalle, los pasos son los siguientes:

1. **Estructura del disolvente.** Mediante un proceso ASEP/MD total o parcialmente acoplado se obtiene la estructura del disolvente alrededor del soluto —en forma de una muestra de configuraciones soluto-disolvente— y una distribución de cargas para la molécula del soluto.
2. **Dipolos inducidos.** Para cada una de las configuraciones seleccionadas, se calcula el campo eléctrico que, sobre cada una de las moléculas de disolvente, genera el resto de las moléculas del sistema. El momento dipolar inducido sobre cada molécula estará relacionado con este campo eléctrico a través de la polarizabilidad. Puesto que los momentos dipolares inducidos generan a su vez un campo eléctrico en sus alrededores, este cálculo debe realizarse de manera iterativa. Cuando termina el proceso, los momentos inducidos y el campo eléctrico son coherentes entre sí.
3. **Potencial promedio.** Una vez que se tiene, para cada molécula de disolvente, un conjunto de cargas permanentes y un momento dipolar inducido, se calcula el potencial electrostático que ambos generan en la zona ocupada por el soluto. El potencial electrostático promedio será, de nuevo, la media aritmética de los resultados de todas las configuraciones consideradas.
4. **Cargas del disolvente.** De manera similar al método con disolvente no polarizable (pág. 78), se determinan las posiciones y los valores de las cargas que reproducen el potencial electrostático promedio del disolvente, calculado en el paso 3.
5. **Cálculo en disolución.** Las cargas obtenidas en el paso 4 se introducen como perturbación externa en un cálculo cuántico de la molécula de soluto. Se obtiene un conjunto de cargas atómicas que representa su distribución electrónica.
6. **Repetición del proceso.** Las nuevas cargas atómicas del soluto se reintroducen en un nuevo cálculo de los dipolos inducidos del disolvente (paso 2), conservándose la estructura del disolvente inicial. El proceso continúa hasta que, como en los métodos acoplados, se alcanza la convergencia.

El cálculo de los dipolos inducidos en las moléculas del disolvente (paso 2) se realiza tal como se detalla en el apéndice B.1 (pág. 193). En cada iteración del proceso se calcula el campo eléctrico generado por las cargas permanentes y por los momentos dipolares inducidos de la iteración precedente (comienzan siendo nulos); con este campo eléctrico se obtienen nuevos valores para los dipolos inducidos y, si son significativamente diferentes de los anteriores, se emplean en un nuevo cálculo del campo eléctrico. Al finalizar el proceso se tienen los momentos dipolares inducidos en equilibrio con el campo eléctrico.

Una vez determinada la polarización de las moléculas del disolvente, se calcula el potencial electrostático generado en la zona ocupada por el soluto. Al incluir dipolos puntuales, la ecuación (4.6) se modifica. Las moléculas de disolvente ya no se representan por su distribución de cargas  $q_s$  (cargas efectivas que incluyen ya el efecto promedio de la polarización), sino por sus cargas permanentes  $q_s^0$  y sus dipolos inducidos  $\mu_k$ ; el potencial en los puntos  $r_j$  puede entonces calcularse como:

$$V(r_j, X) = \sum_s \frac{q_s^0}{r_{sj}} + \sum_k \frac{\mu_k r_{kj}}{r_{kj}^3} \quad r_{kj} = r_j - r_k \quad (4.13)$$

También se calcula el potencial generado únicamente por las cargas permanentes,  $V^0(r_j, X)$ . De nuevo, el potencial electrostático promedio del disolvente  $-V_m(r_j)$  para el total y  $V_m^0(r_j)$  para el debido a las cargas permanentes— se obtiene simplemente como una media aritmética de todas las configuraciones consideradas.

A la hora de obtener las cargas que reproducen el potencial promedio se procede de forma completamente similar al método normal (ver pág. 78), con la salvedad de que en lugar de utilizar las cargas empleadas durante la dinámica molecular para los valores de las cargas interiores, se toman las cargas permanentes de las moléculas de disolvente. Se obtienen además dos conjuntos de cargas exteriores, uno de ellos reproduce el potencial  $V_m(r_j)$  y el otro,  $V_m^0(r_j)$ ; cada uno de ellos formará parte, junto con las cargas interiores, de las cargas que representan al disolvente:  $q_i$  y  $q_i^0$ , respectivamente.

Las cargas  $q_i$  se introducen en el cálculo cuántico, a partir del cual se obtiene una nueva distribución de cargas del soluto. Esta nueva distribución de cargas vuelve a emplearse en el cálculo de los dipolos inducidos del disolvente y el proceso se repite hasta que se alcanza la convergencia.

En este caso se pueden obtener distintas energías de interacción entre soluto y disolvente:

$$E^{\text{elec}} = \sum_i q_i V_i^{\text{sol}} \quad (4.14)$$

$$E_{\rho q}^{\text{elec}} = \sum_i q_i^0 V_i^{\text{sol}} \quad (4.15)$$

donde  $E^{\text{elec}}$  es la energía de interacción electrostática total y  $E_{\rho\text{q}}^{\text{elec}}$  es la energía debida únicamente a las cargas permanentes. La energía de interacción del soluto con los dipolos inducidos del disolvente puede calcularse de manera indirecta:

$$E_{\rho\mu}^{\text{elec}} = E^{\text{elec}} - E_{\rho\text{q}}^{\text{elec}} = \sum_i (q_i - q_i^0) V_i^{\text{sol}} = \sum_i q_i^{\mu} V_i^{\text{sol}} \quad (4.16)$$

siendo  $q_i^{\mu}$  las cargas que reproducen el potencial debido únicamente a los dipolos inducidos del disolvente.

Al tener en cuenta las energías necesarias para polarizar tanto la molécula de soluto como las de disolvente —desarrolladas en el apéndice B.2 (pág. 194)—, se obtiene para la energía potencial total del sistema:

$$E = \frac{1}{2}(E^{\text{elec}} + E_{\rho\text{q}}^{\text{elec}}) + E^{\text{dist}} + \left[ \frac{1}{2}E_{\text{q}\mu} + E_{\text{q}\text{q}} + E^{\text{vdw}} \right] \quad (4.17)$$

Los términos entre corchetes implican únicamente al disolvente, es decir, a la porción del sistema descrita mediante MM, y se calculan fácilmente como promedio de las  $N$  configuraciones consideradas. De estos términos, sólo  $E_{\text{q}\mu}$  se ve afectado por una alteración en la distribución electrónica del soluto, por lo tanto, si se estudian diferencias energéticas entre dos estados del soluto —como es el caso de las transiciones electrónicas—,  $E_{\text{q}\text{q}}$  y  $E^{\text{vdw}}$  serán constantes y pueden desestimarse.

#### 4.4. La componente Stark

En los métodos basados en la aproximación del campo medio (ACM), la energía y la función de onda de la molécula de soluto se obtienen en presencia de la perturbación promedio generada por el disolvente. Como cualquier otra aproximación, la ACM introduce errores en el cálculo de la energía de interacción entre soluto y disolvente, así como en otras propiedades moleculares del soluto. Estos errores son debidos al hecho de que no se tiene en cuenta la correlación que existe entre el movimiento de los núcleos del disolvente y la distribución electrónica del soluto, es decir, la ACM no permite que el soluto se polarice en respuesta a un cambio en la configuración del disolvente, lo que produciría una estabilización adicional del sistema. Esta energía de correlación se conoce generalmente como componente Stark [144, 145].

A pesar de que el uso de la ACM en las teorías de efecto del disolvente está muy extendido, pocos estudios se han centrado en la evaluación de la componente Stark. Linder [144] y más tarde Karlström y Halle [145] propusieron expresiones para esta componente basadas en modelos del continuo. En ambos casos, la magnitud de la componente Stark se relacionaba con la polarizabilidad del soluto y la diferencia entre los factores de reacción del dipolo evaluados a frecuencia nula y a frecuencias

ópticas:

$$\Delta W_{\text{Stark}} = -\frac{1}{2}k_{\text{B}}T [\alpha : (g_0 - g_\infty)] \quad (4.18)$$

donde  $\alpha$  es la polarizabilidad del soluto y  $g$  es el factor de Onsager evaluado a frecuencia nula ( $g_0$ ) y a frecuencias ópticas ( $g_\infty$ ); el símbolo  $:$  denota el producto escalar doble,  $A:B = \sum \sum A_{ij}B_{ij}$ . Como otras expresiones basadas en teorías de continuo, la ecuación (4.18) desprecia los efectos de la estructura microscópica del disolvente y de las interacciones específicas. Además, es necesario asignar un valor al radio de la cavidad que contiene el soluto, y esto no siempre es sencillo [146-149].

El método ASEP/MD, como se ha indicado anteriormente, emplea simulaciones de dinámica molecular para describir el disolvente, por lo que carece de muchas de las limitaciones de los modelos del continuo. Sin embargo, también hace uso de la ACM, lo que lo expone a los errores introducidos por ésta. El éxito, en general de los métodos basados en la ACM hace pensar que estos errores deben de ser poco importantes, pero una evaluación correcta de su magnitud requiere comparar los resultados de un cálculo en el que se aplica esta aproximación con los de otro cálculo realizado en las mismas condiciones, pero sin aplicarla.

#### 4.4.1. Evaluación aproximada de la componente Stark

Para la componente Stark —el error introducido por la ACM en la energía de interacción soluto-disolvente— puede obtenerse una expresión aproximada que permite, de una manera sencilla, evaluar su importancia. Asumiendo que la distribución de carga del soluto está convenientemente representada por su momento dipolar, la energía de interacción para cada configuración del sistema,  $X$ , vendrá dada por:

$$W(X) = -\mu(X) \cdot E(X) + \frac{1}{2}\Delta\mu(X) \cdot E(X) \quad (4.19)$$

donde  $\mu$  es el momento dipolar del soluto en disolución,  $E$  es el campo eléctrico generado por el disolvente sobre el soluto y  $\Delta\mu$  es el momento dipolar inducido en el soluto con respecto a su valor en vacío. El segundo sumando de esta ecuación corresponde a la energía necesaria, según la teoría electrostática clásica, para crear el dipolo inducido en el soluto. Expresando  $\mu$  y  $\Delta\mu$  en función del momento dipolar en vacío,  $\mu^0$ , y de la polarizabilidad,  $\alpha$ , del soluto, se obtiene:

$$\begin{aligned} \mu &= \mu^0 + \Delta\mu & \Delta\mu &= \alpha E \\ W(X) &= -\mu^0 \cdot E(X) - \frac{1}{2} [\alpha E(X)] \cdot E(X) = \\ &= -\mu^0 \cdot E(X) - \frac{1}{2} \alpha : [E(X) \otimes E(X)] \end{aligned} \quad (4.20)$$

donde el símbolo  $\otimes$  denota el producto tensorial,  $a \otimes b = ab^T$ , y se supone que la polarizabilidad del soluto es constante y no varía de una configuración del disolvente a otra.

Finalmente, la energía de interacción promedio se obtiene promediando los valores de todas las configuraciones consideradas:

$$\langle W \rangle = -\boldsymbol{\mu}^0 \cdot \langle \mathbf{E} \rangle - \frac{1}{2} \boldsymbol{\alpha} : \langle \mathbf{E} \otimes \mathbf{E} \rangle \quad (4.21)$$

Cuando se aplica la ACM, sin embargo, la energía de interacción se calcula una sola vez a partir del campo eléctrico promedio:

$$\begin{aligned} W^{\text{ACM}} &= -\boldsymbol{\mu} \cdot \langle \mathbf{E} \rangle + \frac{1}{2} \Delta \boldsymbol{\mu} \cdot \langle \mathbf{E} \rangle = -(\boldsymbol{\mu}^0 + \boldsymbol{\alpha} \langle \mathbf{E} \rangle) \cdot \langle \mathbf{E} \rangle + \frac{1}{2} \boldsymbol{\alpha} \langle \mathbf{E} \rangle \cdot \langle \mathbf{E} \rangle = \\ &= -\boldsymbol{\mu}^0 \cdot \langle \mathbf{E} \rangle - \frac{1}{2} \boldsymbol{\alpha} : (\langle \mathbf{E} \rangle \otimes \langle \mathbf{E} \rangle) \end{aligned} \quad (4.22)$$

El error cometido por esta aproximación se calcula entonces por diferencia entre las expresiones (4.21) y (4.22), obteniéndose una fórmula aproximada que es válida cuando la distribución de carga del soluto está bien descrita por un dipolo eléctrico:

$$\begin{aligned} \Delta W_{\text{Stark}} &= \langle W \rangle - W^{\text{ACM}} = -\frac{1}{2} \boldsymbol{\alpha} : (\langle \mathbf{E} \otimes \mathbf{E} \rangle - \langle \mathbf{E} \rangle \otimes \langle \mathbf{E} \rangle) \\ \Delta W_{\text{Stark}} &= -\frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \alpha_{ij} (\langle E_i E_j \rangle - \langle E_i \rangle \langle E_j \rangle) \end{aligned} \quad (4.23)$$

Según esta expresión,  $\Delta W_{\text{Stark}}$  es proporcional a la polarizabilidad del soluto y a la fluctuación del campo eléctrico generado por el disolvente en la posición ocupada por el soluto. Además, siempre será de signo negativo, de acuerdo con el carácter estabilizador de esta componente.

Para la aplicación de la expresión (4.23) en el método ASEP/MD sólo es necesario obtener, para cada configuración considerada en el cálculo, el valor del campo eléctrico ejercido sobre el soluto, situado en el origen de coordenadas. Puesto que el disolvente se representa por un conjunto de cargas puntuales, el campo eléctrico se obtiene fácilmente por la fórmula:

$$\mathbf{E}(X) = - \sum_i \frac{q_i \mathbf{r}_i}{r_i^3} \quad (4.24)$$

donde  $i$  corresponde a todas las cargas presentes en la configuración del disolvente  $X$ . En todo este desarrollo se supone que la polarizabilidad del soluto se mantiene constante; resulta más adecuado tomar su valor calculado en disolución (con el método ASEP/MD, por ejemplo), pero también puede calcularse en fase gas o emplear datos experimentales, ya que el cálculo teórico de la polarizabilidad puede ser problemático.



Segunda parte

AMPLIACIÓN DEL MÉTODO  
ASEP/MD



*Nadie cree en una hipótesis, excepto quien la propone;  
pero todos creen en un experimento, salvo quien lo realiza*



*William I. B. Beveridge (1908-)*





---

---

# 5

---

## CÁLCULO DE LA ENERGÍA LIBRE

La energía libre es posiblemente la magnitud más importante de la termodinámica, es una medida de la cantidad de trabajo que puede realmente obtenerse de un sistema y juega un papel fundamental en todos los procesos químicos.

Normalmente la energía libre se expresa como la función de Helmholtz,  $A$ , o la función de Gibbs,  $G$ . La energía libre de Helmholtz es apropiada para sistemas en los que se mantienen constantes el número de partículas, la temperatura y el volumen ( $NVT$ ); la energía libre de Gibbs, por su parte, se emplea cuando se mantienen constantes el número de partículas, la temperatura y la presión ( $NPT$ ). La mayoría de los experimentos se llevan a cabo bajo condiciones de presión y temperatura constantes, por lo que la función de Gibbs es la magnitud que más se estudia; sin embargo, las simulaciones a volumen constante son más sencillas que las simulaciones a presión constante, por lo que es más frecuente el cálculo teórico de la energía libre de Helmholtz. En cualquier caso, para sistemas poco compresibles, como la mayor parte de los líquidos, la diferencia entre ambas energías libres no es muy importante y puede considerarse una u otra indistintamente.

Por desgracia, la energía libre es una magnitud difícil de calcular, especialmente en sistemas que tienen muchas configuraciones de energía parecida separadas por barreras energéticas de poca magnitud, como líquidos y macromoléculas flexibles. Otras magnitudes relacionadas, como la entropía y el potencial químico también son difíciles de obtener. La energía libre no puede determinarse de manera precisa a partir de una dinámica molecular o simulación de Monte Carlo corrientes, debido a que estas simulaciones no muestrean de manera adecuada las regiones del espacio de configuraciones que contribuyen de forma importante a la energía libre.

Concretamente, la energía libre de Helmholtz viene dada por:

$$A = -k_B T \ln Q = k_B T \ln \left( \iint e^{\beta E(\mathbf{p}^N, \mathbf{r}^N)} \rho(\mathbf{p}^N, \mathbf{r}^N) d\mathbf{p}^N d\mathbf{r}^N \right) \quad (5.1)$$

donde  $Q = \iint e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N)} d\mathbf{p}^N d\mathbf{r}^N$  es la función de partición del sistema. Según aumenta la energía de las configuraciones,  $\rho(\mathbf{p}^N, \mathbf{r}^N)$  disminuye, pero  $e^{\beta E}$  aumenta, de manera que la contribución a la integral sigue siendo importante incluso para configuraciones de elevada energía. Las simulaciones de dinámica molecular y Monte Carlo, sin embargo, sólo muestrean las regiones de menor energía del espacio de configuraciones y dejan sin explorar estas otras regiones que contribuyen significativamente a la integral (5.1). Por lo tanto, el cálculo de energías libres empleando directamente simulaciones convencionales conduce a valores imprecisos y de mala calidad.

## 5.1. Métodos para el cálculo de la energía libre

Consideremos un problema estrechamente relacionado pero ligeramente diferente: el cálculo de la diferencia de energía libre entre dos estados. Éste es un problema que puede abordarse con métodos que emplean simulaciones de dinámica molecular o de Monte Carlo.

Se han propuesto tres métodos principales para el cálculo de diferencias de energía libre [110, 150-153]: perturbación termodinámica, integración termodinámica y desarrollo lento. Los tres métodos tienen un fundamento similar y suelen dar resultados parecidos; sin embargo, puede haber razones prácticas para elegir uno u otro método.

Una cuestión importante es que la formulación de la energía libre se hace en términos de la función de partición,  $Q$ , y de la energía total,  $E(\mathbf{p}^N, \mathbf{r}^N)$ , que incluyen contribuciones de energía cinética y potencial. Puede considerarse que, en un ciclo termodinámico, las contribuciones de energía cinética son iguales en los dos estados y se anulan al hallar la diferencia, en cuyo caso las distintas ecuaciones se pueden escribir en términos de la integral configuracional,  $Z$ , y la energía potencial, respectivamente.

### 5.1.1. Perturbación termodinámica

Consideremos dos estados bien definidos representados por los sistemas **A** y **B**. **A** contiene  $N$  partículas cuya energía viene dada por  $E_A(\mathbf{p}^N, \mathbf{r}^N)$ . **B** contiene también  $N$  partículas con energía  $E_B(\mathbf{p}^N, \mathbf{r}^N)$ . La diferencia de energía libre,  $\Delta A$ ,

entre los dos estados es:

$$\Delta A = A_B - A_A = -k_B T \ln \frac{Q_B}{Q_A} = -k_B T \ln \left( \frac{\iint e^{-\beta E_B(\mathbf{p}^N, \mathbf{r}^N)} d\mathbf{p}^N d\mathbf{r}^N}{\iint e^{-\beta E_A(\mathbf{p}^N, \mathbf{r}^N)} d\mathbf{p}^N d\mathbf{r}^N} \right) \quad (5.2)$$

Introduciendo  $1 = e^{\beta E_A(\mathbf{p}^N, \mathbf{r}^N)} e^{-\beta E_A(\mathbf{p}^N, \mathbf{r}^N)}$  en el numerador se obtiene:

$$\Delta A = -k_B T \ln \left( \frac{\iint \left( e^{-\beta E_B(\mathbf{p}^N, \mathbf{r}^N)} e^{\beta E_A(\mathbf{p}^N, \mathbf{r}^N)} \right) e^{-\beta E_A(\mathbf{p}^N, \mathbf{r}^N)} d\mathbf{p}^N d\mathbf{r}^N}{\iint e^{-\beta E_A(\mathbf{p}^N, \mathbf{r}^N)} d\mathbf{p}^N d\mathbf{r}^N} \right) \quad (5.3)$$

La ecuación (5.3) puede escribirse en términos de un promedio colectivo:

$$\Delta A = -k_B T \ln \left\langle e^{-\beta[E_B(\mathbf{p}^N, \mathbf{r}^N) - E_A(\mathbf{p}^N, \mathbf{r}^N)]} \right\rangle_A \quad (5.4)$$

La nomenclatura  $\langle \mathcal{X} \rangle_A$  indica que se promedia la magnitud  $\mathcal{X}$  sobre todas las configuraciones representativas del estado **A**. Si se quiere expresar  $\Delta A$  en función del promedio correspondiente al estado **B**, la ecuación queda:

$$\Delta A = k_B T \ln \left\langle e^{-\beta[E_A(\mathbf{p}^N, \mathbf{r}^N) - E_B(\mathbf{p}^N, \mathbf{r}^N)]} \right\rangle_B \quad (5.5)$$

Este método de calcular las diferencias de energía libre se atribuye generalmente a Zwanzig [154]. Para realizar un cálculo de «perturbación termodinámica» o «perturbación de energía libre» (*Free Energy Perturbation*, FEP), debemos definir primero  $E_A$  y  $E_B$ , luego llevar a cabo una simulación en el estado **A** y calculando el promedio de  $e^{-\beta(E_B - E_A)}$  para las configuraciones obtenidas. De igual manera, podríamos efectuar una simulación en el estado **B** y obtener el promedio de  $e^{-\beta(E_A - E_B)}$ .

Si los muestreos de **A** y **B** no solapan en el espacio de configuraciones entonces el valor de la energía libre calculado con la ecuación (5.4) o (5.5) no será muy preciso, porque no se estará muestreando de manera adecuada el espacio de configuraciones de **B** cuando se simula **A** y viceversa. Este problema surge sobre todo cuando la diferencia de energía entre los dos estados es mucho mayor que  $k_B T$ :  $|E_B - E_A| \gg k_B T$ . En estas circunstancias, para obtener una estimación válida de la diferencia de energía libre, se suele introducir una serie de  $n$  estados intermedios,  $i = \{1, 2, \dots, n\}$ , entre **A** y **B** con energía  $E_i(\mathbf{p}^N, \mathbf{r}^N)$  y energía libre  $A_i$ :

$$\begin{aligned} \Delta A &= (A_1 - A_A) + \sum_{i=1}^{n-1} (A_{i+1} - A_i) + (A_B - A_n) = \\ &= -k_B T \ln \left[ \frac{Q_1}{Q_A} \cdot \left( \prod_{i=1}^{n-1} \frac{Q_{i+1}}{Q_i} \right) \cdot \frac{Q_B}{Q_n} \right] = \end{aligned}$$

$$= -k_B T \sum_{i=0}^n \ln \left\langle e^{-\beta[E_{i+1}(\mathbf{p}^N, \mathbf{r}^N) - E_i(\mathbf{p}^N, \mathbf{r}^N)]} \right\rangle_i \quad (5.6)$$

considerando que  $E_A = E_0$  y  $E_B = E_{n+1}$ . La principal característica es que los términos correspondientes a los estados intermedios se cancelan entre sí, por lo que somos libres de elegir tantos estados como sea necesario para conseguir buenos solapamientos entre los espacios de configuraciones y, por lo tanto, valores fiables de las diferencias de energía libre. Además, estos estados intermedios no tienen por qué tener ningún significado físico y suelen describirse por un parámetro de acoplamiento  $\lambda$  ( $0 \leq \lambda \leq 1$ ). El método más sencillo es definir los estados intermedios mediante una interpolación lineal de los estados **A** y **B**:

$$E_i = \lambda_i E_B + (1 - \lambda_i) E_A \quad \lambda_i = \frac{i}{n+1} \quad (5.7)$$

aunque se pueden emplear conexiones más complejas.

De esta manera pueden obtenerse los distintos estados intermedios y luego calcular  $\Delta A$  mediante la ecuación (5.6). El número de estados intermedios se elige de tal manera que cada promedio se calcule sobre diferencias de energía comparables a  $k_B T$ . Para comprobar la calidad del resultado, el cálculo se realiza normalmente en ambas direcciones (**A**  $\rightarrow$  **B** y **B**  $\rightarrow$  **A**), y la diferencia es una medida del error cometido en  $\Delta A$ . No obstante, debe tenerse en cuenta que simulaciones demasiado cortas pueden conducir a valores calculados por perturbación directa e inversa muy similares entre sí, sin que eso signifique que el resultado final es fiable.

### 5.1.2. Integración termodinámica

Una manera alternativa de calcular diferencias de energía libre es la «integración termodinámica». Se supone que la energía del sistema varía con el parámetro  $\lambda$ :  $E = E(\mathbf{p}^N, \mathbf{r}^N, \lambda)$ . La energía libre del sistema,  $A$ , por lo tanto, también depende de  $\lambda$  y, si es una función continua, podemos escribir:

$$\Delta A = \int_0^1 \frac{\partial A(\lambda)}{\partial \lambda} d\lambda = -k_B T \int_0^1 \frac{\partial \ln Q(\lambda)}{\partial \lambda} d\lambda = \int_0^1 \frac{-k_B T}{Q(\lambda)} \frac{\partial Q(\lambda)}{\partial \lambda} d\lambda \quad (5.8)$$

De la definición de  $Q$  se tiene:

$$\begin{aligned} \frac{\partial Q(\lambda)}{\partial \lambda} &= \iint \frac{\partial}{\partial \lambda} \left( e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N, \lambda)} \right) d\mathbf{p}^N d\mathbf{r}^N = \\ &= -\beta \iint \frac{\partial E(\mathbf{p}^N, \mathbf{r}^N, \lambda)}{\partial \lambda} e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N, \lambda)} d\mathbf{p}^N d\mathbf{r}^N \end{aligned} \quad (5.9)$$

Sustituyendo en la ecuación (5.8):

$$\begin{aligned}\Delta A &= \int_0^1 \iint \frac{\partial E(\mathbf{p}^N, \mathbf{r}^N, \lambda)}{\partial \lambda} \frac{e^{-\beta E(\mathbf{p}^N, \mathbf{r}^N, \lambda)}}{Q(\lambda)} d\mathbf{p}^N d\mathbf{r}^N d\lambda = \\ &= \int_0^1 \left\langle \frac{\partial E(\mathbf{p}^N, \mathbf{r}^N, \lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda\end{aligned}\quad (5.10)$$

Para calcular una diferencia de energía libre por integración termodinámica es necesario determinar la integral de la ecuación (5.10). En la práctica, esto se consigue realizando una serie de simulaciones correspondientes a distintos valores de  $\lambda$  entre 0 y 1. Para cada valor de  $\lambda$  se calcula el promedio de  $\partial E(\mathbf{p}^N, \mathbf{r}^N, \lambda)/\partial \lambda$ . Estas derivadas parciales pueden calcularse analíticamente o mediante diferencias finitas ( $\partial E/\partial \lambda = \Delta E/\Delta \lambda$ ). La diferencia de energía libre total,  $\Delta A$ , se calcula entonces por integración numérica:

$$\Delta A \simeq \sum_i^n \left\langle \frac{\partial E(\lambda)}{\partial \lambda} \right\rangle_{\lambda_i} \Delta \lambda_i \quad (5.11)$$

Si la energía está parametrizada como en la ecuación (5.7) resulta bastante sencillo calcular la derivada respecto al parámetro de acoplamiento de manera analítica ( $\partial E/\partial \lambda = E_B - E_A$ ), y el promedio en (5.11) no es más complicado que el de la ecuación (5.4). Además, las mismas simulaciones pueden emplearse para calcular la diferencia de energía libre mediante ambas ecuaciones, lo que permite una medida adicional de la fiabilidad del valor obtenido.

### 5.1.3. Desarrollo lento

Un tercer método para el cálculo de diferencias de energía libre a partir de simulaciones es el método de «desarrollo lento» (*slow growth*). En este método, la forma de la energía del sistema varía de una manera constante y muy lentamente a lo largo de la simulación. Esto quiere decir que en cada etapa la energía  $E_{i+1}$  es prácticamente igual a  $E_i$ . La ecuación (5.6) puede entonces desarrollarse como una serie de Taylor:

$$\begin{aligned}\Delta A &= -k_B T \sum_{i=0}^N \ln \left\langle 1 - \frac{E_{i+1} - E_i}{k_B T} + \dots \right\rangle_i = \\ &= -k_B T \sum_{i=0}^N \ln \left[ 1 - \frac{1}{k_B T} \langle E_{i+1} - E_i \rangle_i + \dots \right] \simeq \\ &\simeq \sum_{i=0}^N \langle E_{i+1} - E_i \rangle\end{aligned}\quad (5.12)$$

En este caso,  $N$  es el número de pasos de la simulación y  $E$  varía de manera que  $E_0 = E_A$  y  $E_{N+1} = E_B$ , por ejemplo según la ecuación (5.7).

## 5.2. Aplicación en el método ASEP/MD

Como uno de los objetivos de este trabajo, nos propusimos desarrollar un método que permitiera observar la variación de la energía libre durante el transcurso del cálculo ASEP/MD.

En un principio, la variación en la energía libre del sistema de una iteración a la siguiente será debida únicamente a la polarización del soluto (y del disolvente, si se trata de un cálculo totalmente acoplado). No obstante, otro de los objetivos planteados consiste en poner a punto un método para realizar la optimización de la geometría del soluto en disolución (tratada en el próximo capítulo, pág. 99), por lo que en la variación de la energía libre habrá también una componente correspondiente a este proceso. En cualquier caso, durante el proceso de polarización ASEP/MD, se espera observar una disminución en la energía libre durante las primeras iteraciones —señal de que el sistema se estabiliza— y una fluctuación en torno a un valor determinado una vez que se ha alcanzado la convergencia, fluctuación debida a que las muestras de configuraciones generadas por las simulaciones son finitas.

Debido a la naturaleza del método ASEP/MD, se consideró el método de perturbación de energía libre (FEP) como el más adecuado, ya que permite, de una manera sencilla, el cálculo de las energías libres tras haber obtenido las configuraciones en equilibrio con los distintos sistemas. Como se ha mencionado, entre dos iteraciones sucesivas del método ASEP/MD tiene lugar una variación en las cargas atómicas del soluto y, posiblemente, en su geometría. El objetivo, por lo tanto, es calcular la diferencia de energía libre entre dos sistemas en los que la geometría del soluto cambia de  $\mathbf{r}_{i-1}$  a  $\mathbf{r}_i$ , y su función de onda —y con ella la distribución electrónica y cargas atómicas— varía en consecuencia.

La expresión empleada para obtener esta diferencia de energía libre, cuya justificación [155] se da en el apéndice C (pág. 197), es la siguiente (sustituyendo ya la energía libre de Helmholtz por la de Gibbs):

$$\begin{aligned} \Delta G_i &= G(\mathbf{r}_i) - G(\mathbf{r}_{i-1}) = \Delta E_{\text{QM}} + k_B T \ln \left\langle e^{\beta[E_{\text{QM/MM}}(\mathbf{r}_i) - E_{\text{QM/MM}}(\mathbf{r}_{i-1})]} \right\rangle_i \quad (5.13) \\ \Delta E_{\text{QM}} &= E_{\text{QM}}(\mathbf{r}_i) - E_{\text{QM}}(\mathbf{r}_{i-1}) = \langle \Psi_i | \hat{H}_i^0 | \Psi_i \rangle - \langle \Psi_{i-1} | \hat{H}_{i-1}^0 | \Psi_{i-1} \rangle \end{aligned}$$

donde  $\Delta E_{\text{QM}}$  corresponde a la diferencia de energía cuántica del soluto, mientras que el segundo sumando corresponde a la diferencia en la interacción soluto-disolvente. Para el cálculo de este segundo sumando, basta con tomar la muestra de

configuraciones en equilibrio con el soluto en el estado  $i$  y, para cada una de ellas calcular la diferencia de  $E_{\text{QM/MM}}$  según se consideren la geometría y distribución de cargas del soluto en el estado  $i$  ( $\mathbf{r}_i$ ) o  $i - 1$  ( $\mathbf{r}_{i-1}$ ).

Este método es válido no sólo para calcular la diferencia de energía entre dos iteraciones consecutivas del proceso ASEP/MD, sino entre dos sistemas cualesquiera. Así, por ejemplo, pueden obtenerse las estructuras (geometrías y distribuciones de carga) de dos confórmeros, **A** y **B** de un mismo soluto y, posteriormente, calcular la diferencia de energía libre entre ellos. Cuando los sistemas son demasiado diferentes, se recurre a la técnica mostrada en la ecuación (5.6), creando estados intermedios, que no tienen necesariamente sentido físico:

$$\begin{aligned} \Delta G_{\text{A} \rightarrow \text{B}} &= E_{\text{QM}}(\mathbf{r}_{\text{B}}) - E_{\text{QM}}(\mathbf{r}_{\text{A}}) + k_{\text{B}}T \sum_{i=1}^{m+1} \ln \left\langle e^{\beta[E_{\text{QM/MM}}(\mathbf{r}_i) - E_{\text{QM/MM}}(\mathbf{r}_{i-1})]} \right\rangle_i = \\ &= E_{\text{QM}}(\mathbf{r}_{\text{B}}) - E_{\text{QM}}(\mathbf{r}_{\text{A}}) - k_{\text{B}}T \sum_{i=0}^m \ln \left\langle e^{-\beta[E_{\text{QM/MM}}(\mathbf{r}_{i+1}) - E_{\text{QM/MM}}(\mathbf{r}_i)]} \right\rangle_i \end{aligned} \quad (5.14)$$

De esta manera, basta con efectuar un cálculo de FEP entre los dos sistemas, considerando únicamente la energía de interacción entre soluto y disolvente, y al resultado sumarle la diferencia en las  $E_{\text{QM}}$ , valores que se obtienen de los cálculos ASEP/MD previos.

Los estados intermedios, como ya se han mencionado, pueden elegirse de la manera que resulte más conveniente, pues la energía libre es función de estado y no depende de la naturaleza de estos estados intermedios; lo más sencillo es realizar una interpolación lineal entre  $\mathbf{r}_{\text{A}}$  y  $\mathbf{r}_{\text{B}}$ . A efectos del cálculo de energía libre, la energía de interacción entre soluto y disolvente,  $E_{\text{QM/MM}}$  en las expresiones (5.13) y (5.14), se calcula clásicamente, representando al soluto mediante su distribución de cargas en disolución.





---

---

# 6

---

## OPTIMIZACIÓN DE GEOMETRÍA

La determinación de estructuras moleculares estables o de transición en disolución es un problema complicado, especialmente por la dificultad de considerar adecuadamente la influencia que tienen la presencia del disolvente y los efectos térmicos. En el caso del efecto del disolvente, es necesario incluir un gran número de moléculas de éste para describir de manera adecuada la perturbación sufrida por el soluto, sobre todo en líquidos polares. Para considerar los efectos térmicos, por otro lado, hay que manejar un gran número de configuraciones del sistema. Estos dos factores implican un coste computacional muy elevado con respecto a los cálculos en fase gaseosa y, además, complican la propia definición del problema. En una situación en la que el entorno de cada molécula de soluto es diferente y está sometido a variaciones temporales, la definición de una única estructura estable puede ser algo poco apropiado.

Existen varias estrategias posibles para aproximarse a lo que puede entenderse como la estructura molecular en disolución. En los métodos QM/MM tradicionales, por ejemplo, los grados de libertad internos del soluto se consideran como variables dinámicas adicionales, sujetas a fluctuaciones térmicas y a interacción con el disolvente. Se realiza entonces una simulación, calculando en cada paso la fuerza total sobre los núcleos del soluto y del disolvente; se obtiene una nueva posición para todos los núcleos por integración de las ecuaciones de movimiento, como en una simulación de dinámica molecular (pág. 63). Promediando sobre las diferentes estructuras del soluto se puede obtener una geometría media. En un segundo conjunto de métodos se define una «superficie de energía libre» sobre la que se localiza la estructura deseada, son conocidos como métodos del «gradiente de la energía libre» (*Free Energy Gradient*, FEG) [156-158]. En este caso se calcula

directamente el gradiente sobre la superficie de energía libre, obtenido como la fuerza promedio de una simulación en la que la estructura de la molécula de soluto es rígida (ver apéndice E, pág. 203). La principal ventaja de este método es que permite obtener tanto geometrías estables como de transición; sin embargo, el tiempo de simulación se incrementa notablemente, ya que es necesario obtener un gran número de configuraciones para poder definir las fuerzas y proseguir con la optimización.

## 6.1. Fundamento

En el marco de la aproximación de Born-Oppenheimer (pág. 33), la energía de los sistemas químicos puede considerarse como una función de las coordenadas que definen el sistema, es decir, las posiciones de los núcleos. Estas posiciones pueden darse en forma de coordenadas cartesianas o bien como «coordenadas internas» que relacionan unos núcleos con otros, como distancias y ángulos de enlace [159]. En cualquier caso, una vez determinado el sistema químico a estudiar —cargas y masas nucleares, número de electrones—, su energía varía con las mencionadas coordenadas nucleares. Para cada estructura o geometría del sistema puede obtenerse una energía y su representación con respecto a las coordenadas nucleares se conoce como «superficie o hipersuperficie de energía potencial» (aunque también incluye la energía cinética de los electrones, formalmente se considera que es la energía potencial de los núcleos).

El problema de la optimización de geometrías moleculares consiste en hallar las estructuras, definidas por las coordenadas nucleares, para las que la energía es mínima, es decir se trata de una búsqueda de mínimos en la superficie de energía potencial. Los mínimos locales en esta superficie corresponden a estructuras estables del sistema, ya que cualquier pequeña deformación de la geometría conduce a un aumento de energía y, por lo tanto, a una desestabilización. Matemáticamente, el problema es la minimización de una función no lineal de múltiples variables, función que generalmente carece de expresión analítica.

No solamente los mínimos en la superficie de energía potencial tienen interés químico, también son objeto de estudio los llamados «puntos de ensilladura» (*saddle points*). La estructura definida por un punto de ensilladura corresponde a un estado de transición, es decir, la estructura de mayor energía dentro del camino de menor energía que une dos mínimos.

La estructura del sistema está determinada por el vector  $n$ -dimensional  $r$ , que define las coordenadas nucleares. La energía potencial es una función de las coordenadas:  $E(r)$ . En general, la función es continua y diferenciable en cualquier punto

$\mathbf{r}$ , por lo que puede calcularse el vector de las primeras derivadas parciales o «gradiente»:

$$\mathbf{g}(\mathbf{r}) = \nabla E(\mathbf{r}) = \begin{pmatrix} \partial E / \partial r_1 \\ \partial E / \partial r_2 \\ \vdots \\ \partial E / \partial r_n \end{pmatrix}_{\mathbf{r}} \quad (6.1)$$

así como la matriz de las segundas derivadas o «hessiana»:

$$\mathbf{G}(\mathbf{r}) = \nabla(\mathbf{g}(\mathbf{r})^T) = \nabla \nabla E(\mathbf{r}) = \begin{pmatrix} \partial^2 E / \partial r_1^2 & \partial^2 E / \partial r_2 \partial r_1 & \cdots & \partial^2 E / \partial r_n \partial r_1 \\ \partial^2 E / \partial r_1 \partial r_2 & \partial^2 E / \partial r_2^2 & \cdots & \partial^2 E / \partial r_n \partial r_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2 E / \partial r_1 \partial r_n & \partial^2 E / \partial r_2 \partial r_n & \cdots & \partial^2 E / \partial r_n^2 \end{pmatrix}_{\mathbf{r}} \quad (6.2)$$

Puesto que el orden de diferenciación no altera el resultado,  $\partial^2 E / \partial r_i \partial r_j = \partial^2 E / \partial r_j \partial r_i$ , la matriz hessiana es simétrica:  $\mathbf{G}^T = \mathbf{G}$ .

Dado un vector unitario  $\mathbf{s}$ , puede definirse una recta como el conjunto de puntos  $\mathbf{r} + \alpha \mathbf{s}$ . En el punto  $\mathbf{r}$ , la pendiente de la superficie de energía potencial a lo largo de dicha recta, en el sentido definido por  $\mathbf{s}$ , es  $\mathbf{s}^T \mathbf{g}(\mathbf{r})$  y su curvatura es  $\mathbf{s}^T \mathbf{G}(\mathbf{r}) \mathbf{s}$ .

Desde el punto de vista químico, los puntos de la superficie que tienen interés son los puntos estacionarios de la superficie. Dado un punto  $\mathbf{r}^*$ , será un punto estacionario si, para cualquier dirección  $\mathbf{s}$ , la pendiente de la superficie en dicho punto es nula:

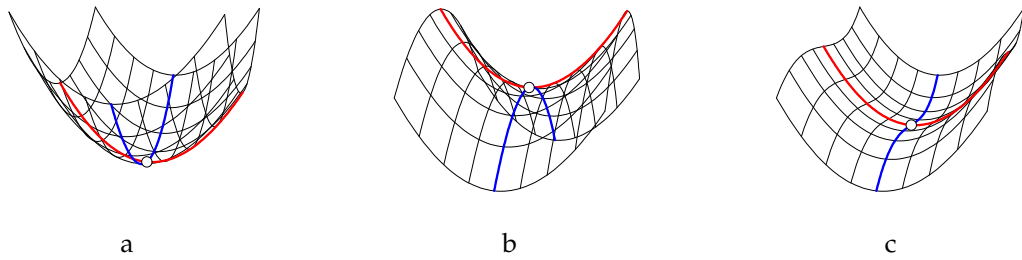
$$\forall \mathbf{s} \quad \mathbf{s}^T \mathbf{g}(\mathbf{r}^*) = 0 \quad \Rightarrow \quad \mathbf{g}^* = \mathbf{g}(\mathbf{r}^*) = 0 \quad (6.3)$$

Además, para que el punto sea un mínimo, la curvatura de la superficie en  $\mathbf{r}^*$  en cualquier dirección debe ser positiva:

$$\forall \mathbf{s} \quad \mathbf{s}^T \mathbf{G}^* \mathbf{s} > 0, \quad \mathbf{G}^* = \mathbf{G}(\mathbf{r}^*) \quad (6.4)$$

Esta última condición equivale a decir que la matriz  $\mathbf{G}^*$  ha de ser definida positiva, o que todos sus valores propios tienen que ser positivos. Si los valores propios de la matriz hessiana son negativos, el punto  $\mathbf{r}^*$  no será un mínimo, sino un máximo local. Si  $m$  valores propios son negativos y el resto positivos, se tratará de un punto de ensilladura de orden  $m$ . Un estado de transición corresponde a un punto de ensilladura de primer orden, caracterizado por una matriz hessiana con un único valor propio negativo.

En la figura 6.1 se representan tres de los tipos de puntos estacionarios que pueden darse en una superficie bidimensional: un mínimo, un punto de ensilladura y un punto de inflexión.



**Figura 6.1:** Tres tipos de puntos estacionarios (marcados con un círculo) en una superficie bidimensional: a) Mínimo. b) Punto de ensilladura. c) Punto de inflexión. En los tres casos, la línea roja marca la dirección con curvatura positiva; la curvatura en la dirección marcada por la línea azul es, respectivamente, positiva, negativa y nula.

## 6.2. Métodos de optimización

El objetivo de los métodos de optimización [152, 160-162], es encontrar el punto —o uno de los puntos—  $\mathbf{r}^*$  para el que se cumple la condición (6.3) y en el que, además, la matriz hessiana,  $\mathbf{G}^*$ , tiene la forma adecuada: todos los valores propios positivos si se quiere localizar una estructura estable, un solo valor propio negativo si el objetivo es encontrar un punto de ensilladura. En lo sucesivo, se supone que se trata del primer caso: la búsqueda de un mínimo de energía. La mayor parte de los métodos de optimización se basan en una búsqueda iterativa del mínimo, es decir, partiendo de un punto inicial,  $\mathbf{r}_0$ , obtienen un nuevo punto,  $\mathbf{r}_1$ , en el que, idealmente, el valor de la energía es menor; este nuevo punto vuelve a emplearse como punto inicial para obtener el siguiente, y así sucesivamente. El proceso termina cuando se cumplen ciertos criterios de convergencia, generalmente relativos a la magnitud de  $\mathbf{g}(\mathbf{r})$ .

En la mayoría de los métodos, cada punto se obtiene a partir del anterior por una fórmula del tipo:

$$\mathbf{r}_{k+1} = \mathbf{r}_k + \alpha_k \mathbf{s}_k \quad (6.5)$$

donde  $\mathbf{s}_k$  es un vector que determina la dirección en la que se da el paso y  $\alpha_k$  es un escalar que determina su longitud. El valor de  $\alpha_k$  puede calcularse mediante una búsqueda lineal: se halla el valor de  $\alpha_k$  que hace  $E(\mathbf{r}_{k+1})$  mínimo, esto requiere normalmente un número relativamente elevado de cálculos de la energía, lo que puede representar un consumo importante de tiempo de cálculo. Otra opción es asignar a  $\alpha_k$  un valor más o menos arbitrario.

### 6.2.1. Métodos de primer orden

En esta categoría se incluyen los métodos que hacen uso del gradiente de la superficie de energía potencial (derivada de primer orden) para encontrar un mínimo.

Uno de los métodos más sencillos es el llamado método de la «pendiente máxima» (*steepest descent*), que consiste en seguir la dirección marcada por el gradiente en cada punto,  $\mathbf{g}_k$ , en sentido descendente, para obtener el siguiente punto:

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \frac{\mathbf{g}_k}{|\mathbf{g}_k|} \quad (6.6)$$

Normalmente este método funciona mejor si el valor de  $\alpha_k$  se obtiene mediante una búsqueda lineal, pero puede simplificarse si se toma  $\alpha_k = |\mathbf{g}_k|$ , ya que entonces  $\mathbf{r}_{k+1} = \mathbf{r}_k - \mathbf{g}_k$ . También pueden tomarse pasos de tamaño arbitrario, que, según progresa la optimización, se van alargando (si  $E(\mathbf{r}_{k+1}) < E(\mathbf{r}_k)$ ) o acortando (si  $E(\mathbf{r}_{k+1}) > E(\mathbf{r}_k)$ ).

Una particularidad de este método es que, cuando el valor de  $\alpha_k$  se halla mediante una búsqueda lineal (de manera que se minimiza  $E(\mathbf{r}_{k+1})$  en cada iteración), los pasos sucesivos son ortogonales entre sí. Esto puede conducir a comportamientos oscilatorios y convergencias muy lentas en determinadas superficies, aunque se trata de un método generalmente robusto en zonas de la superficie alejadas del mínimo.

Para evitar los problemas de comportamiento oscilatorio del método de la pendiente máxima, pueden emplearse los métodos de «gradiente conjugado». En estos métodos la dirección de cada paso se calcula a partir del gradiente y de la dirección en el paso anterior. Tomando la ecuación (6.5) como base, las expresiones de los dos métodos más importantes son:

$$\text{Fletcher-Reeves [163]:} \quad \mathbf{s}_k = -\mathbf{g}_k + \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \mathbf{s}_{k-1} \quad (6.7a)$$

$$\text{Polak-Ribière [164]:} \quad \mathbf{s}_k = -\mathbf{g}_k + \frac{(\mathbf{g}_k - \mathbf{g}_{k-1})^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \mathbf{s}_{k-1} \quad (6.7b)$$

En el caso de funciones cuadráticas, y con búsqueda lineal en cada paso, las fórmulas de Fletcher-Reeves y de Polak-Ribière son equivalentes y se alcanza el mínimo en  $n$  pasos a lo sumo, siendo  $n$  el número de grados de libertad del sistema. Cuando la búsqueda lineal no es exacta o la función no es cuadrática, se necesita un mayor número de pasos. Para funciones no cuadráticas el método de Fletcher-Reeves tiene un mejor comportamiento.

### 6.2.2. Métodos de segundo orden

Los métodos de segundo orden hacen uso de la matriz hessiana (derivada de segundo orden) para localizar los puntos estacionarios en la superficie de energía

potencial. Se basan en la aproximación de la superficie, en el entorno del punto inicial, a una superficie cuadrática, mediante un polinomio de Taylor:

$$\begin{aligned} E(\mathbf{r} + \mathbf{h}) &\simeq E(\mathbf{r}) + \mathbf{h}^T \mathbf{g}(\mathbf{r}) + \frac{1}{2} \mathbf{h}^T \mathbf{G}(\mathbf{r}) \mathbf{h} + \dots \\ \mathbf{g}(\mathbf{r} + \mathbf{h}) &\simeq \mathbf{g}(\mathbf{r}) + \mathbf{G}(\mathbf{r}) \mathbf{h} + \dots \end{aligned} \quad (6.8)$$

La serie se trunca en el término cuadrático, tal como está escrita. En el método de Newton-Raphson se halla el valor de  $\mathbf{h}$  tal que  $\mathbf{g}(\mathbf{r} + \mathbf{h}) = 0$ :

$$\mathbf{g}(\mathbf{r}) + \mathbf{G}(\mathbf{r}) \mathbf{h} = 0 \quad \Rightarrow \quad \mathbf{h} = -\mathbf{G}(\mathbf{r})^{-1} \mathbf{g}(\mathbf{r}) \quad (6.9)$$

Si la función  $E(\mathbf{r})$  es cuadrática, la serie de Taylor (6.8) es una expresión exacta y el punto  $\mathbf{r} + \mathbf{h}$  será un punto estacionario; pero generalmente, la función no es cuadrática y la serie de Taylor es sólo una aproximación, por lo que el punto  $\mathbf{r} + \mathbf{h}$  será también una aproximación al punto estacionario. El método se aplica de manera iterativa y el proceso queda:

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \mathbf{G}_k^{-1} \mathbf{g}_k \quad (6.10)$$

Realmente, el método de Newton-Raphson es un método para hallar las raíces de una función; en el caso de la optimización, se aplica para hallar el punto en el que el gradiente se anula, lo que significa que el punto que se localiza es un punto estacionario, pero no necesariamente un mínimo.

La expresión (6.9) puede escribirse en función de los valores y vectores propios de la hessiana,  $b_i$  y  $\mathbf{v}_i$ , de la siguiente manera:

$$\mathbf{h} = - \sum_{i=1}^n \frac{\mathbf{v}_i^T \mathbf{g}}{b_i} \mathbf{v}_i \quad (6.11)$$

Según esta expresión, el paso de optimización del método de Newton-Raphson está dirigido en contra del gradiente —hacia menores energías— en las direcciones de los modos (vectores propios de la hessiana) correspondientes a los valores propios positivos y a favor de gradiente —hacia mayores energías— en las direcciones de los modos de valor propio negativo. Es decir, tiende a minimizar los modos positivos y a maximizar los negativos. Si la matriz hessiana es definida positiva, éste es el comportamiento deseado para encontrar un mínimo, ya que todos los  $b_i$  serán positivos y la energía se minimizará en todas las direcciones.

En los casos en que el carácter de la hessiana no sea el adecuado, el método de Newton-Raphson no se dirige hacia un mínimo, sino hacia un punto de ensilladura o hacia un máximo. Para evitar este comportamiento, puede modificarse el método,

de manera que:

$$\mathbf{h} = - \sum_{i=1}^n \frac{\mathbf{v}_i^T \mathbf{g} \mathbf{v}_i}{b_i - \lambda} \quad (6.12)$$

en este caso,  $\lambda$  es un parámetro que sirve de desplazamiento para los valores propios de la hessiana,  $b_i$ . Así, dándole a  $\lambda$  un valor menor que cualquier  $b_i$ , puede conseguirse que todos los  $(b_i - \lambda)$  sean positivos, independientemente del signo de los  $b_i$  negativos; de esta manera se puede encontrar un mínimo aunque la hessiana no sea definida positiva. Existen distintos métodos para calcular el valor de  $\lambda$ , uno de ellos, el método RFO, se verá más adelante.

La principal desventaja del método de Newton-Raphson es que necesita la información que proporciona la hessiana, cuyo cálculo no suele ser sencillo y, generalmente, es muy costoso. Además, tal como está planteado, el método requiere efectuar una inversión matricial en cada iteración, operación que no es tampoco sencilla cuando el sistema cuenta con gran número de grados de libertad.

Este problema se puede evitar haciendo uso de los métodos «cuasi-Newton». En estos métodos  $\mathbf{G}^{-1}$  se aproxima mediante una matriz simétrica definida positiva,  $\mathbf{H}$ , que se va actualizando de iteración en iteración. La matriz inicial  $\mathbf{H}_0$  puede ser cualquier matriz simétrica definida positiva, a falta de una mejor estimación puede tomarse la matriz unidad  $\mathbf{I}$ . Si es posible, resulta más adecuado tomar  $\mathbf{H}_0 = \mathbf{G}_0^{-1}$ , efectuando el cálculo de la hessiana sólo una vez, para lo que pueden introducirse también algunas aproximaciones adicionales en caso de ser necesarias (por ejemplo, puede emplearse un método computacionalmente menos exigente para este cálculo inicial).

La fórmula general de los métodos cuasi-Newton es:

$$\begin{aligned} \mathbf{r}_{k+1} &= \mathbf{r}_k - \alpha \mathbf{H}_k \mathbf{g}_k \\ \mathbf{H}_{k+1} &= h(\mathbf{H}_k, \mathbf{r}_k, \mathbf{r}_{k+1}, \mathbf{g}_k, \mathbf{g}_{k+1}) \end{aligned} \quad (6.13)$$

donde  $\alpha$  puede calcularse por búsqueda lineal en cada iteración o tomar un valor arbitrario como, por ejemplo, 1. Una vez obtenido el nuevo punto  $\mathbf{r}_{k+1}$  y calculado  $\mathbf{g}_{k+1}$ , con la información de  $\mathbf{r}_k$ ,  $\mathbf{r}_{k+1}$ ,  $\mathbf{g}_k$  y  $\mathbf{g}_{k+1}$  se actualiza  $\mathbf{H}_k$  para obtener  $\mathbf{H}_{k+1}$ . Para la actualización de la matriz  $\mathbf{H}$  existen diversas fórmulas, cada una más adecuada en distintos casos, quizá las dos más empleadas son la DFP [165-167] y la BFGS [168-172], en el Apéndice D (pág. 201) se muestran estas fórmulas.

#### – El método RFO

Como se ha mencionado, una de las opciones para el cálculo de  $\lambda$  en la ecuación (6.12) es el método de RFO (*Rational Function Optimisation*) [173-175], que se

basa en la aproximación de la superficie de energía potencial,  $E(\mathbf{r})$ , mediante una función racional en lugar de una serie de Taylor, de la siguiente forma:

$$\begin{aligned} E(\mathbf{r} + \mathbf{h}) - E(\mathbf{r}) &\simeq \frac{\mathbf{h}^T \mathbf{g}(\mathbf{r}) + \frac{1}{2} \mathbf{h}^T \mathbf{G}(\mathbf{r}) \mathbf{h}}{1 + \mathbf{h}^T \mathbf{S} \mathbf{h}} = \\ &= \frac{\frac{1}{2} \begin{pmatrix} \mathbf{h}^T & 1 \end{pmatrix} \begin{pmatrix} \mathbf{G}(\mathbf{r}) & \mathbf{g}(\mathbf{r}) \\ \mathbf{g}^T(\mathbf{r}) & 0 \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ 1 \end{pmatrix}}{\begin{pmatrix} \mathbf{h}^T & 1 \end{pmatrix} \begin{pmatrix} \mathbf{S} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ 1 \end{pmatrix}} \end{aligned} \quad (6.14)$$

donde  $\mathbf{S}$  es una matriz simétrica que suele tomarse como la matriz unidad. Si se aplica la condición de punto estacionario,  $\mathbf{g}(\mathbf{r} + \mathbf{h}) = 0$ , se obtiene una ecuación de valores propios de orden  $n + 1$ :

$$\begin{pmatrix} \mathbf{G}(\mathbf{r}) & \mathbf{g}(\mathbf{r}) \\ \mathbf{g}^T(\mathbf{r}) & 0 \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{S} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ 1 \end{pmatrix} \quad (6.15)$$

La matriz que aparece en el término de la izquierda se conoce como matriz hessiana ampliada. Esta ecuación también puede escribirse como:

$$(\mathbf{G}(\mathbf{r}) - \lambda \mathbf{S}) \mathbf{h} + \mathbf{g}(\mathbf{r}) = 0 \quad (6.16a)$$

$$\mathbf{h}^T \mathbf{g}(\mathbf{r}) = \lambda \quad (6.16b)$$

que, tomando  $\mathbf{S} = \mathbf{I}$ , puede expresarse en función de los valores y vectores propios de la hessiana como:

$$\mathbf{h} = \sum_{i=1}^n \frac{\mathbf{v}_i^T \mathbf{g} \mathbf{v}_i}{\lambda - b_i} \quad (6.17)$$

De esta manera no sólo se obtiene de nuevo la ecuación (6.12), sino que se cuenta además con un método para calcular el valor de  $\lambda$ , ya que, según la ecuación (6.15),  $\lambda$  ha de ser uno de los valores propios de la matriz hessiana ampliada. Para calcular estos valores propios a partir de los correspondientes a la hessiana original puede sustituirse (6.17) en (6.16b) y se obtiene:

$$\sum_{i=1}^n \frac{(\mathbf{v}_i^T \mathbf{g})^2}{\lambda - b_i} = \sum_{i=1}^n \frac{d_i^2}{\lambda - b_i} = \lambda \quad (6.18)$$

donde  $d_i = \mathbf{v}_i^T \mathbf{g}$  es la proyección del gradiente sobre el vector  $\mathbf{v}_i$ .

Puesto que la matriz hessiana ampliada tiene un orden  $n + 1$ , los posibles valores de  $\lambda$  son también  $n + 1$ , y tienen la propiedad de que acotan los  $n$  valores propios de la hessiana:  $\lambda_i \leq b_i \leq \lambda_{i+1}$ . En un punto estacionario, los valores de  $\lambda$  coinciden con los  $b_i$  y además existe un  $\lambda_k = 0$ . Así pues, tomando  $\lambda = \lambda_1$  (el



menor de los valores propios) en la ecuación (6.17), se minimiza la energía en las direcciones de todos los modos propios de la hessiana. La propiedad anteriormente mencionada hace que sea igual de sencillo localizar estados de transición, basta con tomar  $\lambda = \lambda_2$ , se maximizará la energía en la dirección dada por el modo de menor valor propio y se minimizará en todas las demás direcciones.

Una variante de este método es el P-RFO (*Partitioned RFO*) [174]. En este caso los distintos modos locales (vectores propios de la hessiana) se dividen en dos grupos, la energía se maximiza en las direcciones de uno de los grupos, designadas con  $\{\alpha, \beta, \dots\}$  y se minimiza en las otras, designadas con  $\{a, b, \dots\}$  Se construyen entonces dos matrices:

$$\begin{pmatrix} b_\alpha & 0 & \cdots & d_\alpha \\ 0 & b_\beta & \cdots & d_\beta \\ \vdots & \vdots & \ddots & \vdots \\ d_\alpha & d_\beta & \cdots & 0 \end{pmatrix} \quad \begin{pmatrix} b_a & 0 & \cdots & d_a \\ 0 & b_b & \cdots & d_b \\ \vdots & \vdots & \ddots & \vdots \\ d_a & d_b & \cdots & 0 \end{pmatrix} \quad (6.19)$$

y se obtienen dos valores de  $\lambda$ :  $\lambda_P$  como el mayor de los valores propios de la matriz de la izquierda y  $\lambda_N$  como el menor de los valores propios de la matriz de la derecha; o, de manera equivalente, como la mayor y la menor, respectivamente, de las soluciones de las ecuaciones:

$$\sum_{\gamma} \frac{(d_\gamma)^2}{\lambda_P - b_\gamma} = \lambda_P \quad \sum_j \frac{(d_j)^2}{\lambda_N - b_j} = \lambda_N \quad (6.20)$$

donde  $\gamma = \{\alpha, \beta, \dots\}$  y  $j = \{a, b, \dots\}$ . El paso del método P-RFO se calcula entonces como:

$$\mathbf{h} = \sum_{\gamma} \frac{d_\gamma \mathbf{v}_\gamma}{\lambda_P - b_\gamma} + \sum_j \frac{d_j \mathbf{v}_j}{\lambda_N - b_j} \quad (6.21)$$

Cuando el objetivo es localizar un estado de transición, existe un único modo en el grupo  $\gamma$  y entonces  $\lambda_P$  se puede calcular como:

$$\lambda_P = \frac{1}{2}b_\gamma + \frac{1}{2}\sqrt{b_\gamma^2 + 4d_\gamma^2} \quad (6.22)$$

mientras que  $\lambda_N$  se calcula como se indica en la ecuación (6.20). La elección del modo  $\gamma$  a lo largo del cual se maximizará la energía no es trivial, pero si la estructura inicial de la molécula está cercana al punto de ensilladura que se quiere localizar suele ser seguro tomar el modo de menor valor propio ( $\gamma = 1$ ).

Al igual que en otros métodos, el proceso es iterativo:  $\mathbf{r}_{k+1} = \mathbf{r}_k + \mathbf{h}_k$ . En cada iteración puede calcularse la hessiana,  $\mathbf{G}_k$  y, a partir de ella, obtener los distintos  $b_i$  y  $v_i$  o, como en los métodos cuasi-Newton, actualizar la matriz  $\mathbf{B}_k$ , que sirve de aproximación para la hessiana, y obtener sus valores y vectores propios.

### 6.3. Aplicación en el método ASEP/MD

Dado que ASEP/MD es un método que hace uso de la aproximación del campo medio, la manera más natural de introducir la optimización de geometrías es mediante el método FEG, es decir, localizando los puntos estacionarios sobre la superficie de energía libre del sistema,  $G(\mathbf{r})$ .

Una vez que, como se expone en el apéndice C (pág. 197), se define la superficie de energía libre del sistema en la forma:

$$G(\mathbf{r}) = E_{\text{QM}}(\mathbf{r}) + G_{\text{QM/MM}}(\mathbf{r}) \quad (6.23)$$

el gradiente y la hessiana sobre dicha superficie pueden calcularse por las siguientes expresiones (conviene no confundir  $G$  y  $\mathcal{G}$ ):

$$\mathbf{g}(\mathbf{r}) = \nabla G(\mathbf{r}) = \frac{\partial E_{\text{QM}}(\mathbf{r})}{\partial \mathbf{r}} + \frac{\partial G_{\text{QM/MM}}(\mathbf{r})}{\partial \mathbf{r}} \quad (6.24a)$$

$$\mathcal{G}(\mathbf{r}) = \nabla \nabla G(\mathbf{r}) = \frac{\partial^2 E_{\text{QM}}(\mathbf{r})}{\partial \mathbf{r}^2} + \frac{\partial^2 G_{\text{QM/MM}}(\mathbf{r})}{\partial \mathbf{r}^2} \quad (6.24b)$$

Por su parte, las derivadas de  $G_{\text{QM/MM}}$  pueden calcularse como se indica en el apéndice E (pág. 203), obteniéndose:

$$\frac{\partial G_{\text{QM/MM}}(\mathbf{r})}{\partial \mathbf{r}} = \left\langle \frac{\partial E_{\text{QM/MM}}(\mathbf{r})}{\partial \mathbf{r}} \right\rangle = \left\langle \frac{\partial E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r})}{\partial \mathbf{r}} \right\rangle + \left\langle \frac{\partial E_{\text{QM/MM}}^{\text{vdw}}(\mathbf{r})}{\partial \mathbf{r}} \right\rangle \quad (6.25a)$$

$$\frac{\partial^2 G_{\text{QM/MM}}(\mathbf{r})}{\partial \mathbf{r}^2} \simeq \left\langle \frac{\partial^2 E_{\text{QM/MM}}(\mathbf{r})}{\partial \mathbf{r}^2} \right\rangle = \left\langle \frac{\partial^2 E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r})}{\partial \mathbf{r}^2} \right\rangle + \left\langle \frac{\partial^2 E_{\text{QM/MM}}^{\text{vdw}}(\mathbf{r})}{\partial \mathbf{r}^2} \right\rangle \quad (6.25b)$$

En este punto hay que considerar cómo se obtienen  $E_{\text{QM/MM}}^{\text{elec}}$  y  $E_{\text{QM/MM}}^{\text{vdw}}$  en el método ASEP/MD. La interacción de Van der Waals,  $E_{\text{QM/MM}}^{\text{vdw}}$ , se obtiene de manera clásica para cada configuración del sistema, generalmente por potenciales Lennard-Jones; puede obtenerse también una expresión analítica para sus derivadas y hallar el promedio de éstas para todas las configuraciones, con lo cual el último sumando no presenta ningún problema. La interacción electrostática,  $E_{\text{QM/MM}}^{\text{elec}}$ , sin embargo, se calcula cuánticamente una sola vez y corresponde a la interacción promedio, por lo que puede representarse como  $\langle E_{\text{QM/MM}}^{\text{elec}} \rangle$ .

Por consiguiente, se establece una nueva aproximación:

$$\left\langle \frac{\partial E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r})}{\partial \mathbf{r}} \right\rangle \simeq \frac{\partial \langle E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r}) \rangle}{\partial \mathbf{r}} \quad \left\langle \frac{\partial^2 E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r})}{\partial \mathbf{r}^2} \right\rangle \simeq \frac{\partial^2 \langle E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r}) \rangle}{\partial \mathbf{r}^2} \quad (6.26)$$

Finalmente, cuando se realiza un cálculo cuántico se obtiene conjuntamente

$E_{\text{QM}} + E_{\text{QM/MM}}^{\text{elec}}$ , por lo que el gradiente y la hessiana proporcionados incluyen ambas contribuciones, es decir:

$$\mathbf{g}_{\text{calc}}(\mathbf{r}) = \frac{\partial (E_{\text{QM}}(\mathbf{r}) + \langle E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r}) \rangle)}{\partial \mathbf{r}} = \frac{\partial E_{\text{QM}}(\mathbf{r})}{\partial \mathbf{r}} + \frac{\partial \langle E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r}) \rangle}{\partial \mathbf{r}} \quad (6.27a)$$

$$\mathbf{G}_{\text{calc}}(\mathbf{r}) = \frac{\partial^2 (E_{\text{QM}}(\mathbf{r}) + \langle E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r}) \rangle)}{\partial \mathbf{r}^2} = \frac{\partial^2 E_{\text{QM}}(\mathbf{r})}{\partial \mathbf{r}^2} + \frac{\partial^2 \langle E_{\text{QM/MM}}^{\text{elec}}(\mathbf{r}) \rangle}{\partial \mathbf{r}^2} \quad (6.27b)$$

Con esto se obtiene la expresión final para las derivadas sobre la superficie de energía libre:

$$\mathbf{g}(\mathbf{r}) \simeq \mathbf{g}_{\text{calc}}(\mathbf{r}) + \langle \nabla E_{\text{QM/MM}}^{\text{vdw}}(\mathbf{r}) \rangle \quad (6.28a)$$

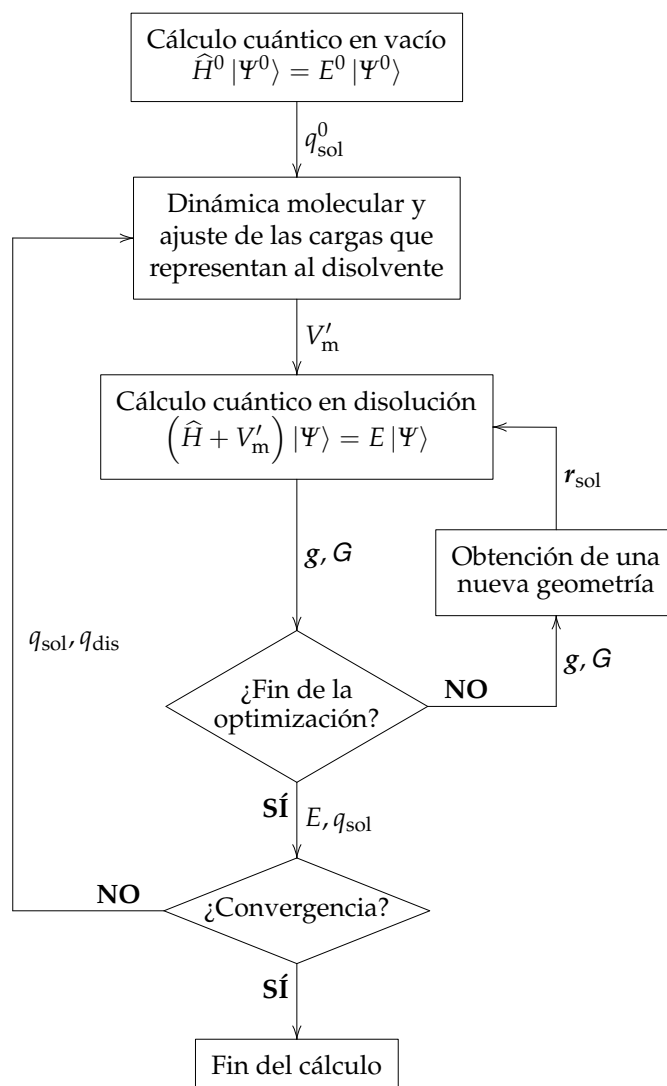
$$\mathbf{G}(\mathbf{r}) \simeq \mathbf{G}_{\text{calc}}(\mathbf{r}) + \langle \nabla \nabla E_{\text{QM/MM}}^{\text{vdw}}(\mathbf{r}) \rangle \quad (6.28b)$$

esto es, sólo es necesario sumar la contribución de Van der Waals al resultado del cálculo cuántico con el disolvente promedio.

Con el gradiente y la hessiana de la energía libre, puede aplicarse cualquiera de los métodos de optimización mencionados de dos maneras diferentes. Una posibilidad consiste en hallar, en cada iteración del proceso ASEP/MD, la geometría óptima para el soluto: en este caso se realiza una optimización completa en cada ciclo. La otra opción es efectuar solamente un número reducido de pasos de optimización en cada iteración ASEP/MD, de manera que el proceso de optimización de geometría se intercala con el de polarización del soluto y de la estructura del disolvente.

En el primer caso se requiere un número mayor de cálculos cuánticos por cada iteración ASEP/MD, pero es previsible que la convergencia final se consiga en un menor número de iteraciones. En el segundo caso, al limitarse la variación de geometría entre cada dos iteraciones, puede facilitarse la convergencia en sistemas que sufran un cambio muy importante en disolución.

En la figura 6.2 se muestra cómo se introduce la optimización de geometría en el método ASEP/MD acoplado.



**Figura 6.2:** Esquema del método ASEP/MD con optimización. El fin de la optimización se produce cuando el gradiente,  $g$ , es menor que cierto valor o cuando las iteraciones llegan a un número determinado.

---

---

**Tercera parte**

**APLICACIONES**

---

---



*1729 es un número muy interesante, es el menor número  
que puede expresarse como suma de dos cubos de dos maneras diferentes*



*Srinivasa Ramanujan (1887-1920)*



---

---

# 7

---

## CONDICIONES GENERALES

Los métodos desarrollados se han aplicado a algunos sistemas de interés. En los próximos capítulos se detallan las condiciones particulares en cada caso y los resultados obtenidos. En éste se presentan las condiciones más generales, como son los programas empleados, los parámetros de Lennard-Jones, etc.

Para todos los cálculos cuánticos se utilizó el programa *Gaussian 98* [176], salvo para las simulaciones DFT/MM del triazeno, que se detallarán en su momento. Se aplicaron los siguientes métodos:

- Hartree-Fock [44-46].
- Teoría de perturbaciones Møller-Plesset de segundo orden, MP2 [52].
- Método del espacio activo completo, CASSCF [56].
- Teoría del funcional de la densidad, DFT [64], con funcionales BP86 (Becke-Perdew) [71, 177], B3LYP [72, 73] y BH&HLYP (*Becke's Half & Half - LYP*) [72, 178].

En estos cálculos se emplearon los siguientes conjuntos de base:

- La base «triple zeta de valencia» con funciones de polarización y difusas de Pople, 6-311++G(2d,2p) [179, 180].
- La base «doble zeta de valencia» coherente con la correlación y con funciones de polarización y difusas de Dunning, aug-cc-pVDZ [181-183].

- Una base contraída de «orbitales atómicos naturales» (*Atomic Natural Orbitals*, ANO) [184], formada a partir de conjuntos primitivos 14s9p4d/8s contraídos a 4s3p1d/2s.
- Una base «triple zeta de valencia» con funciones de polarización de tipo NCO(7111/411/1) H(41/1) optimizada para los cálculos DFT del programa deMon [185, 186].

Las simulaciones de dinámica molecular clásica se efectuaron con el programa *Moldy* [187], que se caracteriza por el uso de moléculas rígidas empleando la técnica de los cuaterniones (ver pág. 66). A los distintos sistemas se les aplicó la condición de límite periódico y las interacciones electrostáticas se calculan por el método de Ewald [15, 109, 110]. Las simulaciones se realizaron a una temperatura constante de 298 K, por medio del termostato de Nosé-Hoover [118, 119], el paso temporal empleado fue de 0,5 fs y, salvo cuando se indica lo contrario, cada simulación constó de 50 000 pasos de equilibrado y 100 000 pasos (50 ps) de producción de datos, de estos 50 ps se consideraron 1000 configuraciones para los análisis posteriores.

Para las interacciones entre las distintas moléculas se utilizó una combinación de cargas atómicas centradas en los núcleos y potenciales de Lennard-Jones. Las moléculas de disolvente, representadas mediante MM, se tomaron con las cargas atómicas y geometrías que se dan en la tabla 7.1. Las mismas geometrías son válidas para el soluto cuando se representó con geometría fija (sin optimización), en el caso de los compuestos carbonílicos, se emplearon las geometrías experimentales, que se dan en la tabla 7.2. Los parámetros de Lennard-Jones de todas las moléculas estudiadas se dan en la tabla 7.3.

La interacción de Lennard-Jones viene dada por un potencial del tipo:

$$E^{\text{LJ}}(r_{ij}) = 4\varepsilon_{ij} \left( \frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) \quad (7.1)$$

donde  $\{i, j\}$  designan a dos tipos de átomos. Los parámetros  $\varepsilon_{ij}$  y  $\sigma_{ij}$  se obtienen a partir de los correspondientes a los átomos individuales por media geométrica y aritmética, respectivamente:

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad \sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \quad (7.2)$$

Finalmente, el método ASEP/MD se aplicó por medio de un programa informático —del mismo nombre— desarrollado al efecto, y cuyo código se proporciona en el apéndice F (pág. 207). Para el cálculo de las cargas atómicas del soluto (necesarias en las dinámicas moleculares) se utilizaron los métodos CHELP [139] y CHELPG [140].



**Tabla 7.1:** Geometrías y cargas atómicas de las moléculas de disolvente. Las denominaciones *trans* y *gauche* lo son con respecto al enlace O–H.

Agua (TIP3P) [188]					
$r$ (O–H)	0,9572 Å	$q$ (O)	–0,834		
$\theta$ (H–O–H)	104,52°	$q$ (H)	+0,457		

Alcoholes [105]			Metanol	Etanol	Propanol
$r$ (O–H)	0,945 Å	$q$ (O)	–0,84	–0,76	–0,80
$r$ (C–O)	1,430 Å	$q$ (C <sup>1</sup> )	+0,54	+0,38	+0,57
$r$ (C–C)	1,530 Å	$q$ (C <sup>2</sup> )	—	–0,20	+0,02
$r$ (C–H)	1,090 Å	$q$ (C <sup>3</sup> )	—	—	+0,21
$\theta$ (C–O–H)	108,5°	$q$ (H [O])	+0,50	+0,46	+0,45
$\theta$ (O–C–H)	109,5°	$q$ (H [C <sup>1</sup> , <i>trans</i> ])	–0,04	—	—
$\theta$ (C–C–O)	108,0°	$q$ (H [C <sup>1</sup> , <i>gauche</i> ])	–0,08	–0,02	–0,10
$\theta$ (C–C–C)	112,0°	$q$ (H [C <sup>2</sup> ])	—	+0,05	–0,03
$\theta$ (H–C–H)	109,5°	$q$ (H [C <sup>3</sup> ])	—	—	–0,06

**Tabla 7.2:** Geometrías experimentales de los compuestos carbonílicos.

	Formaldehído [189]	Acetaldehído [190]	Acetona [191]
$r$ (C=O)	1,208 Å	1,216 Å	1,222 Å
$r$ (C–CO)	—	1,501 Å	1,507 Å
$r$ (H–CO)	1,116 Å	1,114 Å	—
$r$ (C–H)	—	1,086 Å	1,085 Å
$\theta$ (C–C=O)	—	123,90°	121,40°
$\theta$ (H–C=O)	121,75°	118,60°	—
$\theta$ (C–C–H)	—	110,60°	110,83°
$\theta$ (H–C–H)	116,50°	108,32°	108,78°

**Tabla 7.3:** Parámetros de Lennard-Jones utilizados en los cálculos.

	$\epsilon_i$ (kcal/mol)	$\sigma_i$ (Å)		$\epsilon_i$ (kcal/mol)	$\sigma_i$ (Å)
Agua (TIP3P) [188]			Xilopiranososa [195]		
O	0,15207	3,15066	O	0,15923	3,200
H	0,00000	0,00000	C	0,09032	3,600
Alcoholes [104, 105]			H (OH)	0,04981	1,600
O	0,0105	3,07	H (CH <sub>n</sub> )	0,00450	2,935
C	0,0800	3,50	Amoniaco [196]		
H (OH)	0,0150	1,35	N	0,27824	3,40
H (CH <sub>n</sub> )	0,0220	2,55	H	0,00000	0,00
Compuestos carbonílicos [192, 193]			Clorometano [197]		
O	0,20	2,850	Cl	0,300	3,395
C (CO)	0,12	3,296	C	0,066	3,490
C (CH <sub>3</sub> )	0,06	3,207	H	0,030	2,462
H	0,01	2,744	Amoniaco-clorometano [198]		
Formamida y triazeno [106, 194]			Cl	0,1119	4,1215
O	0,210	2,96	N	0,1615	3,0870
N	0,170	3,25	C	0,1000	3,3950
C	0,115	3,80	H(CH <sub>3</sub> )	0,0700	1,8700
H	0,000	0,00	H(NH <sub>3</sub> )	0,0000	0,0000

---

---

# 8

---

## VALIDACIÓN DE LA APROXIMACIÓN DEL CAMPO MEDIO

Muchos de los métodos empleados en el estudio de los efectos del disolvente hacen uso de la aproximación del campo medio. Las distintas versiones cuánticas de los modelos del continuo (SCRF, PCM, expansiones multipolares, etc.) o el método de los dipolos de Langevin son algunos ejemplos representativos.

En la última década se han propuesto nuevos métodos que combinan la ACM con una descripción detallada de la estructura del disolvente, como el método RISM-SCF o el ASEP/MD. En cualquier caso, el empleo de la ACM simplifica en gran medida el tratamiento, ya que permite reducir drásticamente el número de cálculos cuánticos necesarios.

Uno de los objetivos planteados en el presente trabajo fue comprobar la validez de la aproximación del campo medio, concretamente en el marco del método ASEP/MD. Con tal fin se llevó a cabo un estudio de los errores introducidos por dicha aproximación en las energías de interacción soluto-disolvente, momentos dipolares y derivadas de la energía libre. Para ello se realizaron dos tipos de cálculos: en un caso se aplicó la ACM a través del método ASEP/MD, en otro se efectuaron tantos cálculos cuánticos como configuraciones consideradas, de manera similar a los métodos QM/MM.

### 8.1. Detalles de los cálculos

Se estudiaron dos tipos de sistemas: líquidos puros y disoluciones. Los líquidos puros fueron metanol, etanol y propanol. Los cálculos cuánticos de estos alcoholes

se realizaron con el método Hartree-Fock y, cuando se indica, el MP2, con el conjunto de base aug-cc-pVDZ. En las simulaciones de dinámica molecular se incluyeron 128 moléculas, se truncaron las interacciones intermoleculares a 9,0 Å, 11,5 Å y 12,5 Å para metanol, etanol y propanol, respectivamente.

Como disoluciones, se consideraron formaldehído, acetaldehído y acetona en agua. Para los cálculos cuánticos se empleó un método CASSCF(6,4) con una base ANO, tomando como orbitales de partida los resultantes de un cálculo HF. En las dinámicas moleculares se incluyó una molécula de soluto y 214 de agua, las interacciones se truncaron a 9,0 Å.

Adicionalmente, se consideró la formamida en disolución acuosa. Se trató mediante DFT, con el funcional BP86 y una base NCO(7111/411/1) H(41/1). También se incluyeron 214 moléculas de disolvente en las simulaciones y las interacciones se truncaron a 9,0 Å. En este sistema únicamente se estudió el efecto de la ACM sobre el gradiente de la energía libre. Se realizó además una dinámica más larga, de 300 ps, con la que poder estudiar las fluctuaciones del gradiente promedio.

Para cada uno de estos sistemas se efectuó un cálculo ASEP/MD, completamente acoplado en el caso de los líquidos puros y parcialmente acoplado en el de las disoluciones; en todos los casos se empleó el método CHELP para el cálculo de las cargas atómicas.

En los resultados obtenidos estarán presentes los errores introducidos por la ACM. Para obtener las distintas propiedades cuando no se aplica la ACM, se escogieron 100 configuraciones (1000 para el metanol) de la última simulación de dinámica molecular realizada en cada caso —las mismas configuraciones a partir de las cuales se calculó  $V_m$  en el método ASEP/MD— y se efectuó un cálculo cuántico para cada una de ellas; el promedio de los valores obtenidos se compara con el resultado del método ASEP/MD.

## 8.2. Resultados y discusión

En primer lugar, se comprobó que los métodos empleados en los cálculos cuánticos proporcionan una descripción adecuada de los distintos solutos. Como se ha mostrado en otros trabajos [41, 199], la adecuada representación de las distribuciones electrónicas en vacío tiene una importante influencia en las propiedades calculadas en disolución. En la tabla 8.1 se dan los momentos dipolares y las polarizabilidades obtenidas en los cálculos para moléculas en fase gas, así como los correspondientes valores experimentales. Los momentos dipolares se reproducen bien, aunque en el caso de los alcoholes están sobrestimados en alrededor de 0,1 D. Las polarizabilidades están subestimadas en unos  $3 a_0^3$  en los alcoholes y algo más en los compuestos carbonílicos.

**Tabla 8.1:** Momentos dipolares ( $\mu^0$ ) y polarizabilidades ( $\alpha^0$ ) en vacío de las moléculas estudiadas. Momentos dipolares en D, polarizabilidades en  $a_0^3$ .

	$\mu^0$	$\alpha^0 = 1/3 \text{ tr } \alpha^0$	$\mu^0$ (exp.)	$\alpha^0$ (exp.) [200]
Metanol	1,81	19,52	1,71 [201]	22,30
Etanol	1,72	31,00	1,68 [202]	34,48
Propanol	1,65	42,12	1,55 [203]	45,49 [204]
Formaldehído	2,35	14,40	2,36 [205]	16,53
Acetaldehído	2,76	25,38	2,75 [206]	31,04
Acetona	2,90	35,58	2,93 [207]	42,92

### 8.2.1. Influencia del número de configuraciones

Un aspecto a tener en cuenta es la importancia que puede tener en la calidad de los resultados el número de configuraciones del disolvente consideradas en los promedios. Este efecto se estudió en el caso del metanol. Una vez realizados los 1000 cálculos cuánticos, uno para cada configuración, se hallaron los promedios de la energía de interacción electrostática entre soluto y disolvente y del momento dipolar. Los promedios se calcularon variando el número de configuraciones,  $N$ , pero siempre abarcando el intervalo completo de simulación de 50 ps; los resultados se muestran en la tabla 8.2-A. También se calcularon los promedios parciales de distintos subintervalos de 5 ps, tomando 100 configuraciones en cada uno de ellos; estos resultados se muestran en la tabla 8.2-B.

Tomando los valores obtenidos con las 1000 configuraciones como referencia, se observa que cuando se incluyen sólo 50 configuraciones en el promedio el error cometido es de 0,27 kcal/mol en la energía y de 0,1 D en el momento dipolar. Con 100 configuraciones los errores disminuyen a 0,06 kcal/mol y 0,00 D, respectivamente. Con 200 configuraciones los resultados son prácticamente indistinguibles del cálculo de referencia. Es evidente que, si las configuraciones elegidas no están correlacionadas —como es el caso—, los resultados deben converger rápidamente y es suficiente un número muy pequeño de configuraciones para obtener una descripción adecuada de las propiedades promedio del sistema. La misma conclusión obtuvieron Coutinho y col. [208, 209] a partir de un estudio de los desplazamientos solvatocrómicos de la primera transición electrónica del benceno en distintos disolventes. A causa de la pequeña diferencia entre los promedios de 100 y 1000 configuraciones, y con el fin de ahorrar tiempo de cálculo, en el resto de los sistemas se hicieron cálculos cuánticos solamente para 100 configuraciones.

La variación de los resultados es mayor cuando se toman diferentes intervalos temporales para calcular los promedios. Como se puede observar en la tabla 8.2-B, se encuentran variaciones de unas 3 kcal/mol en  $\langle E_{QM/MM}^{elec} \rangle$  al calcular el promedio

**Tabla 8.2:** Promedios de la energía de interacción electrostática (en kcal/mol) y del momento dipolar (en D) de una molécula de metanol en estado líquido. Resultados obtenidos a nivel Hartree-Fock.

A			B		
$N$	$\langle E_{QM/MM}^{elec} \rangle$	$\langle \mu \rangle$	Intervalo (ps)	$\langle E_{QM/MM}^{elec} \rangle$	$\langle \mu \rangle$
50	-19,28	2,47	0:5	-18,95	2,41
100	-19,07	2,46	5:10	-18,78	2,44
200	-19,02	2,46	10:15	-17,44	2,39
250	-19,02	2,46	15:20	-17,86	2,46
334	-19,02	2,46	20:25	-18,73	2,48
500	-19,01	2,46	25:30	-19,26	2,50
1000	-19,01	2,46	30:35	-20,79	2,56
			35:40	-20,13	2,49
			40:45	-20,64	2,52
			45:50	-17,59	2,38

- A** En función del número de configuraciones,  $N$ , incluidas en el cálculo del promedio.
- B** En función del intervalo temporal escogido para calcular el promedio (100 configuraciones en cada intervalo).

en diferentes intervalos de 5 ps, en el momento dipolar la variación es de aproximadamente 0,2 D. De estos resultados se concluye que, para obtener valores promedio representativos, más importante que incluir muchas configuraciones en el cálculo es elegir las configuraciones de manera que abarquen el intervalo temporal más amplio posible y que, por lo tanto, no estén correlacionadas.

### 8.2.2. Componente Stark

Como ya se ha comentado en la sección 4.4, la componente Stark es el error introducido por la ACM en la energía de interacción entre soluto y disolvente. En la tabla 8.3 se muestran los resultados obtenidos para la energía de interacción electrostática y el momento dipolar cuando se aplica la ACM,  $E_{ACM}^{elec}$  y  $\mu_{ACM}$ , y cuando no se hace,  $\langle E_{QM/MM}^{elec} \rangle$  y  $\langle \mu \rangle$ . La contribución de Van der Waals a la energía de interacción soluto-disolvente,  $E_{QM/MM}^{vdw}$ , al no depender de las coordenadas electrónicas del soluto, puede calcularse de manera sencilla sin recurrir a la ACM, como promedio de los valores para las  $N$  configuraciones consideradas. Por lo tanto, la componente Stark es simplemente:

$$\Delta W_{Stark} = \langle E_{QM/MM}^{elec} \rangle - E_{ACM}^{elec} \quad (8.1)$$

**Tabla 8.3:** Energías de interacción electrostática (en kcal/mol) y momentos dipolares (en D) de las moléculas estudiadas, calculados como valor promedio ( $\langle E_{\text{QM/MM}}^{\text{elec}} \rangle$  y  $\langle \mu \rangle$ ) y con la aproximación del campo medio ( $E_{\text{ACM}}^{\text{elec}}$  y  $\mu_{\text{ACM}}$ ). Para el metanol se dan los resultados con 100 y 1000 configuraciones.

	$\langle E_{\text{QM/MM}}^{\text{elec}} \rangle$	$E_{\text{ACM}}^{\text{elec}}$	$\Delta W_{\text{Stark}}$	$\langle \mu \rangle$	$\mu_{\text{ACM}}$	$\langle \mu \rangle - \mu_{\text{ACM}}$
HF						
Metanol (1000)	-19,0	-18,6	0,4 (2,1 %)	2,46	2,45	0,01 (0,4 %)
Metanol (100)	-19,1	-18,7	0,4 (2,1 %)	2,46	2,45	0,01 (0,4 %)
Etanol	-16,5	-16,0	0,5 (3,0 %)	2,27	2,25	0,02 (0,9 %)
Propanol	-14,3	-14,0	0,3 (2,1 %)	2,15	2,13	0,02 (0,9 %)
MP2						
Metanol (1000)	-18,3	-17,9	0,4 (2,2 %)			
Metanol (100)	-18,3	-17,9	0,4 (2,2 %)			
Etanol	-15,8	-15,4	0,4 (2,5 %)			
Propanol	-13,7	-13,5	0,2 (1,5 %)			
CASSCF						
Formaldehído	-9,2	-8,8	0,4 (4,3 %)	2,99	2,99	0,00 (0,0 %)
Acetaldehído	-8,9	-8,5	0,4 (4,5 %)	3,46	3,46	0,00 (0,0 %)
Acetona	-21,9	-21,1	0,8 (3,6 %)	4,48	4,47	0,01 (0,2 %)

Los valores de la componente Stark están entre 0,3 kcal/mol y 0,8 kcal/mol, lo que representa menos del 5 % en todos los casos. Las diferencias en los momentos dipolares son incluso más bajas, de 0,2 D como máximo, lo que no llega al 1 % de error. La inclusión de la correlación electrónica con el método MP2 hace disminuir ligeramente la energía de interacción, algo que ya se había encontrado en estudios realizados con modelos de continuo [210], pero tiene un efecto muy pequeño en el valor de la componente Stark. Los momentos dipolares no varían con la introducción del método MP2.

Resulta claro que la componente Stark representa una parte pequeña de la energía de interacción soluto-disolvente. La ACM se comporta mejor aún en el caso de propiedades moleculares como el momento dipolar. En consecuencia, puede concluirse que esta aproximación no introduce errores significativos ni en la energía ni en las propiedades moleculares, mientras que, como ya se ha mencionado, permite un ahorro considerable de tiempo de cálculo.

Pese a que los errores son pequeños, sería deseable contar con un método que permita reducirlos o, al menos, evaluar su magnitud. El procedimiento correcto para la evaluación de la componente Stark requiere efectuar un cálculo cuántico para cada una de las configuraciones del sistema consideradas, lo que anula los beneficios de la ACM. Sin embargo, según la ecuación (4.23) (pág. 88), la compo-

nente Stark es proporcional a la polarizabilidad de la molécula de soluto (obtenida mediante un único cálculo cuántico) y a la fluctuación del campo eléctrico generado por el disolvente (calculada de manera clásica). Por lo tanto, mediante esta ecuación puede obtenerse un valor aproximado para la componente Stark sin necesidad de renunciar a los beneficios de la ACM. En la tabla 8.4 se dan los valores de polarizabilidad y de fluctuación del campo eléctrico para cada uno de los sistemas, así como los valores exacto,  $\Delta W_{\text{Stark}}$ , y aproximado,  $\Delta W_{\text{Stark}}^{\text{aprox}}$ , de la componente Stark.

**Tabla 8.4:** Polarizabilidad en disolución (en  $a_0^3$ ), fluctuación del campo eléctrico (en  $10^{-5} \cdot E_h/e \cdot a_0$ ) y comparación entre los valores de la componente Stark calculada exactamente y con la fórmula aproximada de la ecuación (4.23) (en kcal/mol).

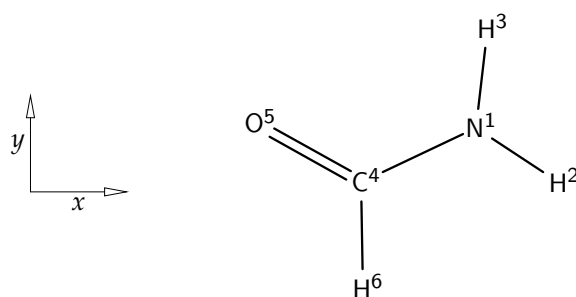
	$\alpha = 1/3 \text{tr } \alpha$	$1/3 \text{tr } (\langle E^2 \rangle - \langle E \rangle^2)$	$\Delta W_{\text{Stark}}$	$\Delta W_{\text{Stark}}^{\text{aprox}}$ (ec. 4.23)
Metanol (1000)	19,45	3,87	0,44	0,20 (45 %)
Metanol (100)	19,45	4,10	0,40	0,20 (50 %)
Etanol	31,00	2,77	0,51	0,25 (49 %)
Propanol	42,11	1,15	0,29	0,15 (52 %)
Formaldehído	14,40	6,60	0,40	0,28 (70 %)
Acetaldehído	25,28	3,24	0,40	0,25 (63 %)
Acetona	35,58	4,18	0,78	0,45 (58 %)

Se aprecia que, en general, los valores aproximados están en torno a la mitad de los valores exactos. La razón es que la ecuación (4.23) procede de un tratamiento muy simplista, en el que se aproxima el soluto por un dipolo ideal y se desprecian todos los términos multipolares de mayor orden. Sin embargo, añadiendo este valor aproximado de la componente Stark a los resultados obtenidos para  $E_{\text{ACM}}^{\text{elec}}$ , el error con respecto a  $\langle E_{\text{QM/MM}}^{\text{elec}} \rangle$  se reduce a menos de 1,5 %. Por lo tanto, la fórmula (4.23) proporciona un método sencillo y poco costoso de reducir los pequeños errores en la energía de interacción soluto-disolvente debidos al empleo de la ACM.

### 8.2.3. Gradiente de la energía libre

El efecto que la ACM tiene en el gradiente calculado sobre la superficie de energía libre es importante cuando se realiza una optimización de geometría. Principalmente, la magnitud del gradiente permite identificar los puntos estacionarios en la superficie, en ellos el gradiente debe anularse. Además, el gradiente proporciona una información básica sobre la forma de la superficie y juega un importante papel en los distintos algoritmos de optimización. Por ello, para poder aplicar con confianza la ACM en los procesos de optimización de geometrías en disolución, es necesario comprobar que esta aproximación no introduce errores significativos en el gradiente.





**Figura 8.5:** Molécula de formamida orientada según sus ejes principales de inercia. Se muestra también la numeración de los átomos.

Con este objetivo se llevó a cabo un estudio del gradiente sobre la superficie de energía libre de la formamida en disolución acuosa. Como en el caso de las energías, se compararon los valores obtenidos para el gradiente como promedio de 1000 cálculos cuánticos individuales y mediante la aplicación de la ACM. Hay que tener en cuenta que el gradiente calculado no incluye la contribución de las interacciones de Van der Waals, puesto que esta contribución se considera aditiva y no se ve afectada por la ACM. Los resultados se muestran en la tabla 8.6-A.

Comparando los valores de  $\langle g \rangle$  y de  $g_{ACM}$  se observa que el mayor error cometido en las componentes del gradiente es de  $0,39 \cdot 10^{-3} E_h/a_0$  (en la componente  $x$  sobre  $O^5$ ), lo que representa un error de alrededor del 4% con respecto a la «raíz cuadrática media» (*root mean square*, RMS) del gradiente. La diferencia en la RMS, sin embargo, es de sólo  $0,09 \cdot 10^{-3} E_h/a_0$ . En general, el valor del gradiente obtenido mediante la ACM, con un solo cálculo cuántico, es muy similar al resultante de un promedio de 1000 cálculos cuánticos.

Con el fin de evaluar la importancia del error introducido por la ACM en el gradiente, se analizó la influencia que el intervalo temporal de las simulaciones tiene en el valor obtenido. Se realizó una simulación de 300 ps y se dividió en seis intervalos de 50 ps; para cada uno de estos intervalos se hizo un cálculo aplicando la ACM. De los seis valores obtenidos puede extraerse la media y la desviación típica. El mismo proceso se repitió dividiendo la simulación inicial en tres intervalos de 100 ps. Ambos resultados se muestran en la tabla 8.6-B. En este caso se da el gradiente total, incluyendo la contribución de Van der Waals.

La desviación típica del gradiente permite evaluar error introducido por el tamaño finito de las muestras configuracionales y puede considerarse como una medida de la precisión que puede alcanzarse para el gradiente sobre la superficie de energía libre. Para los cálculos sobre muestras de 50 ps, el valor máximo de la desviación típica es de  $2,0 \cdot 10^{-3} E_h/a_0$  (en la componente  $x$  sobre  $C^4$ ), un error cinco veces mayor que el introducido por la ACM. Al duplicar el tamaño de la muestra,

**Tabla 8.6:** Gradiente en coordenadas cartesianas de la energía libre de la formamida en disolución acuosa, en  $10^{-3} \cdot E_h/a_0$ .  $\langle g \rangle$  es el promedio de 1000 cálculos cuánticos,  $g_{ACM}$  es el resultado de aplicar la ACM. El sistema de referencia está orientado según los ejes principales de inercia de la molécula, tal como se muestra en la figura 8.5.

		A		B			
		$\langle g \rangle$	$g_{ACM}$	50 ps		100 ps	
				Media	Desv.	Media	Desv.
N <sup>1</sup>	x	12,914	12,827	-0,765	1,154	-0,706	0,173
	y	3,858	3,831	-0,892	0,372	-0,798	0,262
	z	0,084	0,085	-0,237	0,367	-0,212	0,045
H <sup>2</sup>	x	-3,365	-3,132	-0,054	0,625	-0,084	0,175
	y	3,049	2,928	0,616	0,348	0,632	0,132
	z	0,053	0,049	0,182	0,342	0,151	0,098
H <sup>3</sup>	x	-1,116	-1,081	-0,316	0,169	-0,288	0,097
	y	-5,347	-5,081	-0,424	0,545	-0,498	0,343
	z	-0,006	-0,004	0,110	0,141	0,101	0,074
C <sup>4</sup>	x	-26,864	-26,816	-3,009	2,004	-3,040	0,261
	y	5,473	5,487	-0,682	0,557	-0,662	0,273
	z	-0,002	-0,010	0,003	0,046	-0,011	0,028
O <sup>5</sup>	x	18,255	17,862	-1,820	1,510	-1,695	0,496
	y	-10,052	-9,788	-1,389	0,727	-1,400	0,179
	z	-0,041	-0,026	0,164	0,105	0,245	0,136
H <sup>6</sup>	x	1,567	1,571	-0,054	0,063	-0,051	0,048
	y	-2,136	-2,240	0,102	0,135	0,107	0,036
	z	-0,004	0,002	-0,002	0,063	0,000	0,007
RMS		8,898	8,807	0,970	0,736	0,959	0,202

- A** Comparación de los resultados obtenidos con la ACM y sin ella. No se incluye la contribución de Van der Waals.
- B** Valor medio y desviación típica de los gradientes obtenidos aplicando la ACM sobre 50 o 100 ps.

la desviación típica máxima se reduce a  $0,5 \cdot 10^{-3} E_h/a_0$  (en la componente  $x$  sobre  $O^5$ ), del mismo orden de magnitud que el error debido a la ACM.

Se puede deducir, por lo tanto, que la importancia relativa del error introducido por la ACM en el cálculo del gradiente de la energía libre depende de la duración de las simulaciones sobre las que se aplica la aproximación. Para simulaciones de 50 ps, este error es menos importante que el debido al tamaño finito de las simulaciones. Sin embargo, para simulaciones mayores de 100 ps, la ACM puede ser el factor limitante en la precisión del gradiente y, por consiguiente, en la precisión de la optimización de geometrías.

### 8.3. Conclusiones

La aproximación del campo medio permite una importante reducción del número de cálculos necesarios en el estudio teórico de sistemas líquidos. Se ha comprobado que los errores introducidos por esta aproximación son pequeños, más aún teniendo en cuenta otras aproximaciones incorporadas en la mayoría de los métodos QM/MM —desestimación del intercambio electrónico intermolecular, parametrización de las interacciones de Van der Waals, etc.

Además, pueden extraerse la siguientes conclusiones:

- El error en los momentos dipolares es prácticamente despreciable.
- El error en la energía de interacción soluto-disolvente es de alrededor de 0,5 kcal/mol para todos los sistemas estudiados. Este error puede reducirse casi un 50 % por medio de la fórmula aproximada (4.23), que es una manera rápida y sencilla de obtener una estimación para la componente Stark.
- El error cometido en el cálculo de la energía está dentro de lo que podría llamarse «precisión química» (1 kcal/mol) y es comparable a los errores introducidos por el empleo de campos de fuerzas clásicos. Así, por ejemplo, en un estudio sobre los confórmeros de la glucosa, Barrows y col. [211] proporcionan los siguientes errores para algunos campos de fuerzas: 0,6 kcal/mol para OPLS-AA, entre 0,7 kcal/mol y 0,8 kcal/mol para AMBER, 0,9 kcal/mol para MMFF, 1,5 kcal/mol para MM3 y 2,1 kcal/mol para GROMOS.
- El error en el gradiente sobre la superficie de energía libre es menor que el debido a la duración finita de las simulaciones. Con simulaciones de más 150 ps, la ACM podría imponer un límite a la precisión que puede alcanzarse en las optimizaciones de geometría.



---

---

# 9

---

## OPTIMIZACIÓN DE GEOMETRÍAS EN DISOLUCIÓN

La optimización de geometrías moleculares en disolución es un problema complejo. Especialmente en el caso de líquidos polares, es necesario considerar un elevado número de moléculas de disolvente para describir adecuadamente el campo eléctrico generado sobre el soluto. El problema puede simplificarse haciendo uso de la ACM, cuya validez se ha demostrado en el capítulo 8.

En el capítulo 6 se presentaron algunos de los algoritmos más empleados para la optimización de geometrías y, en la sección 6.3, se indicó la manera de incluirlos en el método ASEP/MD. En este capítulo se verifica el buen comportamiento del método propuesto, aplicándolo a algunos sistemas relativamente sencillos. En estos sistemas existen importantes interacciones electrostáticas por enlace de hidrógeno y es previsible que la influencia del disolvente sobre el soluto sea especialmente significativa.

### 9.1. Detalles de los cálculos

Se estudiaron tanto líquidos puros como disoluciones; en todos los casos, sistemas con fuertes interacciones intermoleculares. Como ejemplos de líquidos puros se consideraron el agua y el metanol.

Los cálculos cuánticos de la molécula de agua se realizaron mediante el método MP2 y en la dinámica molecular se simularon 215 moléculas. Para los cálculos del metanol se empleó DFT con el funcional BP86 y se simularon 128 moléculas. En ambos casos se utilizó la base aug-cc-pVDZ y las interacciones intermoleculares se truncaron a 9,0 Å.

Como ejemplo de disolución se estudió la formamida en disolución acuosa, en las mismas condiciones detalladas en el capítulo anterior (pág. 118).

En los tres sistemas estudiados se comenzó por optimizar la geometría del soluto en vacío para, a continuación, efectuar un cálculo ASEP/MD parcialmente acoplado con optimización de geometría; el método utilizado para el cálculo de las cargas atómicas fue CHELP. Los algoritmos de optimización utilizados en cada caso fueron los siguientes:

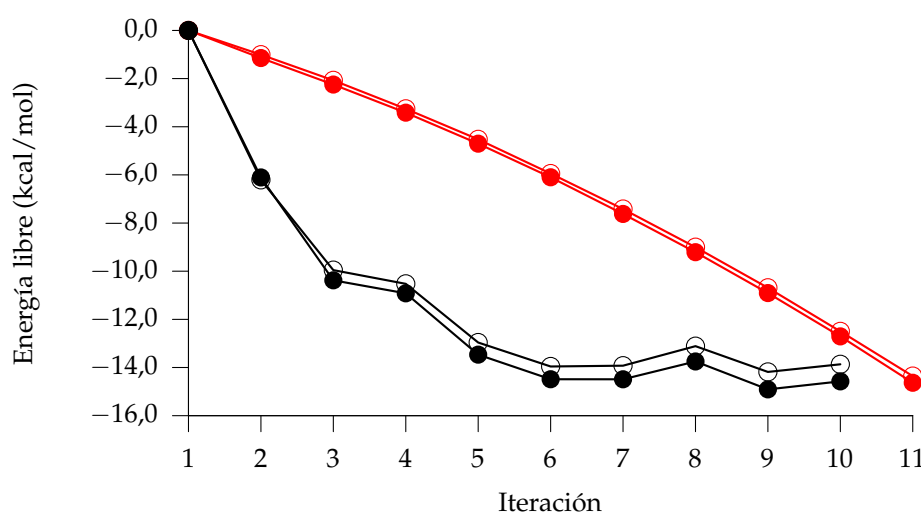
- Para el agua se aplicó una sola iteración del método de la pendiente máxima, con búsqueda lineal del mínimo de energía.
- Para el metanol se empleó un método cuasi-Newton, con actualización BFGS y también con una única iteración tras cada dinámica.
- Para la formamida se utilizó igualmente un método cuasi-Newton con actualización BFGS. Se hicieron tres tipos de cálculo:
  - Con dinámicas de 75 ps, se realizó una optimización completa —hasta la convergencia— tras cada simulación.
  - El cálculo anterior se continuó con dinámicas más largas, de 150 ps.
  - Con dinámicas de 25 ps, se efectuó una sola iteración de optimización tras cada simulación, aplicando una búsqueda lineal.

El objeto de todos estos cálculos es comprobar el efecto que, sobre el comportamiento del método y sobre el resultado final, tienen la duración de las simulaciones y las diferentes opciones empleadas durante la optimización.

## 9.2. Resultados y discusión

### 9.2.1. Variación de la energía libre

En primer lugar se procedió a comprobar la validez del método empleado para el cálculo de las variaciones de energía libre. Para ello se realizó un cálculo de optimización en disolución de la formamida y se obtuvo la diferencia de energía libre entre cada dos iteraciones sucesivas del proceso ASEP/MD, como se expone en la sección 5.2. Posteriormente, y a efectos de comparación, se realizó también un cálculo directo de la diferencia de energía libre entre las estructuras inicial y final de la formamida, obtenidas en el cálculo previo; las estructuras de los estados intermedios se obtuvieron en este caso por interpolación lineal de la geometría y cargas atómicas. En cada caso, la diferencia de energía libre entre cada dos estructuras se calculó partiendo de una dinámica molecular con la primera estructura



**Figura 9.1:** Cálculo de la variación de energía libre entre dos estructuras de la formamida durante la optimización (negro) y mediante un cálculo directo (rojo). En cada caso se realizó un cálculo hacia adelante (símbolos rellenos) y hacia atrás (símbolos vacíos).

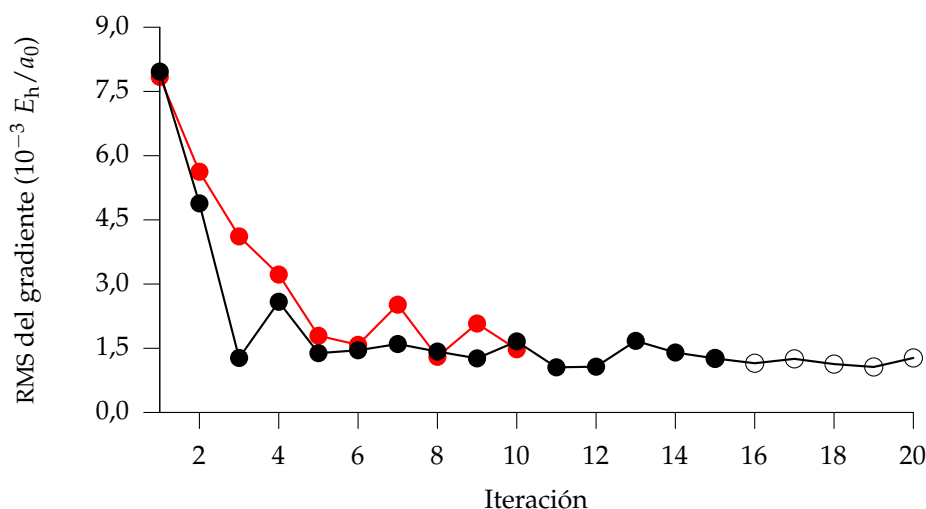
(versión «hacia adelante») o con la segunda (versión «hacia atrás»). Los resultados obtenidos se muestran en la figura 9.1.

En la figura 9.1, la magnitud a comparar es la variación de energía libre entre los puntos inicial y final de cada cálculo, mientras que la trayectoria de cada línea pone de manifiesto la diferencia en los métodos empleados. El cálculo directo con interpolación lineal dio como resultado  $-14,6$  kcal/mol y  $-14,4$  kcal/mol, para las versiones hacia adelante y hacia atrás, respectivamente. Cuando la diferencia de energía libre se obtiene durante el proceso de optimización, los resultados fueron  $-14,6$  kcal/mol y  $-13,9$  kcal/mol. Como se observa, la diferencia entre los valores hacia adelante y hacia atrás es mayor en este último caso, no obstante los errores son suficientemente pequeños para poder confiar en los valores obtenidos. En resumen, se comprueba la validez del procedimiento utilizado para el cálculo de las diferencias de la energía libre entre distintas estructuras del soluto.

### 9.2.2. Formamida

Se analizan en primer lugar los resultados de la formamida en disolución acuosa. En la figura 9.2 se presenta la variación de la RMS del gradiente sobre la superficie de energía libre durante el proceso ASEP/MD. En cada iteración se efectúa una optimización completa de la geometría de la formamida, por lo que el gradiente representado es el inicial, antes de realizar cada optimización. En las simulaciones de 50 ps se alcanza la convergencia en 5 o 6 iteraciones y, a partir de entonces, el

gradiente fluctúa alrededor de un valor cercano a  $0,0015 E_h/a_0$ . Este valor es algo menor que el obtenido por otros autores que también emplean el método FEG. Por ejemplo, Okuyama-Yoshida y col. [157], en la optimización de la glicina, obtuvieron un gradiente de  $0,0025 E_h/a_0$ , probablemente debido a la menor duración de sus simulaciones: 10 ps. El valor obtenido en nuestro caso es cinco veces mayor que el umbral predeterminado para las optimizaciones en vacío en el programa *Gaussian 98* ( $0,0003 E_h/a_0$  para la RMS del gradiente).



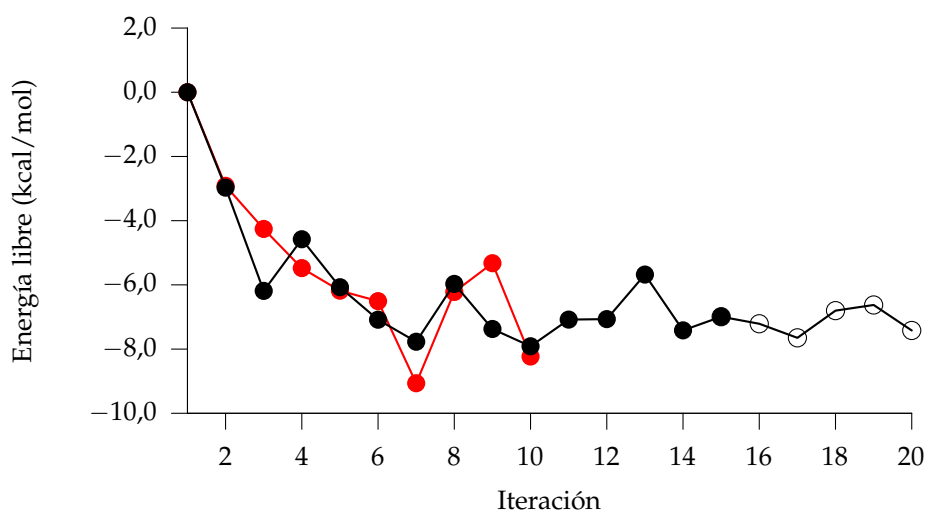
**Figura 9.2:** Evolución de la RMS del gradiente sobre la superficie de energía libre de la formamida en disolución acuosa durante el proceso de optimización. Se muestra el gradiente inicial tras cada dinámica molecular para simulaciones de 50 ps (negro), 150 ps (círculos vacíos) y 25 ps (rojo).

Tal y como es de esperar, se observa que al aumentar la duración de las simulaciones disminuyen las fluctuaciones del gradiente de la energía libre. Sin embargo, no ocurre lo mismo con la magnitud del gradiente, que se mantiene prácticamente constante en alrededor de  $0,0015 E_h/a_0$ . Esto es, probablemente, un artefacto del método, relacionado con el diferente tratamiento que se da al soluto durante las simulaciones (un conjunto clásico de monopolos eléctricos) y durante la optimización (un tratamiento cuántico detallado de la función de onda electrónica).

En la figura 9.3 se muestra la evolución de la energía libre durante el proceso de optimización, tomando como cero de energía el valor correspondiente al sistema formamida-agua en el que la estructura de la formamida es la optimizada en vacío. En este caso, la diferencia entre el estado inicial y final se puede considerar una «energía libre de polarización»,  $\Delta G^{\text{pol}}$ . Como ocurre con el gradiente, la convergencia se alcanza en 5 o 6 iteraciones y a partir de entonces  $\Delta G^{\text{pol}}$  comienza a fluctuar. También en este caso se observa una disminución en las fluctuaciones al aumentar los tiempos de simulación, pero el valor medio de los últimos ciclos es prácticamen-



te constante:  $-7,2 \pm 1,4$  kcal/mol,  $-7,1 \pm 0,7$  kcal/mol y  $-7,2 \pm 0,4$  kcal/mol para las simulaciones de 25, 50 y 150 ps, respectivamente. Es interesante recordar que las fluctuaciones debidas a los efectos térmicos son aproximadamente de magnitud  $k_B T$ , por lo que a 298 K son del orden de 0,6 kcal/mol.

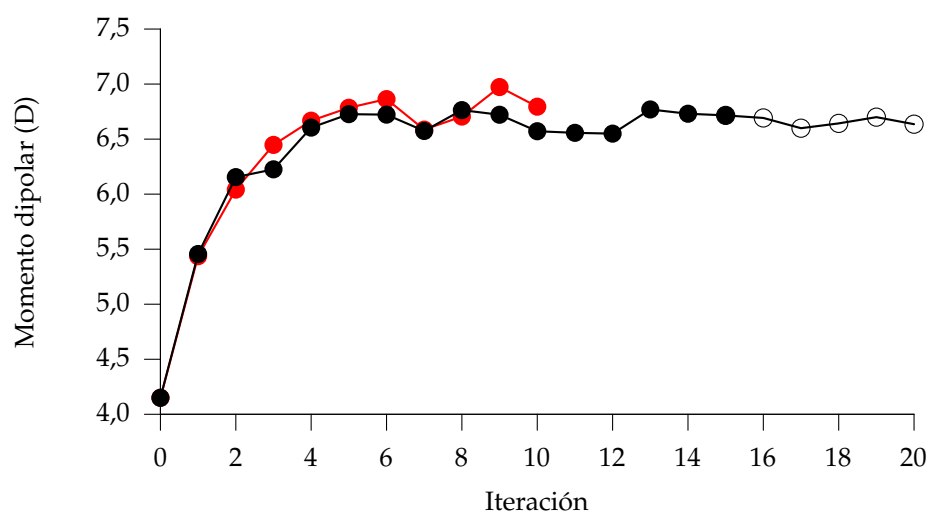


**Figura 9.3:** Evolución de la energía libre de polarización de la formamida en disolución acuosa durante el proceso de optimización. Simulaciones de 50 ps (negro), 150 ps (círculos vacíos) y 25 ps (rojo).

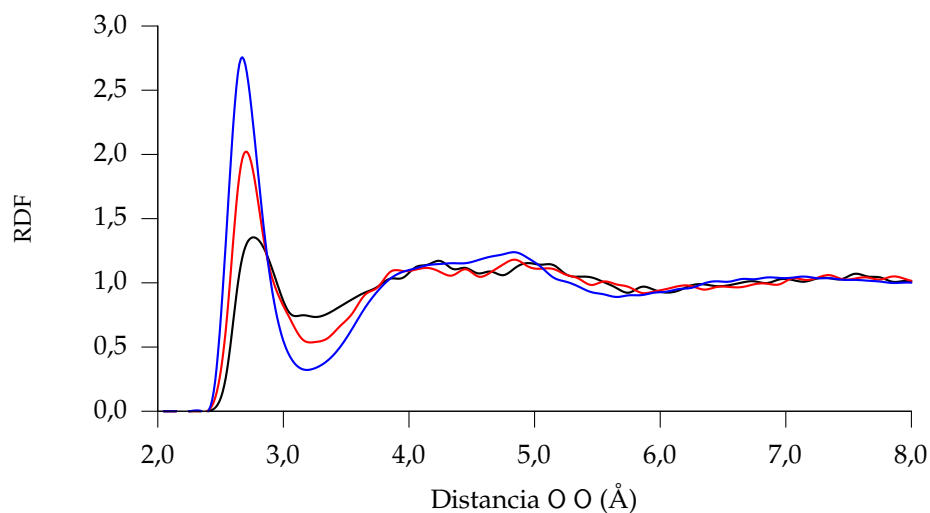
Las tendencias en la energía libre y en su gradiente confirman que, efectivamente, el proceso de optimización se dirige a un mínimo en la superficie de energía libre. Ambas magnitudes disminuyen durante la optimización y el gradiente tiende a anularse, aunque, como se ha comentado, no llega a hacerlo completamente.

Durante el proceso de optimización ASEP/MD, el soluto sufre cambios tanto en su geometría como en su distribución electrónica. La polarización del soluto se aprecia claramente en la figura 9.4, donde se representa el valor del momento dipolar de la formamida en cada iteración. El momento dipolar promedio inducido en el soluto es de 2,5 D (de 4,1 D a 6,6 D), un aumento de casi el 60% respecto al valor calculado en fase gas. Esta polarización del soluto está acompañada por un incremento paralelo de la estructura del disolvente alrededor del soluto. Esto se observa en la evolución de la función de distribución radial del oxígeno del agua alrededor del oxígeno de la formamida, representada en la figura 9.5. Según progresa la polarización, aumenta la altura del primer pico de la RDF y se desplaza a distancias más cortas. Durante este proceso la distribución electrónica del soluto y la estructura del disolvente se equilibran mutuamente.

Por su parte, las geometrías calculadas para la formamida en fase gaseosa y en disolución se muestran en la figura 9.6 y en la tabla 9.7; las geometrías en disolución

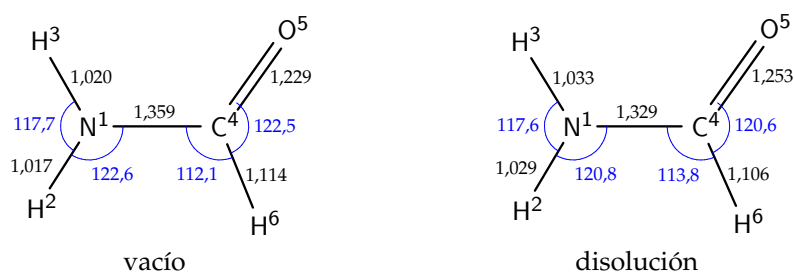


**Figura 9.4:** Evolución del momento dipolar de la molécula de formamida en disolución acuosa durante el proceso de optimización. Simulaciones de 50 ps (negro), 150 ps (círculos vacíos) y 25 ps (rojo).



**Figura 9.5:** Evolución de la función de distribución radial entre el oxígeno de la formamida y el oxígeno del agua durante el proceso de optimización, con simulaciones de 50 ps. Se muestra la RDF en la primera iteración (negro), la segunda (rojo) y el promedio de las iteraciones 5-15 (azul).

corresponden al promedio de los últimos ciclos (5 para las simulaciones de 150 ps, 10 para las de 50 ps), en los que apenas hay variación; se muestran estos promedios para los cálculos con simulaciones de 50 ps y 150 ps. Como comparación se incluyen también los resultados obtenidos con el método DFT/MM desarrollado en Nancy [136]. La primera conclusión es que la geometría optimizada apenas depende de la duración de las simulaciones, los resultados obtenidos con 50 ps y 150 ps son prácticamente idénticos. Además, a pesar de las fluctuaciones en el gradiente asociadas al tamaño de la simulación, las geometrías del soluto son muy estables, las desviaciones típicas son menores que  $0,001 \text{ \AA}$  para las distancias y  $0,2^\circ$  para los ángulos de enlace. Aunque estos resultados no son directamente comparables con los obtenidos por [Chalmet y Ruiz-López](#), debido a pequeñas diferencias en el conjunto de base y en el funcional de correlación-intercambio, puede comprobarse que son muy similares.



**Figura 9.6:** Geometrías en vacío y en disolución de la molécula de formamida. Distancias en  $\text{\AA}$ , ángulos en grados.

Como puede observarse, el disolvente induce una disminución de la distancia C–N, un aumento de la distancia C=O y variaciones despreciables en los ángulos de enlace. Estos resultados pueden interpretarse de manera sencilla por la estabilización de la estructura *zwitteriónica* del soluto ( $\text{H}_2\text{N}-\text{CH}=\text{O} \leftrightarrow \text{H}_2\text{N}^\oplus=\text{CH}-\text{O}^\ominus$ ) gracias a la interacción electrostática con el disolvente, un hecho confirmado por el comportamiento de las cargas atómicas del soluto (tabla 9.8): la densidad electrónica en el átomo de N disminuye respecto a la estructura en fase gaseosa mientras que aumenta en el átomo de O.

### 9.2.3. Agua y metanol

El siguiente sistema analizado es el agua líquida. Como ocurre en la formamida, la convergencia se alcanza en unas 5 iteraciones del proceso ASEP/MD. En la figura 9.9 se muestra la variación del gradiente de la energía libre durante la optimización, el valor final que alcanza el gradiente es muy parecido al caso de la formamida, en torno a  $0,0015 E_h/a_0$ . La magnitud de las fluctuaciones de la energía

**Tabla 9.7:** Evolución de la geometría de la formamida en disolución acuosa durante el proceso de optimización. La iteración 0 es el cálculo en vacío. Se muestra la geometría promedio de los últimos ciclos y la diferencia entre la geometría en vacío y en disolución. Como comparación se muestran también los resultados obtenidos con el método DFT/MM [136]. Distancias en Å, ángulos en grados.

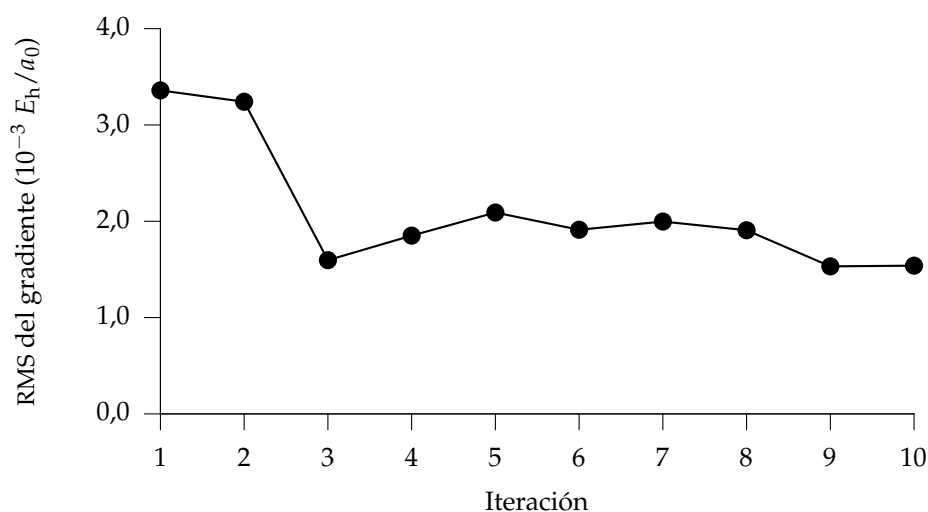
Distancia o ángulo	ASEP/MD (50 ps)							
	0	1	2	3	4	5	10	15
N1-H2	1,017	1,022	1,025	1,028	1,030	1,031	1,030	1,029
N1-H3	1,020	1,025	1,029	1,029	1,030	1,032	1,032	1,033
N1-C4	1,359	1,343	1,335	1,334	1,331	1,329	1,331	1,329
C4=O5	1,229	1,240	1,248	1,248	1,253	1,254	1,252	1,254
C4-H6	1,114	1,109	1,107	1,106	1,106	1,105	1,106	1,105
C4-N1-H2	122,6	121,9	121,5	121,6	121,1	120,6	121,0	120,6
C4-N1-H3	119,7	120,4	120,9	120,7	121,4	121,6	121,3	122,0
N1-C4=O5	125,4	125,5	125,5	125,1	125,6	125,6	125,3	125,6
N1-C4-H6	112,1	113,0	113,5	113,8	113,7	113,8	113,9	113,8

Distancia o ángulo	ASEP/MD				DFT/MM [136]		
	50 ps		150 ps		Vacío	Disol.	Difer.
	Prom.	Difer.	Prom.	Difer.			
N1-H2	1,029	+0,013	1,030	+0,013	1,016	1,040	+0,024
N1-H3	1,033	+0,013	1,033	+0,013	1,018	1,041	+0,023
N1-C4	1,329	-0,029	1,330	-0,029	1,366	1,338	-0,028
C4=O5	1,253	+0,024	1,253	+0,024	1,227	1,262	+0,035
C4-H6	1,106	-0,008	1,106	-0,008	1,114	1,112	-0,002
C4-N1-H2	120,8	-1,9	120,8	-1,9	121,7	119,8	-1,9
C4-N1-H3	121,6	+2,0	121,7	+2,0	119,0	120,5	+1,5
N1-C4=O5	125,5	+0,2	125,6	+0,2	124,4	124,6	+0,2
N1-C4-H6	113,8	+1,8	113,8	+1,7	112,5	114,2	+1,7

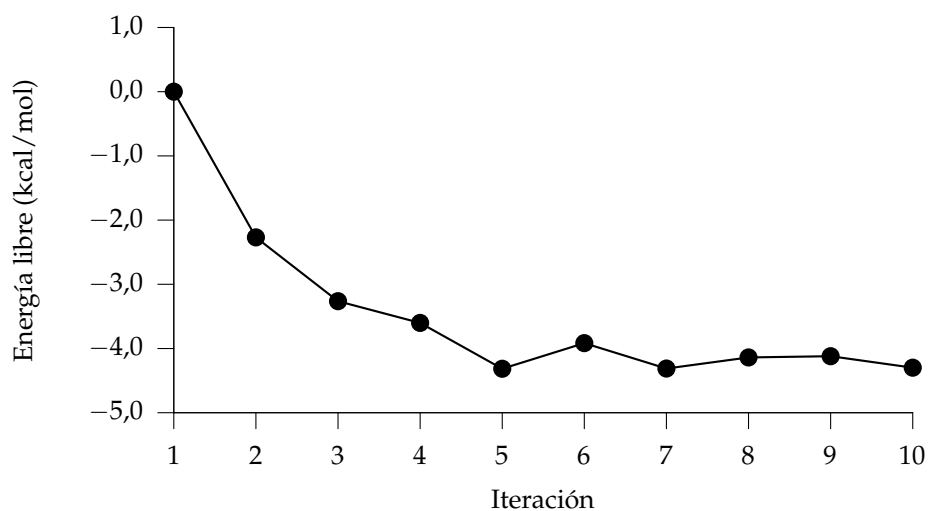
**Tabla 9.8:** Cargas atómicas (en  $e$ ) calculadas para la formamida en vacío y en disolución.

	Vacío	Disolución
N <sup>1</sup>	-0,749	-0,719
H <sup>2</sup>	0,325	0,368
H <sup>3</sup>	0,333	0,388
C <sup>4</sup>	0,805	0,911
O <sup>5</sup>	-0,601	-0,851
H <sup>6</sup>	-0,113	-0,098

libre (figura 9.10) es de unas 0,2 kcal/mol, claramente menor que las fluctuaciones debidas a efectos térmicos.



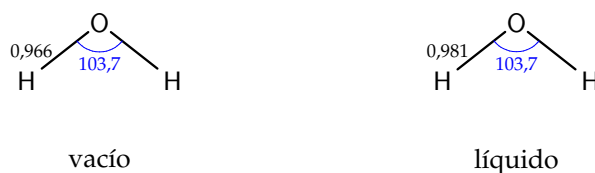
**Figura 9.9:** Evolución de la RMS del gradiente sobre la superficie de energía libre de una molécula de agua en estado líquido durante el proceso de optimización.



**Figura 9.10:** Evolución de la energía libre de polarización de la molécula de agua en estado líquido durante el proceso de optimización.

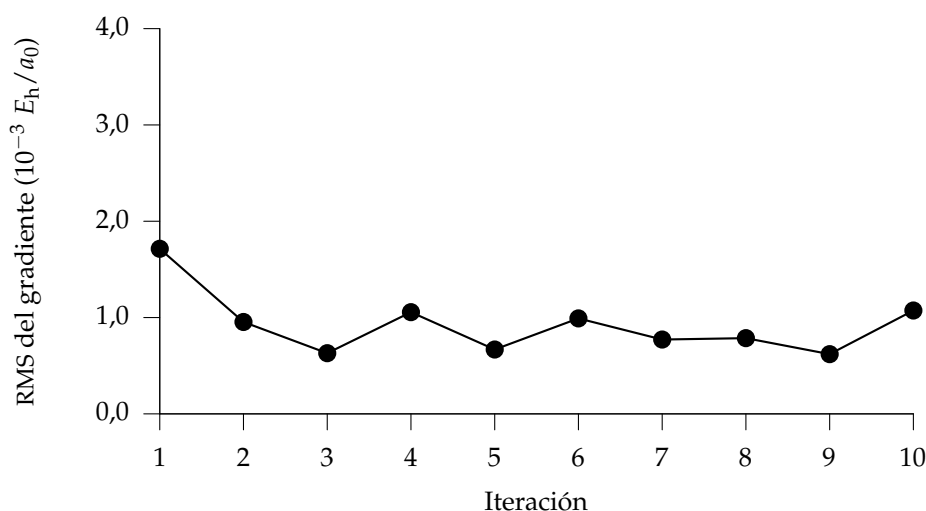
La molécula de agua experimenta además una fuerte polarización al pasar de la fase gaseosa a la fase líquida. El momento dipolar inducido calculado es de 0,76 D, valor que representa un 27% con respecto al momento dipolar en vacío y que es similar a los resultados obtenidos por otros autores [212-216] y al valor experimental estimado: entre 0,70 D y 1,0 D [217-220]. Los cambios en la geometría del agua

se muestran en la figura 9.11; la distancia O–H aumenta en 0,015 Å, mientras el ángulo H–O–H apenas sufre cambios apreciables.

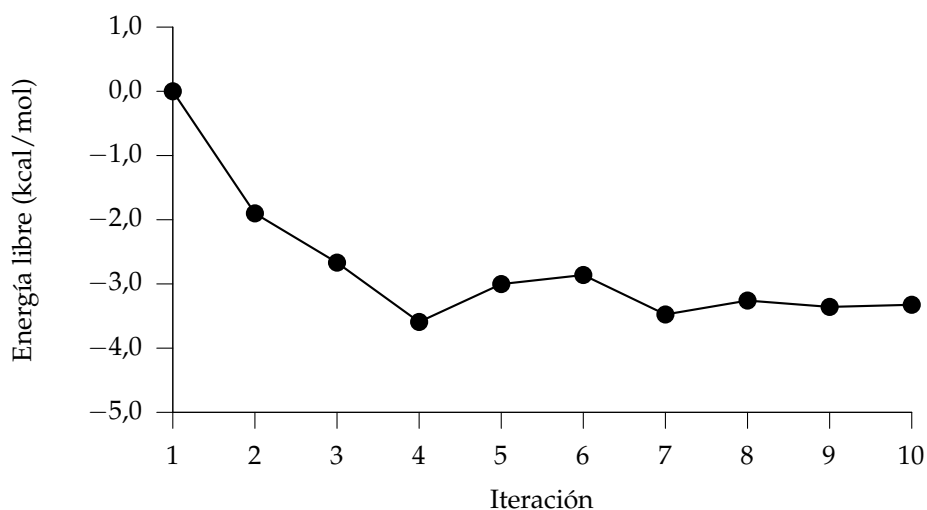


**Figura 9.11:** Geometrías en vacío y en estado líquido de la molécula de agua. Distancias en Å, ángulos en grados.

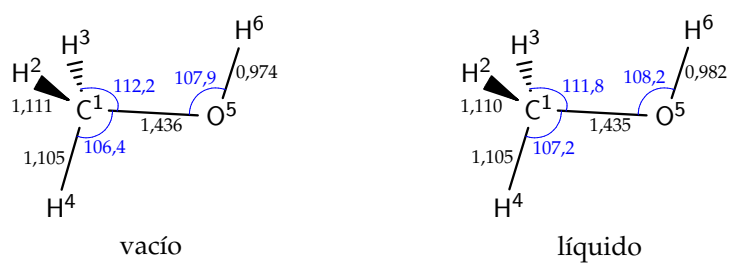
Por último se analiza el metanol en estado líquido. En este caso la convergencia se alcanza en 4 o 5 iteraciones y el gradiente resultante es algo menor que el obtenido para la formamida y el agua, ligeramente por debajo de  $0,001 E_h/a_0$ , como se muestra en la figura 9.12. Las fluctuaciones de la energía libre son de una magnitud similar a las del agua, de alrededor de 0,3 kcal/mol (figura 9.13). El momento dipolar inducido en la molécula de metanol al pasar a la fase líquida es de 0,55 D un valor algo menor que el obtenido para el agua, pero que representa, en este caso, un 35% con respecto al valor en fase gas. Los cambios originados por el disolvente sobre la geometría del soluto se muestran en la figura 9.14, sólo la distancia de enlace O–H sufre una variación apreciable, aumentando en 0,008 Å. Como en los casos anteriores, las fluctuaciones del gradiente no tienen apenas efecto en la geometría calculada y las desviaciones típicas de distancias y ángulos de enlace son mínimas.



**Figura 9.12:** Evolución de la RMS del gradiente sobre la superficie de energía libre de una molécula de metanol en estado líquido durante el proceso de optimización.



**Figura 9.13:** Evolución de la energía libre de polarización de la molécula de metanol en estado líquido durante el proceso de optimización.



**Figura 9.14:** Geometrías en vacío y en estado líquido de la molécula de metanol. Distancias en Å, ángulos en grados.

### 9.3. Conclusiones

En términos generales, se ha comprobado el buen funcionamiento del método propuesto para la optimización de geometrías moleculares en disolución. La incorporación de la optimización al método ASEP/MD sólo requiere, como dato adicional, la obtención del gradiente y, opcionalmente, la hessiana en cada cálculo cuántico.

Las principales conclusiones particulares son:

- En todos los casos estudiados, la convergencia se alcanza en 5 o 6 iteraciones del proceso ASEP/MD, independientemente del método de optimización empleado (cálculo de la hessiana, optimización completa en cada iteración). Para sistemas con un número reducido de átomos, como los estudiados, el cálculo de la hessiana y el empleo de la optimización completa en cada iteración no representan un gran esfuerzo computacional y, sin embargo, deben proporcionar un método más robusto. Para sistemas mayores, en los que los cálculos cuánticos consumen la mayor parte del tiempo y para los que el cálculo de la hessiana puede ser prohibitivo, será recomendable emplear otras variantes menos exigentes (actualización de hessianas aproximadas, número limitado de pasos de optimización).
- El valor final de la RMS del gradiente está entre  $0,001 E_h/a_0$  y  $0,002 E_h/a_0$ , menor que el obtenido por otros autores que emplean el método FEG, pero mayor que la precisión normalmente exigida para cálculos en vacío.
- Tanto las fluctuaciones observadas en la energía libre de polarización (del orden de  $0,5$  kcal/mol para dinámicas de  $50$  ps) como en las geometrías moleculares (alrededor de  $0,001$  Å para las distancias de enlace y  $0,2^\circ$  para los ángulos) indican que el resultado de la optimización es estable.
- En los sistemas estudiados se observa una fuerte polarización del soluto al pasar de la fase gaseosa a la fase líquida. Sin embargo, los cambios en las geometrías, aunque significativos, no son sustanciales y son menores que la influencia puede tener el conjunto de base o el método de cálculo. Por ejemplo, la distancia O–H en el  $H_2O$  en fase gaseosa puede variar en  $0,031$  Å según se optimice la geometría con el método HF, MP2 o DFT.



---

---

# 10

---

## EQUILIBRIO ANOMÉRICO EN LA D-XILOPIRANOSA

Los monosacáridos son uno de los tipos de moléculas más importantes en los sistemas biológicos y son la base de los carbohidratos o hidratos de carbono. Polisacáridos, glucoproteínas y los propios monosacáridos juegan un papel esencial en gran cantidad de procesos bioquímicos, tales como rutas metabólicas, procesos de reconocimiento molecular, almacenamiento de energía, función estructural, etc. El monosacárido más abundante en la naturaleza y, sin duda, el de mayor importancia biológica, es la D-glucosa y no resulta sorprendente que haya sido el foco de atención de gran cantidad de estudios teóricos. No obstante, quedan aún sin resolver muchas cuestiones sobre su conformación, estructura y comportamiento, tanto en fase gaseosa como en disolución.

En disolución, la D-glucosa, como la mayoría de los monosacáridos, sufre una autocondensación espontánea, y se encuentra casi por completo en forma de estructura cíclica: la D-glucopiranososa. En la formación de dicho ciclo se crea un nuevo centro quiral en el carbono  $C^1$  (correspondiente a un grupo carbonilo en la forma abierta), esto da lugar a la existencia de dos especies estereoquímicas, llamadas anómeros. Los dos anómeros se distinguen por la orientación que toma el grupo OH unido al  $C^1$ : en la conformación denominada  ${}^4C_1$ , la más estable, el anómero  $\beta$  tiene todos los grupos hidroxilo en disposición ecuatorial con respecto al anillo, mientras que el anómero  $\alpha$  tiene el grupo hidroxilo del  $C^1$  en disposición axial. Adicionalmente, existe un grupo  $CH_2OH$  unido al  $C^5$  que puede tomar tres conformaciones diferentes por giro alrededor del enlace C-C, dando lugar a tres diferentes rotámeros.

El análisis de las conformaciones de la D-glucopiranososa se complica por la li-

bertad rotacional que tienen los grupos hidroxilo, de manera que, considerando todas las posibles rotaciones de estos grupos, habría que tener en cuenta más de 700 posibles conformeros [221], que pueden encontrarse en dos formas de «silla» ( ${}^1C_4$  y  ${}^4C_1$ ) y en dos formas anoméricas ( $\alpha$  y  $\beta$ ). Algunos de estos conformeros son de alta energía y no contribuyen significativamente a la distribución de poblaciones, pero la existencia de otros conformeros con energías muy similares entre sí dificulta el análisis de las poblaciones relativas de cada uno de ellos.

La situación se complica aún más en disolución, donde las interacciones con el disolvente compiten con las intermoleculares, normalmente disminuyendo las diferencias de estabilidad de los distintos conformeros [211, 221-223]. Por consiguiente, el estudio teórico de la D-glucopiranososa implica el cálculo de diferencias de energía muy pequeñas, de manera que son necesarios métodos muy precisos.

La diferente estabilidad de los conformeros en vacío está dominada por la red de interacciones entre grupos hidroxilo. Es obvio que una descripción adecuada de estas interacciones es de suma importancia en este tipo de cálculos. Está ampliamente aceptado que, para conseguir resultados aceptables de las energías relativas de los conformeros de los monosacáridos es necesario incluir la correlación electrónica en los cálculos y emplear conjuntos de base con funciones difusas [224]. También los efectos del disolvente juegan un papel extremadamente importante a la hora de determinar las estabildades y poblaciones relativas de los posibles conformeros y anomeros de los monosacáridos. Por ejemplo, el efecto anomérico [225], que determina la preferencia por la disposición axial de los sustituyentes electronegativos en el  $C^1$  del anillo de piranososa, hace que el anomero  $\alpha$  sea más estable que el  $\beta$  en vacío. Sin embargo, se ha observado experimentalmente [226] que, en disolución acuosa, el orden de estabilidad se invierte y el anomero  $\beta$  pasa a ser el más abundante.

Debido a la necesidad de utilizar bases con funciones difusas, al tamaño relativamente grande del sistema y a la importancia que tiene el entorno en la estabilidad de los conformeros, el empleo de métodos mecanocuánticos rigurosos para el estudio de los monosacáridos ha sido limitado. El primer estudio *ab initio* dedicado a la D-glucosa en vacío es debido a Polavarapu y Ewig [227], quienes encontraron que, en vacío, el anomero  $\alpha$  es 0,4 kcal/mol más estable que el  $\beta$  y que la orientación antihoraria de los grupos hidroxilo exocíclicos resulta energéticamente favorable. Barrows y col. [228], por su parte, tras realizar un extenso análisis conformacional con mecánica molecular y depurar los resultados con métodos cuánticos, encontraron que la conformación de silla  ${}^4C_1$  de la  $\beta$ -D-glucopiranososa es unas 8 kcal/mol más estable que la  ${}^1C_4$ . Otros autores [211, 229] han obtenido resultados similares posteriormente con métodos más sofisticados.

En cuanto a los cálculos teóricos en disolución, la mayoría de ellos se ha basado en modelos de continuo. Cramer y Truhlar [221] exploraron la estabilidad relativa de

distintos conformeros de la glucosa con los modelos de solvatación semiempíricos AM1-SM2 y PM3-SM3. En trabajos posteriores refinaron la descripción de soluto y disolvente empleando mejores métodos cuánticos y modelos de continuo [211, 228, 230]. Con los métodos más sofisticados reprodujeron la preferencia por los anómeros  $\beta$  en disolución que se encuentra experimentalmente, obteniendo una diferencia de energía promedio de 0,2 kcal/mol entre los dos anómeros. Wladkowski y col. [222] estudiaron los efectos de la solvatación sobre los rotámeros de la D-glucopiranososa usando un modelo de continuo y llegaron a la conclusión de que es necesario incluir explícitamente algunas moléculas del disolvente para describir adecuadamente estos efectos.

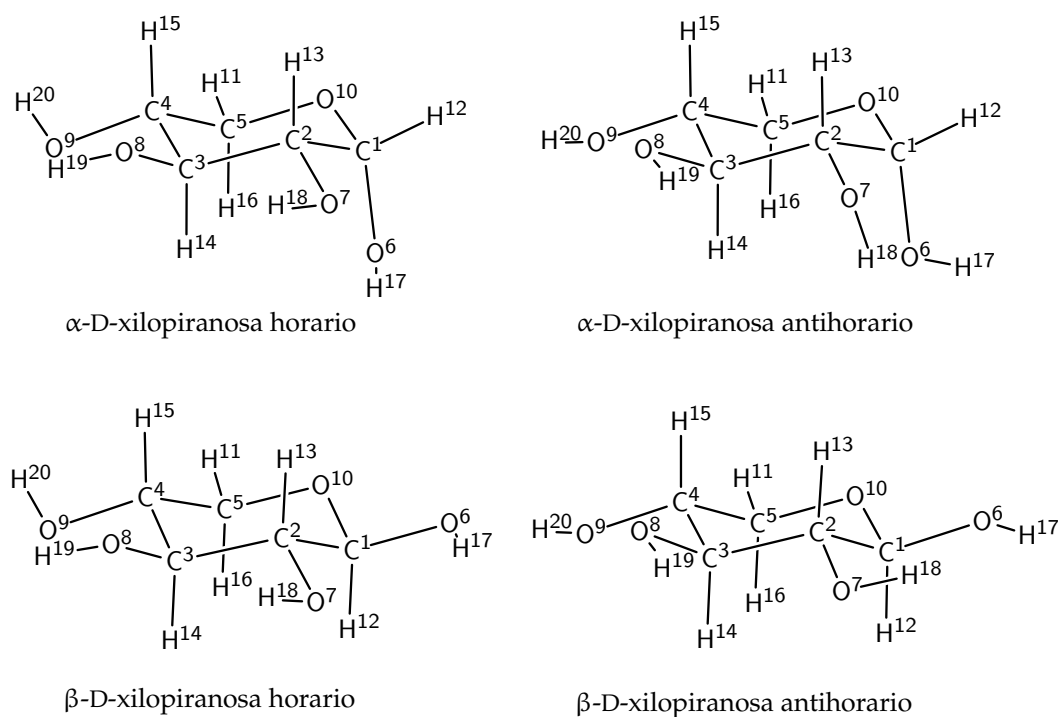
Para abordar este tipo de problema, en el que es necesario combinar una descripción de alto nivel del soluto con una descripción detallada del disolvente a su alrededor, la combinación de mecánica cuántica, mecánica molecular y la aproximación del campo medio, formalizada en el método ASEP/MD, se presenta como una herramienta muy adecuada. Por este motivo, y por el interés que tiene el estudio teórico de los monosacáridos en disolución, se ha llevado a cabo el estudio ASEP/MD del equilibrio anomérico de la D-xilosa.

La D-xilosa es también uno de los monosacáridos más abundantes en la naturaleza. Se trata de una pentosa (monosacárido con cinco átomos de carbono) con configuración equivalente a la de los cinco primeros átomos de carbono de la D-glucosa y, al igual que ésta, forma una estructura cíclica en disolución: la D-xilopiranososa. Los anómeros y conformaciones de la D-xilopiranososa son equivalentes a los de la D-glucopiranososa, sólo se diferencian en que esta última, cuenta con un grupo hidroximetilo del que carece la primera. Esta semejanza estructural hace de la D-xilopiranososa un modelo idóneo para el estudio teórico de la glucosa, ya que se mantienen las principales características estructurales, mientras que la ausencia del grupo  $\text{CH}_2\text{OH}$  reduce el número de conformaciones a tener en cuenta y simplifica los cálculos.

## 10.1. Detalles de los cálculos

Las moléculas estudiadas fueron la  $\alpha$ -D-xilopiranososa y la  $\beta$ -D-xilopiranososa. Se consideró la conformación  ${}^4\text{C}_1$  —la más estable— del anillo de piranososa. Además, en cada caso, existen dos posibles orientaciones de los enlaces O—H para formar enlaces de hidrógeno intramoleculares, los enlaces se disponen en sentido «horario» o «antihorario». Las cuatro estructuras estudiadas se muestran en la figura 10.1.

Para los cálculos cuánticos se empleó una técnica DFT con el funcional B3LYP y el conjunto de base 6-311++G(2d,2p). En las simulaciones se incluyó 1 molécula de D-xilopiranososa y 214 de agua TIP3P, las interacciones intermoleculares se truncaron



**Figura 10.1:** Estructuras de la D-xilopiranos estudiadas. Para cada uno de los anómeros,  $\alpha$  y  $\beta$ , se consideran dos confórmeros definidos por la orientación de los enlaces O–H.

a 9,0 Å. Cada simulación de dinámica molecular constó de 25 ps de equilibrado y 100 ps de producción de datos.

Se aplicó en cada caso el método ASEP/MD parcialmente acoplado, con optimización de geometría. Tras cada dinámica molecular se efectuaron cuatro iteraciones de optimización por el método RFO (pág. 105), utilizando la matriz unidad como hessiana inicial y actualizándola mediante la fórmula BFGS modificada propuesta por Prat-Resina y col. [231]. Las cargas atómicas del soluto se calcularon con el método CHELPG.

Una vez calculada la geometría óptima de cada anómero en disolución, se procedió a calcular la estabilidad relativa de cada uno de ellos, para lo cual se empleó el método de perturbación termodinámica. Se introdujeron 19 estados intermedios entre los anómeros  $\alpha$  y  $\beta$ , obtenidos por interpolación lineal. El cálculo se hizo empleando el método de la «topología sencilla», en el que cada estado intermedio tiene una geometría y cargas atómicas interpoladas entre los estados inicial y final. También se realizó un cálculo con «topología dual», en el que cada estado intermedio es una superposición de los estados inicial y final [232, 233]. Al resultado del cálculo FEP se le sumó la diferencia en la energía de la parte cuántica,  $\Delta E_{QM}$ , tal y

como se indica en la ecuación (5.14) (pág. 97):

$$\begin{aligned}\Delta G &= \Delta E_{\text{QM}} + \Delta G_{\text{QM/MM}} \\ E_{\text{QM}} &= \langle \Psi | \hat{H}_{\text{QM}} | \Psi \rangle\end{aligned}\quad (10.1)$$

También se calcularon las diferencias de energía mediante el método PCM incluido en *Gaussian 98*, con un factor de escala para los radios atómicos de 1,2.

## 10.2. Resultados y discusión

Las energías obtenidas para las cuatro estructuras estudiadas, optimizadas en vacío, se presentan en la tabla 10.2. Claramente, la disposición antihoraria de los enlaces O–H es la más estable, por 2,20 kcal/mol en el anómero  $\alpha$  y 3,59 kcal/mol en el anómero  $\beta$ . Se predice que, en vacío, la  $\alpha$ -D-xilopiranososa es ligeramente más estable que la  $\beta$ -D-xilopiranososa, con una diferencia de menos de 1 kcal/mol. Los cálculos en disolución se realizaron únicamente para los confórmeros antihorarios de ambos anómeros.

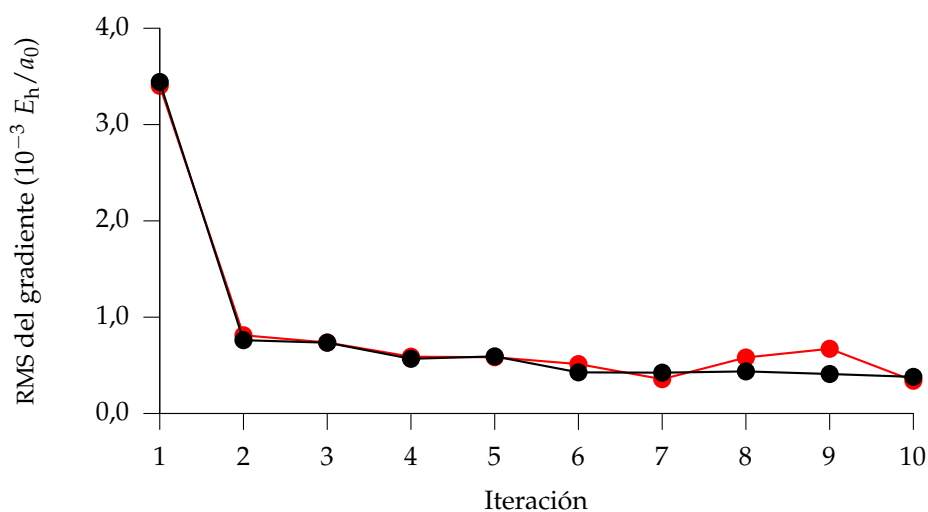
**Tabla 10.2:** Energías en vacío (en kcal/mol) de dos confórmeros de los dos anómeros de la D-xilopiranososa. Las energías son relativas a la estructura más estable.

	Antihorario	Horario
$\alpha$ -D-xilopiranososa	0,00	2,20
$\beta$ -D-xilopiranososa	0,87	4,46

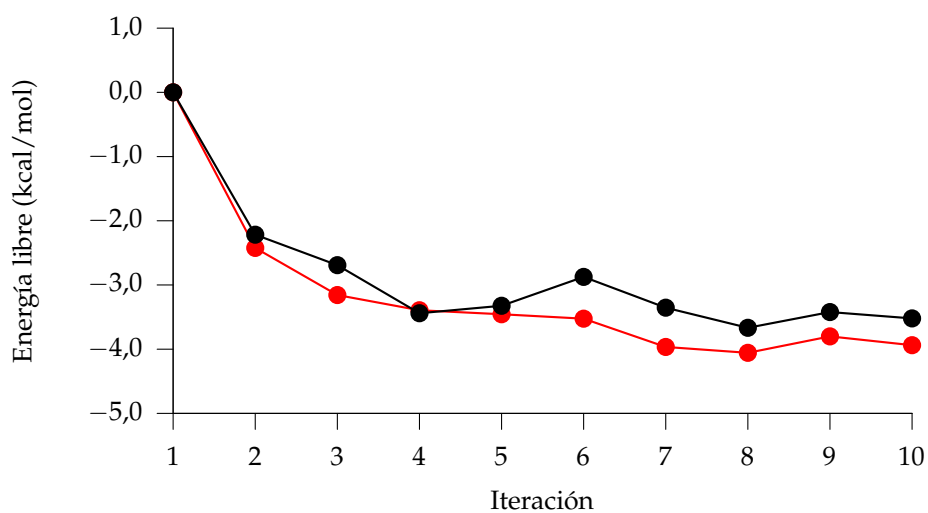
### 10.2.1. Resultados energéticos

La variación del gradiente en el proceso ASEP/MD se muestra en la figura 10.3. En tan sólo 2 o 3 iteraciones la RMS del gradiente toma un valor de entre  $0,0005 E_h/a_0$  y  $0,0010 E_h/a_0$ . La energía libre de polarización de ambos anómeros presenta un comportamiento similar, como se muestra en la figura 10.4. La convergencia se alcanza en 5 o 6 iteraciones. Como resultados finales de las distintas propiedades, tanto para el anómero  $\alpha$  como para el  $\beta$ , se tomaron los promedios de los valores obtenidos en las últimas cinco iteraciones, cuando el sistema se encuentra ya estabilizado.

Para calcular la diferencia de energía libre entre los dos anómeros en disolución se empleó el método de perturbación termodinámica, como se ha mencionado anteriormente. En la tabla 10.5 se dan los resultados, junto con la diferencia de energía obtenida mediante el método PCM. Este método predice que la solvatación



**Figura 10.3:** Evolución de la RMS del gradiente sobre la superficie de energía libre de una molécula de D-xilopiranososa en disolución acuosa durante el proceso de optimización. Se muestran los resultados para la conformación antihoraria de la  $\alpha$ -D-xilopiranososa (negro) y de la  $\beta$ -D-xilopiranososa (rojo).



**Figura 10.4:** Evolución de la energía libre de polarización de la D-xilopiranososa en disolución acuosa durante el proceso de optimización. Se muestran los resultados para la conformación antihoraria de la  $\alpha$ -D-xilopiranososa (negro) y de la  $\beta$ -D-xilopiranososa (rojo).

favorece al anómero  $\alpha$ , dando una energía de solvatación para la  $\alpha$ -D-xilopiranososa 1,1 kcal/mol mayor (en valor absoluto) que para la  $\beta$ -D-xilopiranososa.

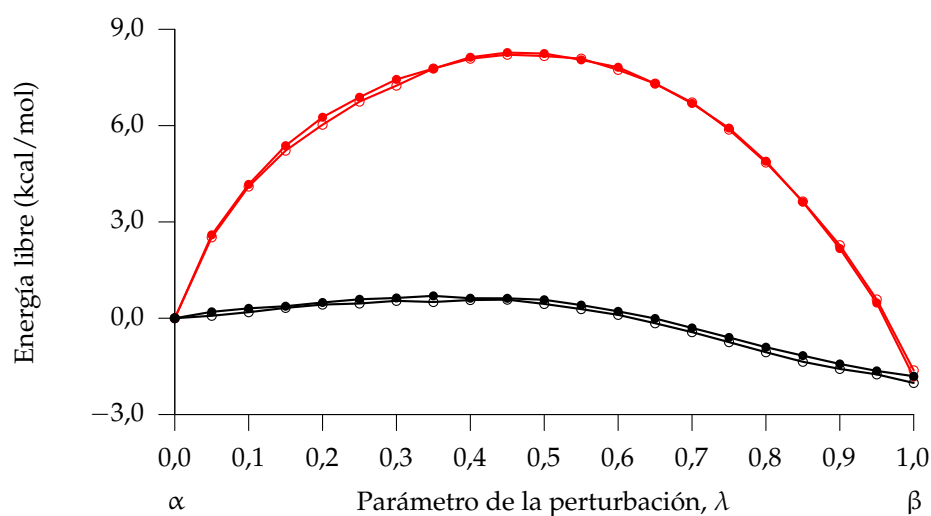
Por el contrario, con el método ASEP/MD se obtiene la tendencia correcta: el anómero  $\beta$  es el más estable en disolución. En el cálculo de perturbación termodinámica se encuentra además un acuerdo muy bueno (una diferencia de sólo 0,2 kcal/mol) entre los cálculos en los dos sentidos posibles, indicativo de la precisión conseguida. Los dos métodos utilizados en este cálculo —topología sencilla y dual— dan resultados muy similares; si se toma el valor medio de ambos métodos, el anómero  $\beta$  es 0,77 kcal/mol más estable que el  $\alpha$ , valor acorde con el resultado experimental de 0,4 kcal/mol [234]. En la figura 10.6 se muestran, gráficamente, los resultados de los cálculos de perturbación termodinámica. Se observa que, a pesar de las distintas técnicas empleadas, el resultado final es independiente del camino seguido, como corresponde a una función de estado.

**Tabla 10.5:** Energía libre en disolución de la  $\beta$ -D-xilopiranososa con respecto a la  $\alpha$ -D-xilopiranososa, en kcal/mol. Para el cálculo de perturbación termodinámica, la transformación se realizó en los dos sentidos (ver figura 10.6).

Método	Topología	$\alpha \rightarrow \beta$	$\beta \rightarrow \alpha$	Media
ASEP/MD	Sencilla	-0,76	-0,97	-0,86
	Dual	-0,78	-0,57	-0,68
PCM				2,0
Experimental [234]				-0,4

A tenor de los resultados, se deduce que, en fase gas, el efecto anomérico estabiliza la forma  $\alpha$ . En disolución, sin embargo, las interacciones con el disolvente favorecen la presencia de la forma  $\beta$ . Esta solvatación preferente del anómero  $\beta$  se ve confirmada por las energías de interacción soluto-disolvente calculadas con el método ASEP/MD. Los resultados se muestran en la tabla 10.7. En ambos anómeros, la energía de interacción está dominada por la componente electrostática, que es 3,0 kcal/mol mayor para la  $\beta$ -D-xilopiranososa, a pesar de que su momento dipolar es casi 0,6 D menor que el del anómero  $\alpha$ . El resto de los términos resulta favorable al anómero  $\alpha$ , reduciéndose finalmente la diferencia en la energía de interacción a 1,8 kcal/mol.

Los resultados de la primera iteración del cálculo ASEP/MD son los obtenidos cuando la estructura del disolvente está en equilibrio con la geometría y distribución de cargas del soluto en vacío; de esta manera se considera el efecto que el disolvente tiene sobre el soluto, pero no el de la polarización mutua entre ambos. Tomando estos valores, la diferencia en la componente electrostática de la energía de interacción es algo menor, 2,4 kcal/mol, aunque el valor total para  $\Delta E^{\text{int}}$  permanece casi constante: 1,6 kcal/mol



**Figura 10.6:** Resultados de los cálculos de perturbación termodinámica entre la  $\alpha$ -D-xilopiranos ( $\lambda = 0$ ) y la  $\beta$ -D-xilopiranos ( $\lambda = 1$ ). La energía libre representada (en kcal/mol) corresponde a  $\Delta G_{\text{QM/MM}}$ . Se realizaron los cálculos con topología sencilla (negro) y dual (rojo), hacia adelante (círculos rellenos) y hacia atrás (círculos vacíos).

**Tabla 10.7:** Momento dipolar (en D) y energías de interacción soluto-disolvente (en kcal/mol), tras la primera iteración y como promedio de las últimas cinco iteraciones del proceso ASEP/MD, para la D-xilopiranos.

	$\alpha$ -D-xilopiranos		$\beta$ -D-xilopiranos	
	1. <sup>a</sup> iteración	Disolución	1. <sup>a</sup> iteración	Disolución
$\mu$	2,15 *	2,95	1,94 *	2,37
$E_{\text{QM/MM}}^{\text{elec}}$	-23,9	-34,1	-26,3	-37,1
$E_{\text{QM/MM}}^{\text{vdw}}$	-11,1	-8,8	-10,8	-7,8
$E^{\text{dist}}$	1,3	2,5	1,5	2,7
$E^{\text{int}}$ **	-33,7	-40,4	-35,6	-42,2

\* Momento dipolar en vacío,  $\mu^0$   
 \*\*  $E^{\text{int}} = E_{\text{QM/MM}}^{\text{elec}} + E_{\text{QM/MM}}^{\text{vdw}} + E^{\text{dist}}$



### 10.2.2. Resultados estructurales

Las distancias y ángulos de enlace optimizados en vacío y en disolución para los dos anómeros se dan en la tabla 10.8. En general, las variaciones de la geometría inducidas por el disolvente son pequeñas. Las distancias C–C y C–H disminuyen en alrededor de 0,005 Å y 0,008 Å, respectivamente. Por el contrario, las distancias C–OH, salvo en el grupo anomérico, aumentan ligeramente, efecto probablemente debido a la participación del O en enlaces de hidrógeno intermoleculares. En cuanto a los ángulos, se observa un incremento de unos 2,5° en los ángulos C–O–H. La mayor diferencia entre los comportamientos de ambos anómeros se encuentra en la distancia C–OH del grupo anomérico: en la  $\beta$ -D-xilopiranososa apenas hay variación con respecto al valor calculado en vacío, pero en la  $\alpha$ -D-xilopiranososa disminuye en casi 0,010 Å.

Las cargas atómicas obtenidas en vacío y en disolución para los dos anómeros se presentan en la tabla 10.9. En general, las distribuciones de carga de las formas  $\alpha$  y  $\beta$  son similares, la mayor diferencia se encuentra en el C anomérico, cuya carga en vacío en el anómero  $\beta$  es casi el doble que en el  $\alpha$ , en disolución esta diferencia se reduce. Al introducirse en el seno de la disolución se produce, principalmente, una polarización de los enlaces O–H, que es mayor en el anómero  $\beta$  que en  $\alpha$ .

En cuanto a la estructura del disolvente, las funciones de distribución radial entre el O anomérico del soluto y el O del agua se muestran en la figura 10.10. Se observa que el primer pico es más alto en el anómero  $\beta$  que en el  $\alpha$ , señal de la mejor solvatación de la forma  $\beta$ . Este resultado contrasta con el descrito recientemente para simulaciones de dinámica molecular *ab initio* de la glucosa [235], donde se encontró que la RDF del anómero  $\beta$  era más ancha y baja que la del  $\alpha$ . El análisis de los números de coordinación también confirma la solvatación preferente del anómero  $\beta$ : se obtiene un número de coordinación de 4,43 para la  $\alpha$ -D-xilopiranososa y de 5,43 para la  $\beta$ -D-xilopiranososa.

La RDF entre el O<sup>8</sup> y el O del agua se representa en la figura 10.11. En este caso la diferencia entre los dos anómeros es mínima, aunque el primer pico es ligeramente más alto en el anómero  $\alpha$ . Es de destacar que existen diferencias significativas entre estas RDF y las proporcionadas por Liu y Brady [236, 237], cuyos picos son más pronunciados; los resultados obtenidos con nuestro método son más similares a los publicados por Molteni y Parrinello [235]. Los números de coordinación en este caso son de 3,46 para ambas formas. El resto de los grupos hidroxilo tienen RDF similares.

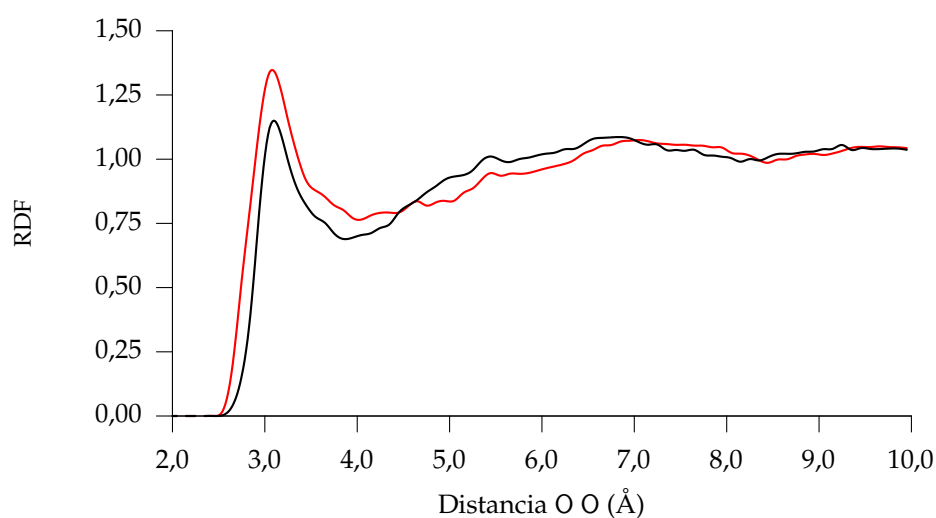
A partir de estos datos, parece claro que la estabilización relativa del anómero  $\beta$  está relacionada con la mejor solvatación del grupo OH anomérico en esta forma, la influencia del resto de la molécula tiene menor importancia.

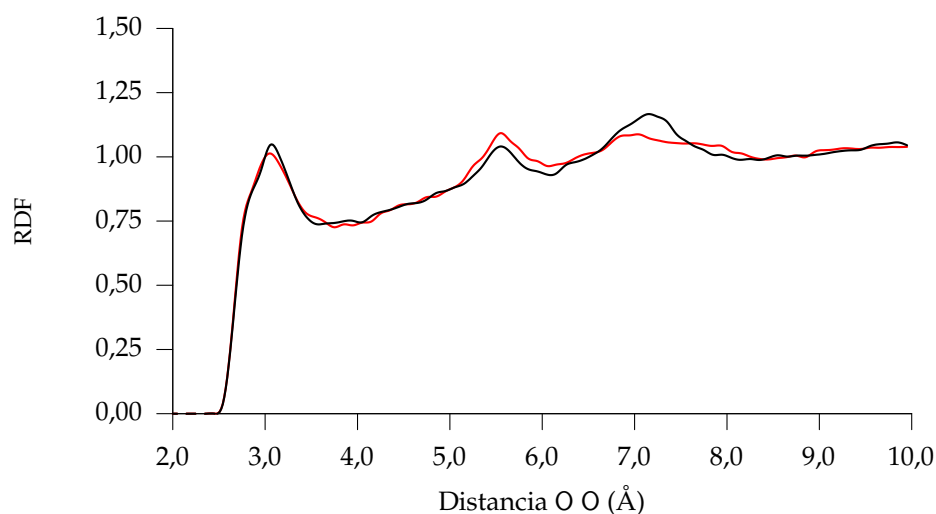
**Tabla 10.8:** Geometrías optimizadas, en vacío y en disolución, para las conformaciones antihorarias de los dos anómeros de la D-xilopiranos. Distancias en Å, ángulos en grados.

	$\alpha$ -D-xilopiranos		$\beta$ -D-xilopiranos	
	Vacío	Disolución	Vacío	Disolución
C <sup>1</sup> -C <sup>2</sup>	1,536	1,529	1,530	1,526
C <sup>2</sup> -C <sup>3</sup>	1,531	1,527	1,523	1,521
C <sup>3</sup> -C <sup>4</sup>	1,524	1,519	1,527	1,522
C <sup>4</sup> -C <sup>5</sup>	1,528	1,523	1,532	1,525
C <sup>5</sup> -O <sup>10</sup>	1,436	1,433	1,427	1,428
C <sup>1</sup> -O <sup>10</sup>	1,407	1,415	1,419	1,421
C <sup>1</sup> -O <sup>6</sup>	1,419	1,408	1,395	1,394
C <sup>2</sup> -O <sup>7</sup>	1,421	1,422	1,420	1,423
C <sup>3</sup> -O <sup>8</sup>	1,421	1,425	1,420	1,423
C <sup>4</sup> -O <sup>9</sup>	1,417	1,425	1,416	1,421
C <sup>1</sup> -H <sup>12</sup>	1,097	1,088	1,107	1,095
C <sup>2</sup> -H <sup>13</sup>	1,095	1,089	1,100	1,093
C <sup>3</sup> -H <sup>14</sup>	1,102	1,093	1,103	1,095
C <sup>4</sup> -H <sup>15</sup>	1,103	1,093	1,102	1,093
C <sup>5</sup> -H <sup>11</sup>	1,092	1,085	1,092	1,085
C <sup>5</sup> -H <sup>16</sup>	1,097	1,091	1,102	1,094
O <sup>6</sup> -H <sup>17</sup>	0,966	0,964	0,967	0,965
O <sup>7</sup> -H <sup>18</sup>	0,970	0,966	0,967	0,964
O <sup>8</sup> -H <sup>19</sup>	0,968	0,965	0,968	0,964
O <sup>9</sup> -H <sup>20</sup>	0,968	0,965	0,968	0,965
C <sup>2</sup> -C <sup>1</sup> -O <sup>10</sup>	112,1	110,8	110,4	109,9
C <sup>1</sup> -C <sup>2</sup> -C <sup>3</sup>	110,4	111,1	109,4	110,0
C <sup>2</sup> -C <sup>3</sup> -C <sup>4</sup>	110,3	109,9	110,4	110,4
C <sup>3</sup> -C <sup>4</sup> -C <sup>5</sup>	108,7	109,3	109,3	109,8
C <sup>4</sup> -C <sup>5</sup> -O <sup>10</sup>	110,9	110,6	110,7	110,2
C <sup>1</sup> -O <sup>10</sup> -C <sup>5</sup>	114,9	114,4	112,4	112,6
O <sup>6</sup> -C <sup>1</sup> -H <sup>12</sup>	110,6	110,9	111,4	110,9
O <sup>7</sup> -C <sup>2</sup> -H <sup>13</sup>	106,5	106,3	111,3	110,3
O <sup>8</sup> -C <sup>3</sup> -H <sup>14</sup>	110,4	109,6	110,2	109,5
O <sup>9</sup> -C <sup>4</sup> -H <sup>15</sup>	111,0	109,8	111,1	110,1
H <sup>11</sup> -C <sup>5</sup> -H <sup>16</sup>	108,7	108,9	108,2	108,6
C <sup>1</sup> -O <sup>6</sup> -H <sup>17</sup>	108,2	109,5	108,2	109,7
C <sup>2</sup> -O <sup>7</sup> -H <sup>18</sup>	105,6	108,1	106,9	109,0
C <sup>3</sup> -O <sup>8</sup> -H <sup>19</sup>	106,3	108,7	106,7	109,3
C <sup>4</sup> -O <sup>9</sup> -H <sup>20</sup>	106,1	108,9	106,1	109,1

**Tabla 10.9:** Cargas atómicas (en  $e$ ), en vacío y en disolución, para los dos anómeros de la D-xilopiranososa.

	$\alpha$ -D-xilopiranososa		$\beta$ -D-xilopiranososa	
	Vacío	Disolución	Vacío	Disolución
C <sup>1</sup>	0,233	0,343	0,464	0,474
C <sup>2</sup>	0,262	0,200	0,188	0,189
C <sup>3</sup>	0,132	0,293	0,199	0,298
C <sup>4</sup>	0,348	0,293	0,324	0,321
C <sup>5</sup>	0,169	0,237	0,176	0,176
O <sup>6</sup>	-0,641	-0,687	-0,659	-0,742
O <sup>7</sup>	-0,627	-0,711	-0,679	-0,803
O <sup>8</sup>	-0,670	-0,783	-0,704	-0,805
O <sup>9</sup>	-0,709	-0,836	-0,714	-0,833
O <sup>10</sup>	-0,480	-0,573	-0,517	-0,572
H <sup>11</sup>	0,069	0,066	0,076	0,093
H <sup>12</sup>	0,080	0,078	-0,016	0,029
H <sup>13</sup>	0,057	0,073	0,059	0,081
H <sup>14</sup>	0,051	0,049	0,029	0,030
H <sup>15</sup>	-0,007	0,029	0,004	0,028
H <sup>16</sup>	0,042	0,039	0,015	0,035
H <sup>17</sup>	0,442	0,480	0,429	0,489
H <sup>18</sup>	0,385	0,429	0,425	0,496
H <sup>19</sup>	0,422	0,476	0,454	0,504
H <sup>20</sup>	0,440	0,503	0,446	0,510

**Figura 10.10:** Función de distribución radial entre el O anomérico de la D-xilopiranososa y el O del agua para el anómero  $\alpha$  (negro) y el anómero  $\beta$  (rojo).



**Figura 10.11:** Función de distribución radial entre el O<sup>8</sup> de la D-xilopiranos y el O del agua para el anómero  $\alpha$  (negro) y el anómero  $\beta$  (rojo).

### 10.3. Conclusiones

El equilibrio anomérico en de la D-xilopiranos en vacío está dominado por el efecto anomérico, que estabiliza a la forma  $\alpha$ ; en disolución, sin embargo, se favorece la forma  $\beta$ . Frente a los estudios realizados por otros autores, en nuestro caso el método empleado en asigna cargas diferentes a los dos anómeros y permite la polarización del soluto en disolución.

Las conclusiones particulares más importantes que pueden extraerse son las siguientes:

- La distribución de carga de la D-xilopiranos sufre un cambio importante en disolución acuosa. Éste cambio es mayor para el anómero  $\alpha$ , lo que, junto con el mayor momento dipolar en vacío, explica que el método PCM prediga una estabilización adicional de la  $\alpha$ -D-xilopiranos en disolución.
- La estabilización de la  $\beta$ -D-xilopiranos en disolución se debe, sobre todo, a la solvatación preferente del grupo O–H anomérico, como se desprende de las RDF, la cual, a su vez, está originada por la diferencia en las cargas atómicas sobre el C y el O anoméricos y por la diferente disposición espacial de los grupos OH.

---

---

# 11

---

## BARRERA DE ACTIVACIÓN DE LA REACCIÓN DE MENSCHUTKIN

La determinación de puntos de ensilladura en la superficie de energía libre de reacciones en disolución es aún una cuestión abierta en la química computacional. En disolución, un buen modelo debe combinar descripciones precisas tanto de la reacción química propiamente dicha como de las interacciones que tienen lugar entre soluto y disolvente.

En la mayor parte de los casos la descripción cuántica completa —de tipo Car-Parrinello— del sistema soluto-disolvente resulta poco práctica o simplemente inviable, debido al gran número de moléculas de disolvente (varios cientos) y de configuraciones (varios miles) que es necesario considerar para obtener resultados estadísticamente significativos. El problema se complica cuando se trata de determinar la variación de la energía libre a lo largo de un camino de reacción, es decir, el potencial de fuerza media. En este último caso el número de cálculos necesarios aumenta drásticamente, de manera que una descripción cuántica completa del sistema es imposible salvo para los casos más sencillos.

Por estos motivos, en el estudio teórico de las reacciones químicas en disolución, es imprescindible introducir aproximaciones en la descripción del soluto o en la del disolvente. Esto hace que los métodos más comúnmente empleados sean los llamados métodos mixtos, cuyas características y diferentes variantes ya han sido detalladas. Este capítulo está dedicado, en concreto, al empleo del método ASEP/MD para el estudio de reacciones en disolución. Como ya se ha mencionado, el método ASEP/MD permite emplear niveles de cálculo cuántico similares a los utilizados en vacío, a la vez que se trata con detalle la estructura del disolvente alrededor del soluto y las interacciones entre ambos.

El procedimiento desarrollado para la optimización de geometrías mediante el método ASEP/MD puede también aplicarse para la localización de puntos de ensilladura o estados de transición. Una vez que se dispone del gradiente y la hessiana sobre la superficie de energía libre, basta con emplear una técnica de búsqueda de puntos de ensilladura en lugar de mínimos. Tras caracterizar con este método el estado de transición de una reacción, la energía libre de activación puede calcularse —en el contexto de la «Teoría del Estado de Transición» (*Transition State Theory*, TST) [238]— hallando la diferencia de energía libre entre el estado de transición y los reactivos, previamente optimizados en disolución.

Como aplicación del método y algoritmos desarrollados, se realizó un estudio del estado de transición y los reactivos de la reacción de Menshutkin [239, 240]:  $\text{NH}_3 + \text{CH}_3\text{Cl} \rightarrow \text{CH}_3\text{NH}_3^{\oplus} + \text{Cl}^{\ominus}$ . Se trata de un tipo especial de reacción  $\text{S}_{\text{N}}2$  en la que los reactivos son neutros y, en su transcurso, se produce una separación de cargas. Este proceso es muy desfavorable en vacío; sin embargo, los disolventes polares estabilizan los productos y reducen la barrera energética. Esta reacción ha sido estudiada en disolución con modelos del continuo [241-243], métodos QM/MM [158, 198] y RISM-SCF [244], lo que permite comparar los resultados obtenidos con diferentes procedimientos.

Cálculos previos con modelos de continuo han empleado distintos niveles de cálculo cuántico: HF, MP2, MCSCF, MP4 y DFT. Tomando como referencia el resultado obtenido con MP4, Truong y col. [241] llegaron a la conclusión de que el método más eficiente es DFT, que proporciona incluso mejores resultados que MP2 y MCSCF. En los cálculos QM/MM se empleó el método semiempírico AM1 para describir el soluto. Así, Gao y Xia [198] construyeron una superficie bidimensional en la que localizaron el estado de transición, mientras que Hirao y col. [158] obtuvieron el estado de transición por medio de una variante del método del gradiente de la energía libre. Naka y col. [244] estimaron el perfil de energía libre para la reacción, asumiendo una coordenada de reacción distinguida, con el método RISM-MP2.

La aplicación del método ASEP/MD al estudio de los estados de transición permite considerar y analizar la estructura del disolvente alrededor del soluto. Por otro lado, gracias al empleo de la ACM junto con el método FEG, la localización de los estados de transición no depende de la definición de una coordenada de reacción específica, sino que se localizan de manera directa verdaderos puntos de ensilladura en la superficie de energía libre.

### 11.1. Detalles de los cálculos

Se llevaron a cabo cálculos en disolución tanto para los reactivos por separado,  $\text{NH}_3$  y  $\text{CH}_3\text{Cl}$ , como para el estado de transición,  $(\text{NH}_3-\text{CH}_3-\text{Cl})^{\ddagger}$ . Para los cálculos

cuánticos se empleó un método DFT con el funcional BH&HLYP y el conjunto de base aug-cc-pVDZ. En las simulaciones de dinámica molecular se rodeó el soluto de 214 moléculas de agua y las interacciones intermoleculares se truncaron a 9,0 Å.

En los cálculos ASEP/MD se tomó como geometría inicial la optimizada en vacío. Para las optimizaciones en disolución se empleó el método P-RFO, realizando una optimización completa de la geometría tras cada simulación de dinámica molecular. Las cargas atómicas del soluto se calcularon con el método CHELPG.

La diferencia de energía libre entre reactivos y estado de transición en vacío se calculó como diferencia entre las energías potenciales,  $\Delta^\ddagger E_{\text{QM}}$ , añadiendo la contribución a la energía libre de la energía del punto cero y de los efectos térmicos,  $\Delta^\ddagger V$ , para los que se aplicó la aproximación armónica a la función de partición vibracional [245]:

$$\Delta^\ddagger G^{\text{vac}} = \Delta^\ddagger E_{\text{QM}}^{\text{vac}} + \Delta^\ddagger V^{\text{vac}} \quad (11.1)$$

La diferencia de energía libre en disolución se calculó mediante FEP. Se consideraron los reactivos separados una distancia de 6,0 Å, los estados intermedios se obtuvieron por interpolación lineal de la geometría, las cargas atómicas y los parámetros de Lennard-Jones. Se introdujeron 39 estados intermedios ( $\Delta\lambda = 0,025$ ). Del cálculo de FEP se obtuvo el término  $\Delta^\ddagger G_{\text{QM/MM}}$ . También se calculó el término  $\Delta^\ddagger V^{\text{dis}}$ , con las mismas expresiones que en vacío, pero empleando las frecuencias vibracionales calculadas en disolución; ésta es, además, una forma aproximada de incluir la contribución a la energía libre de las fluctuaciones del soluto (ver ec. C.9, pág. 199). La diferencia de energía libre en disolución se calcula entonces como:

$$\Delta^\ddagger G^{\text{dis}} = \Delta^\ddagger E_{\text{QM}}^{\text{dis}} + \Delta^\ddagger G_{\text{QM/MM}} + \Delta^\ddagger V^{\text{dis}} \quad (11.2)$$

Conviene destacar que el cálculo de  $\Delta^\ddagger G^{\text{dis}}$  se realiza una vez que se han hallado las estructuras óptimas para los reactivos y estado de transición. Además, la trayectoria seguida en el cálculo FEP es completamente arbitraria y no proporciona ninguna información sobre el camino de reacción o sobre el perfil de energía libre a lo largo de éste.

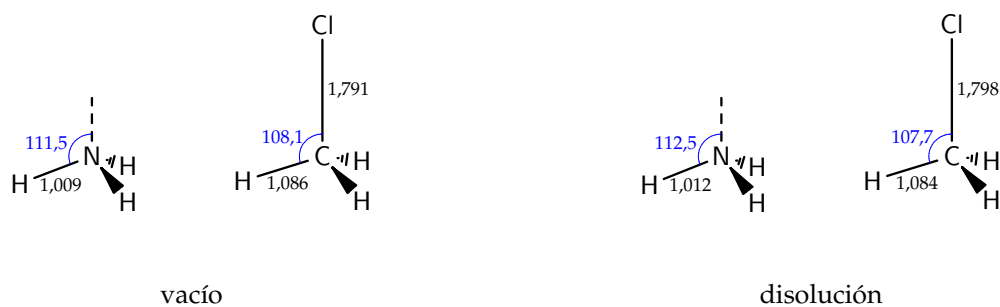
## 11.2. Resultados y discusión

Se optimizaron, en vacío y en disolución, los dos reactivos por separado. También se localizó el estado de transición en vacío y en disolución. Las cargas atómicas, momentos dipolares y energías de interacción soluto-disolvente de los tres sistemas se dan en la tabla 11.1. Las correspondientes geometrías, en vacío y en disolución, se muestran en las figuras 11.2 y 11.3. Los datos en disolución corresponden al promedio de las últimas cinco iteraciones del proceso ASEP/MD.

**Tabla 11.1:** Energías de interacción soluto-disolvente (en kcal/mol), momentos dipolares (en D) y cargas atómicas (en  $e$ ) para reactivos y estado de transición de la reacción de Menshutkin, en vacío y en disolución.

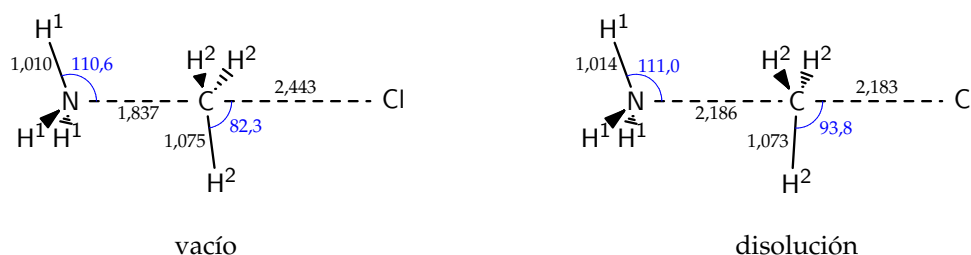
	Vacío	Disolución		Vacío	Disolución
NH <sub>3</sub>			(NH <sub>3</sub> -CH <sub>3</sub> -Cl) <sup>‡</sup>		
$E_{\text{QM/MM}}$		-15,47	$E_{\text{QM/MM}}$		-51,01
$\mu$	1,54	2,22	$\mu$	12,48	11,09
$q(\text{N})$	-0,898	-1,185	$q(\text{N})$	-0,361	-0,704
$q(\text{H})$	0,299	0,395	$q(\text{H}^1)$	0,239	0,298
CH <sub>3</sub> Cl			$q(\text{C})$	-0,022	-0,100
$E_{\text{QM/MM}}$		-9,36	$q(\text{H}^2)$	0,126	0,198
$\mu$	2,07	2,66	$q(\text{Cl})$	-0,710	-0,686
$q(\text{C})$	-0,162	-0,106			
$q(\text{H})$	0,115	0,118			
$q(\text{Cl})$	-0,183	-0,247			

El cálculo en vacío sobrestima en 0,18 D el momento dipolar experimental del CH<sub>3</sub>Cl, que es de 1,89 D [246]. La molécula de NH<sub>3</sub>, sin embargo, se describe mejor en este sentido, ya que el valor experimental del momento dipolar es de entre 1,53 D y 1,56 D [247, 248]. Estos resultados son ligeramente mejores que los obtenidos por Truong y col. [241], quienes emplearon un conjunto más reducido de funciones de base.

**Figura 11.2:** Geometrías en vacío y en disolución de los reactivos de la reacción de Menshutkin, NH<sub>3</sub> y CH<sub>3</sub>Cl. Distancias en Å, ángulos en grados.

El efecto que la solvatación tiene sobre los reactivos es apreciable. Los momentos dipolares del CH<sub>3</sub>Cl y del NH<sub>3</sub> aumentan en un 28 % y un 44 % respectivamente. La carga sobre los átomos de H del CH<sub>3</sub>Cl apenas varía, por lo que el cambio en el momento dipolar está asociado casi exclusivamente a la polarización del enlace C-Cl, cuya distancia aumenta en 0,007 Å. El ángulo Cl-C-H se reduce en 0,4°.





**Figura 11.3:** Geometrías en vacío y en disolución del estado de transición de la reacción de Menshutkin. Distancias en Å, ángulos en grados.

En el  $\text{NH}_3$  se observa un alargamiento de los enlaces  $\text{N-H}$  en  $0,003 \text{ \AA}$  y una reducción del ángulo  $\text{H-N-H}$  en alrededor de  $1,0^\circ$ . Las desviaciones típicas de estos valores (correspondientes al promedio de las últimas cinco iteraciones) son mínimas, menores de  $0,001 \text{ \AA}$ ,  $0,1^\circ$  y  $0,05 \text{ D}$  para distancias de enlace, ángulos de enlace y momentos dipolares, respectivamente. Esto indica que la duración de las simulaciones ( $50 \text{ ps}$ ) es suficiente y que las fluctuaciones asociadas a su tamaño finito tienen un efecto despreciable en los parámetros geométricos.

El efecto del disolvente sobre la estructura del estado de transición es más importante. La distancia  $\text{C-Cl}$  disminuye en  $0,260 \text{ \AA}$  y la distancia  $\text{C-N}$  aumenta en  $0,349 \text{ \AA}$ . El ángulo  $\text{Cl-C-H}^2$  aumenta en  $11,5^\circ$ , pasando de ser menor de  $90^\circ$  a ser mayor. La carga positiva del grupo  $\text{NH}_3$  disminuye en  $0,166 e$ , mientras que la del grupo  $\text{CH}_3$  aumenta en  $0,138 e$ , lo que implica una reducción en el momento dipolar; sin embargo, dentro de cada grupo la polarización se intensifica. Todos los valores anteriores indican que el estado de transición es más simétrico en disolución que en vacío. Dado que en esta reacción se produce una separación de carga, los resultados indican que, en disolución, el estado de transición es más temprano, lo que está de acuerdo con el postulado de Hammond [249].

Los resultados energéticos se presentan en la tabla 11.4. El valor de la energía libre de activación obtenido en vacío,  $44,9 \text{ kcal/mol}$  a  $298 \text{ K}$ , está en buen acuerdo con el de Truong y col. [241],  $45,7 \text{ kcal/mol}$ , como era de esperar dada la similitud en el nivel de cálculo. También existe una buena coincidencia con el resultado MP4SDTQ/6-31+G\* obtenido por Gao y Xia [198, 250],  $46,7 \text{ kcal/mol}$ .

En este sistema, la solvatación hace que disminuya drásticamente la energía libre de activación. El valor en disolución obtenido es de  $25,6 \text{ kcal/mol}$ , resultado que está entre los proporcionados por GCOSMO [241] ( $24,8 \text{ kcal/mol}$ ) y AM1/MM (entre  $26,3 \text{ kcal/mol}$  [198] y  $27,1 \text{ kcal/mol}$  [158]). Amovilli y col. [242], empleando el modelo PCM y una base 6-311G\*\*, obtuvieron  $16,8 \text{ kcal/mol}$  con el método Hartree-Fock y  $20,5 \text{ kcal/mol}$  con el CASSCF. Un valor similar a este último,  $20,9 \text{ kcal/mol}$ , fue el que Naka y col. [244] obtuvieron con el método RISM-MP2. No hay datos experimentales para este sistema, pero sí existen para la reacción entre

el  $\text{NH}_3$  y el  $\text{CH}_3$ ], en este caso la energía de activación experimental en disolución acuosa es 23,5 kcal/mol [251].

**Tabla 11.4:** Componentes de la energía libre de activación a 298 K (ver ec. 11.1 y 11.2 para la reacción de Menshutkin, en kcal/mol. También se dan las diferencias entre las energías de interacción soluto-disolvente del estado de transición y de los reactivos.

	Vacío	Disolución		Disolución
$\Delta^\ddagger E_{\text{QM}}$	32,70	27,48	$\Delta^\ddagger E_{\text{QM/MM}}^{\text{elec}}$	-30,50
$\Delta^\ddagger V$	12,23	11,86	$\Delta^\ddagger E_{\text{QM/MM}}^{\text{vdw}}$	4,32
$\Delta^\ddagger G_{\text{QM/MM}}$		-13,70	$\Delta^\ddagger E_{\text{QM/MM}}$	-26,18
$\Delta^\ddagger G$	44,93	25,64		

Las diferencias entre los valores obtenidos con los modelos del continuo GCOSMO y PCM son debidas en parte al distinto criterio para la elección del radio de la cavidad en la que se sitúa el soluto. En el GCOSMO, los radios atómicos se ajustan para reproducir las energías libres de hidratación de un conjunto representativo de moléculas pequeñas e iones. En el PCM, suelen tomarse como 1,2 veces el radio de Van der Waals. Debido a esta importante diferencia, la comparación directa con los métodos QM/MM no es inmediata. En general, los modelos del continuo conducen a menores energías libres de activación que los métodos QM/MM. Sin embargo, las diferencias son pequeñas y, como ya se ha hecho anteriormente, se puede concluir que el principal factor de la estabilización del estado de transición es la interacción electrostática no específica.

En el cálculo ASEP/MD se encuentran incluidas las correcciones entrópicas a la energía que libre, mientras que estas correcciones se despreciaron en los cálculos AM1/MM. Sin embargo, Gao y Xia mostraron que los resultados del método AM1 sin corrección entrópica son una buena aproximación para la energía MP4SDTQ/6-31+G\*. Pese a la diferencia en el tipo de cálculo cuántico, se observa una buena concordancia entre los resultados aquí presentados y los de los cálculos AM1/MM. Por lo tanto, considerando como válidos estos últimos resultados, se puede concluir que las aproximaciones existentes en el método ASEP/MD no introducen errores significativos en la energía libre de activación de este sistema.

En todos los casos se produce una disminución muy importante de la energía libre de activación al pasar de la fase gaseosa a la disolución acuosa. Se observa además que la principal contribución a esta disminución proviene de la diferencia en la energía libre de solvatación,  $\Delta^\ddagger G_{\text{QM/MM}}$ , entre los reactivos y el estado de transición, lo que indica que este último está mejor solvatado que los reactivos. En la tabla 11.4 también se muestra la diferencia de energía de interacción soluto-disol-

vente, así como sus componentes electrostática y de Van der Waals, entre reactivos y estados de transición, calculada para las respectivas estructuras optimizadas. Se puede apreciar que la componente más importante es la interacción electrostática,  $-30,50$  kcal/mol; de hecho, la componente de Van der Waals es positiva, estabiliza a los reactivos con respecto al estado de transición.

Esta energía de interacción no incluye la energía necesaria para polarizar el disolvente (en este caso, dado que se simularon moléculas no polarizables, se trata de una polarización de su estructura, no de una polarización electrónica). En la teoría electrostática clásica, la energía de polarización puede calcularse como  $-1/2$  de la energía de interacción. Al añadir esta contribución se obtiene una estabilización neta de  $-13,09$  kcal/mol, valor muy cercano al de  $\Delta^\ddagger G_{\text{QM/MM}}$  calculado mediante FEP,  $-13,70$  kcal/mol.

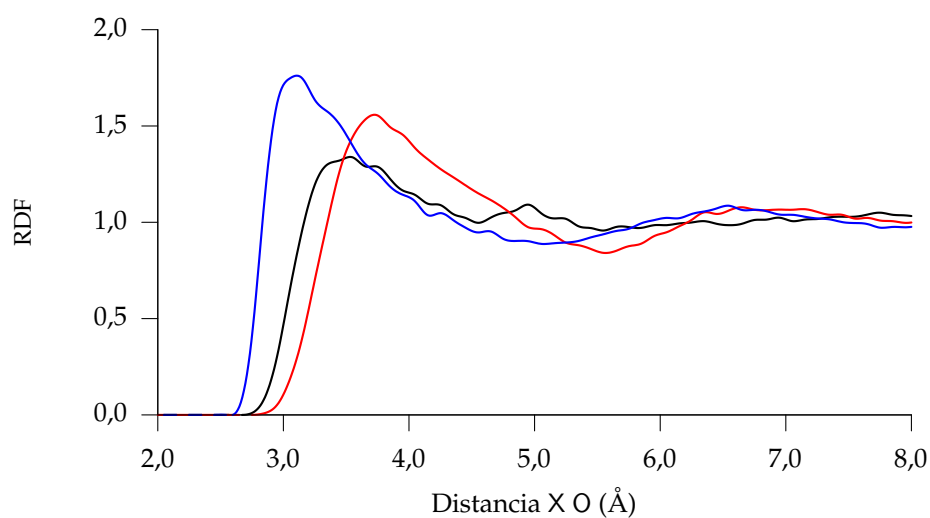
La diferente solvatación de reactivos y productos se pone también de manifiesto en las funciones de distribución radial. En la figura 11.5 se muestran las RDF entre los átomos pesados de los reactivos (N, C y Cl) y el O del agua. Asimismo, en la figura 11.6 se representan las correspondientes RDF para el estado de transición. La diferencia más destacable entre ambas es el desplazamiento que sufre la posición del primer pico de las RDF del Cl y del N, que se encuentra a menores distancias en el estado de transición que en los reactivos. Se observa también un estrechamiento de dichos picos y un aumento de la altura del correspondiente al Cl. Todos estos efectos apuntan a que la interacción con el disolvente es mayor en el estado de transición que en los reactivos, de acuerdo con los datos energéticos. Las RDF obtenidas son, además, muy similares a las publicadas por otros autores [158].

### 11.3. Conclusiones

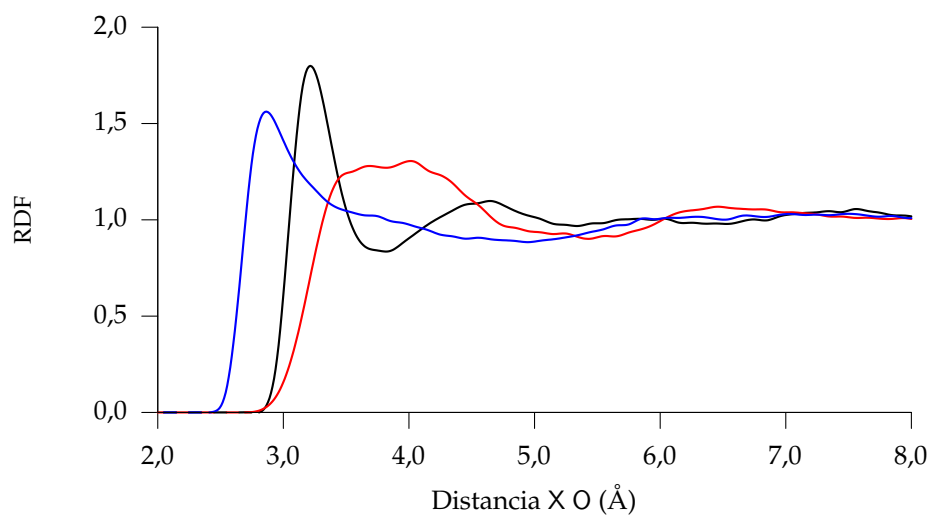
El método desarrollado para la optimización de geometrías en disolución puede adaptarse fácilmente a la localización de estados de transición. Además, el empleo de las técnicas de FEP permite obtener, de manera relativamente rápida, la diferencia de energía libre entre distintas estructuras y, por consiguiente, la energía libre de activación.

Del estudio realizado para el caso de la reacción de Menshutkin entre  $\text{NH}_3$  y  $\text{CH}_3\text{Cl}$  pueden extraerse las siguientes conclusiones particulares:

- La aproximación del campo medio no introduce errores significativos en las magnitudes estudiadas, mientras que permite reducir el número de cálculos cuánticos y simplifica la exploración de la superficie de energía potencial.
- Tanto los resultados energéticos como los estructurales muestran un buen acuerdo con otros estudios que emplean métodos diferentes y son coherentes con lo esperable según el postulado de Hammond.



**Figura 11.5:** Funciones de distribución radial entre los átomos pesados de los reactivos de la reacción de Menshutkin y el O del agua. Se muestran las RDF para el Cl (negro), C (rojo) y N (azul).



**Figura 11.6:** Funciones de distribución radial entre los átomos pesados del estado de transición de la reacción de Menshutkin y el O del agua. Se muestran las RDF para el Cl (negro), C (rojo) y N (azul).

- De estos resultados se deduce que, en este caso, los factores electrostáticos no específicos son los principales responsables de la disminución de la energía libre de activación al pasar de la fase gaseosa a la disolución.



---

---

# 12

---

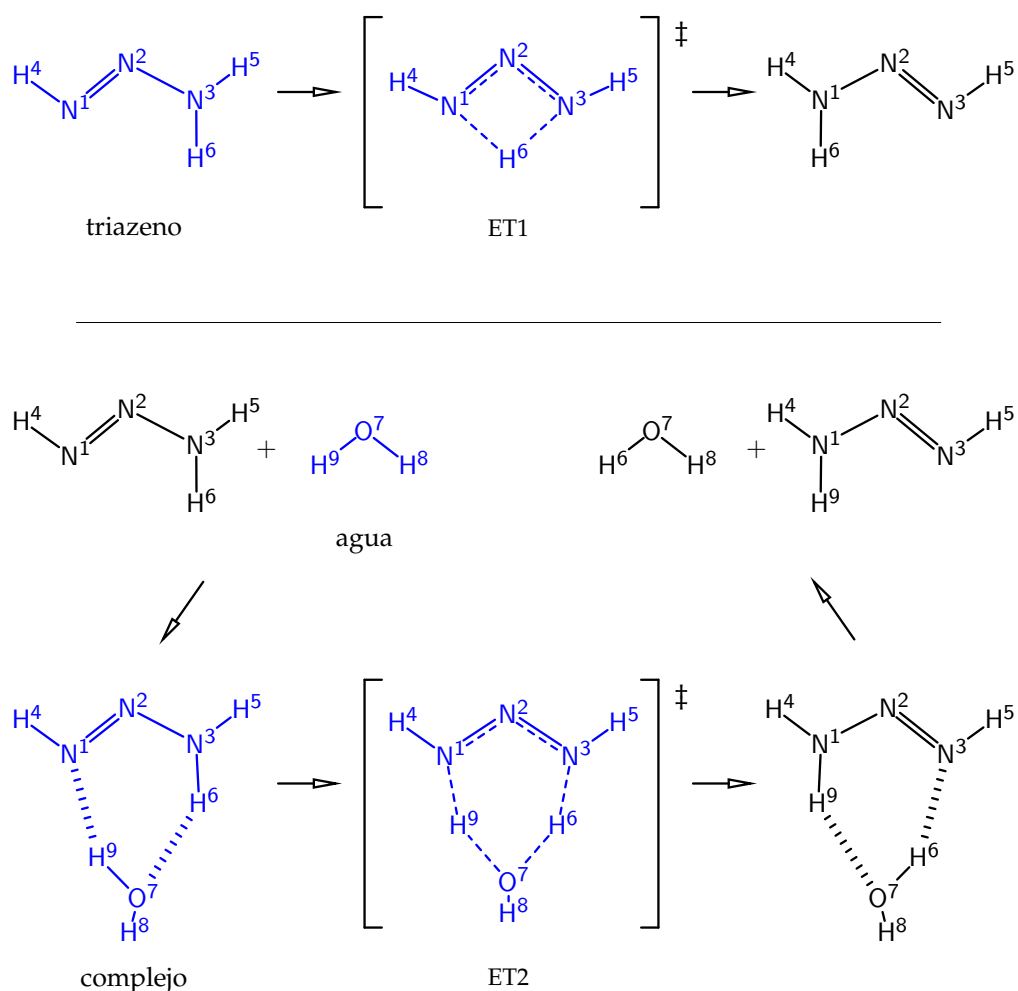
## DESPLAZAMIENTO PROTÓNICO 1,3 EN EL TRIAZENO

Los triazenos son compuestos caracterizados por contener un grupo diazoamino ( $\text{RN}=\text{N}-\text{NR}'\text{R}''$ ), el miembro más sencillo de la familia ( $\text{R} = \text{R}' = \text{R}'' = \text{H}$ ) es conocido simplemente como triazeno. Estos compuestos han sido utilizados como grupos protectores en síntesis orgánica, en química combinatoria, en síntesis de polímeros y como fármacos antitumorales [252]. Su actividad biológica se debe a que, en su descomposición, generan sales de diazonio que pueden producir alquilaciones en el ADN. Normalmente, los triazenos adoptan la configuración *trans* en el doble enlace  $\text{N}=\text{N}$ , pero pueden sufrir isomerizaciones reversibles termo y fotoinducidas [253-255], lo que los convierte en compuestos interesantes como posibles dispositivos para la electrónica molecular.

Existen diversos estudios experimentales [253-259] y teóricos [260-264] sobre la estabilidad, equilibrio y descomposición de triazenos con distintos sustituyentes. Una conclusión general es que la transferencia protónica entre el triazeno y el disolvente juega un papel muy importante tanto en los procesos de descomposición como en los de isomerización. Los triazenos pueden existir en dos formas tautoméricas ( $\text{RN}=\text{N}-\text{NHR}' \rightleftharpoons \text{RHN}-\text{N}=\text{NR}'$ ) y a través de procesos de protonación y desprotonación pueden dar lugar a isomerizaciones *cis-trans* y distintos productos de descomposición.

Dado el interés y la versatilidad de los triazenos y la escasez, por no decir ausencia, de estudios teóricos que incluyan explícitamente el efecto del disolvente, se realizó un análisis de la reacción de desplazamiento protónico 1,3 en el *trans*-triazeno ( $\text{HN}=\text{N}-\text{NH}_2 \rightleftharpoons \text{H}_2\text{N}-\text{N}=\text{NH}$ ) en vacío y en disolución acuosa. Esta tautomería es la base de algunos procesos de isomerización y parece probable que esté

favorecida por el disolvente. Se estudiaron dos procesos (figura 12.1): un proceso unimolecular en el que uno de los protones se desplaza del N<sup>3</sup> al N<sup>1</sup> y un proceso bimolecular en el que, tras la formación de un complejo triazeno-agua, se produce un intercambio protónico entre ambos.



**Figura 12.1:** Los dos mecanismos estudiados para el desplazamiento protónico 1,3 en el *trans*-triazeno: mecanismo unimolecular (arriba) y mecanismo bimolecular (abajo). En azul, las cinco especies estudiadas; las estructuras en negro son equivalentes a alguna de las anteriores.

## 12.1. Método empleado

En los capítulos previos se ha mostrado la utilidad del método ASEP/MD para la obtención de propiedades promedio de moléculas en disolución con un reducido



coste computacional, tanto de estructuras estables como de transición. No obstante, debido precisamente al empleo de la aproximación del campo medio, este método no permite estudiar de manera detallada los movimientos moleculares individuales que tienen lugar durante el transcurso de una reacción química. Los métodos de dinámica QM/MM (pág. 66), por su parte, resultan mucho más costosos en tiempo de cálculo, pero permiten analizar el comportamiento de moléculas y procesos individuales en lugar de las propiedades promedio. Puesto que ambos métodos tienen sus ventajas e inconvenientes, que hasta cierto punto resultan complementarios, se ha intentado combinar los dos procedimientos en este estudio.

En primer lugar, se aplicó el método ASEP/MD a las especies químicas implicadas: triazeno, estado de transición unimolecular (ET1), agua, complejo triazeno-agua y estado de transición bimolecular (ET2), mostradas en la figura 12.1. Se obtuvieron así las geometrías optimizadas en disolución y las diferencias de energía libre entre las distintas especies. Se realizó también una simulación de la molécula de triazeno en disolución acuosa mediante el método DFT/MM [131, 134-136], lo que permite comparar los resultados de ambos métodos. A continuación se simuló un número de «trayectorias» de la reacción bimolecular en disolución para observar con detalle la evolución del sistema. Se empleó de nuevo el método DFT/MM, que ya ha sido aplicado satisfactoriamente para este tipo de estudios [265, 266].

Para reacciones con una energía de activación sensiblemente mayor que las fluctuaciones térmicas ( $k_B T$ ), resulta extremadamente improbable observar un proceso reactivo en el transcurso de una simulación corriente, por lo que se recurre a técnicas como la de «suceso extraordinario» (*rare event approach*) [267-269]. En este caso, se toma como punto de partida para la simulación la estructura de transición de la reacción y se integran las ecuaciones con intervalos temporales ( $\Delta t$ ) positivos y negativos. Es decir, partiendo de la situación en la que el sistema ha llegado al estado de transición, se simula su pasado y su futuro, formando una historia del sistema o trayectoria continua que pasa por el estado de transición, lo que permite aumentar considerablemente la probabilidad de observar procesos reactivos. El problema principal de este procedimiento estriba en la definición de una estructura de transición adecuada. Para este fin es frecuente determinar la geometría del estado de transición con métodos de continuo o definir una coordenada de reacción unidimensional y hallar el máximo de energía libre. En este estudio se utilizó como punto de partida el estado de transición ET2 optimizado con el método ASEP/MD, aprovechando la semejanza de resultados entre este método y el DFT/MM [270].

## 12.2. Detalles de los cálculos

Para los cálculos cuánticos se empleó un método DFT con funcional BP86 y una base NCO(7111/411/1) H(41/1). En las simulaciones se incluyeron, además del

soluto, 215 moléculas de agua TIP3P en una celdilla cúbica de 18,7 Å de lado, las interacciones intermoleculares se truncaron a 9,0 Å.

Se optimizó la geometría de las cinco especies mostradas en la figura 12.1 en vacío. También se hizo con el método de continuo SCRF de Nancy [83-85, 271], con cavidad adaptada a la forma del soluto y  $\epsilon_r = 78,39$ . Para el método ASEP/MD se optó por una optimización completa en cada ciclo del proceso, utilizando el método RFO para las estructuras estables y el P-RFO para las de transición.

Las diferencias de energía libre entre las distintas especies se calcularon como se indica en las ecuaciones (11.1) y (11.2), en los cálculos SCRF también se añadió el término  $\Delta^\ddagger V^{\text{dis}}$  que aparece en dichas ecuaciones. El término  $\Delta G_{\text{QM/MM}}$  se obtuvo por FEP, con 19 estados intermedios ( $\lambda = 0,05$ ) obtenidos por interpolación lineal de geometrías, cargas atómicas y parámetros de Lennard-Jones. En el caso de la diferencia entre los reactivos (triazeno y agua) y el complejo, se consideraron los reactivos separados a una distancia  $\text{N}^1\text{-O}^7$  de 6 Å y se introdujeron 39 estados intermedios ( $\lambda = 0,025$ ).

### 12.2.1. Cálculos DFT/MM

Para las simulaciones DFT/MM se utilizó el programa DFMM [134], con el mismo sistema, método cuántico y parámetros de Lennard-Jones que para los cálculos ASEP/MD, pero sustituyendo los núcleos de hidrógeno (H) por deuterio (D) para reducir las frecuencias de vibración y facilitar la convergencia de las simulaciones. Se empleó un paso temporal de 0,2 fs. Para la simulación de la molécula de triazeno se partió de la geometría optimizada con el ASEP/MD, con una configuración soluto-disolvente obtenida mediante una dinámica molecular clásica, de manera que apenas fueron necesarios 0,1 ps para equilibrar el sistema. La simulación definitiva se efectuó durante 20 ps adicionales.

Una vez optimizado el estado de transición bimolecular, ET2, con ASEP/MD se realizó otra dinámica molecular clásica durante 500 ps. De esta dinámica se tomaron distintas configuraciones del sistema para que sirvieran como punto de partida para las simulaciones DFT/MM. Sin embargo, dado que en la simulación clásica las moléculas se consideran rígidas, estas configuraciones iniciales no incluyen la energía vibracional del soluto. Para intentar representar más adecuadamente el estado vibracional del soluto, se añadieron velocidades relativas a los núcleos del soluto, de la siguiente manera:

- Por diagonalización de la hessiana en disolución, obtenida con el método ASEP/MD, se obtuvieron las frecuencias y modos vibracionales del soluto. De éstos se separaron los modos traslacionales y rotacionales, que se descartaron, y el modo de transición, caracterizado por una constante de fuerza negativa.

- Al modo de transición se le asignó una energía cinética aleatoria en cada configuración siguiendo una distribución de Maxwell-Boltzmann: la probabilidad de una determinada energía  $E$  es proporcional a  $e^{-E/k_B T}$ .
- A cada uno de los restantes modos vibracionales se le asignó una energía cinética  $E_i = 1/4 h\nu_i$ , siendo  $\nu_i$  la correspondiente frecuencia. Esta energía equivale a la energía cinética media de un oscilador armónico de la misma frecuencia en el nivel base. El sentido de cada vibración se eligió aleatoriamente.

Este último punto puede requerir una justificación, ya que la energía cinética de un oscilador armónico cuando éste se encuentra en el mínimo de energía potencial es  $1/2 h\nu$ . Las estructuras optimizadas en disolución corresponden a mínimos en la superficie de energía libre con respecto a las coordenadas vibracionales, pero esta superficie está definida por el promedio termodinámico de todas las configuraciones soluto-disolvente. Para cada configuración concreta, en general, la estructura optimizada no se tratará de un mínimo, sino que contendrá, en cada modo vibracional, una cierta energía potencial. Por lo tanto, a cada configuración individual le corresponde una energía cinética vibracional menor que  $1/2 h\nu$ . Es por ello que se optó por asignar a cada modo vibracional la energía correspondiente a la energía cinética promedio de un oscilador armónico a lo largo de un período de su vibración,  $1/4 h\nu$ .

De esta forma, aunque en todas las configuraciones iniciales consideradas el soluto tiene la misma geometría, con la variación existente en la disposición del disolvente a su alrededor y la introducida en los movimientos nucleares del soluto se consigue un conjunto estadísticamente representativo de configuraciones iniciales para las simulaciones DFT/MM. Para cada una de estas configuraciones iniciales se simuló un tiempo total de 1 ps: 0,5 ps con  $\Delta t = -0,2$  fs y 0,5 ps con  $\Delta t = 0,2$  fs.

## 12.3. Resultados y discusión

### 12.3.1. Triazeno en disolución acuosa

La molécula de triazeno se optimizó en vacío y en disolución acuosa. Su geometría y cargas atómicas se muestran en la tabla 12.2. Los cálculos con el método de continuo SCRF se llevaron a cabo con dos tamaños de cavidad: una cavidad del tamaño predeterminado en el método (radios atómicos de Bondi [272] multiplicados por  $f = 1,30$ ) y otra cavidad más pequeña, de tamaño similar al predeterminado en el método PCM ( $f = 0,98$ ). Esto permite apreciar la dependencia de los resultados respecto al tamaño de la cavidad, lo que representa uno de los puntos débiles de los métodos de continuo.

Los resultados del método DFT/MM corresponden a la media y desviación típica de cada uno de los parámetros geométricos, calculadas a partir de una simulación de 20 ps. Los resultados ASEP/MD, como es habitual, corresponden al promedio de las últimas cinco iteraciones del proceso, tras haberse alcanzado la convergencia en diez iteraciones en este caso (se observan variaciones de menos de 0,001 Å en las distancias de enlace y de menos de 0,1° en los ángulos).

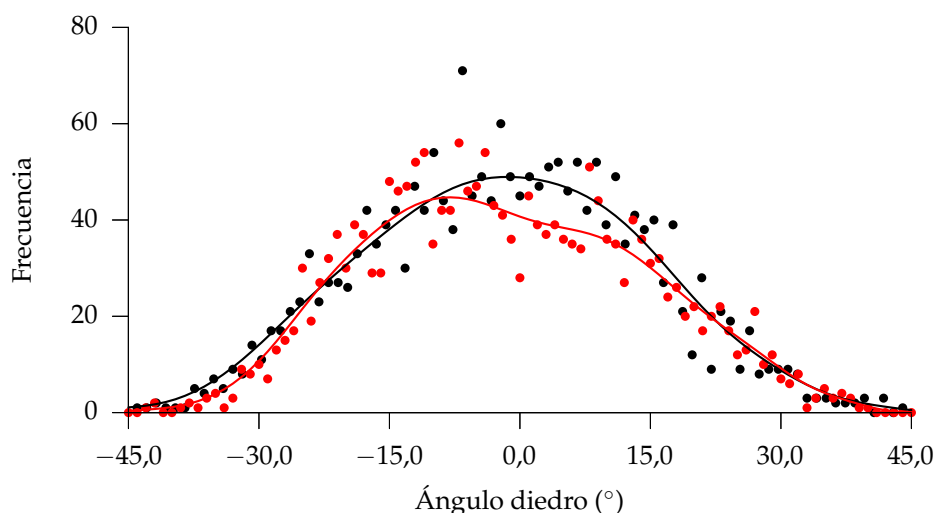
**Tabla 12.2:** Geometrías optimizadas, cargas atómicas y momento dipolar, en vacío y en disolución, de la molécula de *trans*-triazeno. Para el método DFT/MM se da la geometría media y la desviación típica. Distancias en Å, ángulos en grados, cargas en  $e$ , momento dipolar en D.

	Vacío	SCRF		ASEP/MD	DFT/MM
		$f = 1,30$	$f = 0,98$		
N <sup>1</sup> =N <sup>2</sup>	1,271	1,275	1,279	1,287	1,283±0,031
N <sup>2</sup> -N <sup>3</sup>	1,354	1,345	1,331	1,322	1,331±0,030
N <sup>1</sup> -H <sup>4</sup>	1,035	1,034	1,031	1,039	1,047±0,026
N <sup>3</sup> -H <sup>5</sup>	1,016	1,017	1,018	1,020	1,044±0,027
N <sup>3</sup> -H <sup>6</sup>	1,028	1,027	1,027	1,031	1,049±0,028
N <sup>1</sup> =N <sup>2</sup> -N <sup>3</sup>	112,6	113,5	114,9	114,3	115,5±2,7
N <sup>2</sup> =N <sup>1</sup> -H <sup>4</sup>	104,8	104,9	105,6	105,6	105,9±3,6
N <sup>2</sup> -N <sup>3</sup> -H <sup>5</sup>	113,3	114,1	114,8	116,6	114,8±4,1
N <sup>2</sup> -N <sup>3</sup> -H <sup>6</sup>	116,8	118,0	119,6	122,0	119,9±4,3
N <sup>1</sup> =N <sup>2</sup> -N <sup>3</sup> -H <sup>5</sup>	159,0	160,8	163,3	179,3	178,8±15,9
N <sup>1</sup> =N <sup>2</sup> -N <sup>3</sup> -H <sup>6</sup>	16,2	16,1	14,4	0,6	-1,1±15,7
N <sup>1</sup>	-0,398	-0,469	-0,613	-0,636	
N <sup>2</sup>	-0,066	-0,075	-0,092	-0,127	
N <sup>3</sup>	-0,387	-0,384	-0,331	-0,128	
H <sup>4</sup>	0,286	0,312	0,383	0,391	
H <sup>5</sup>	0,281	0,306	0,320	0,247	
H <sup>6</sup>	0,284	0,309	0,333	0,253	
Mom. dipolar ( $\mu$ )	1,70	2,25	2,98	2,96	3,19±0,41

En la tabla 12.2 se aprecia que, tomando el resultado DFT/MM como valor de referencia, tanto ASEP/MD como el método SCRF dan resultados bastante buenos. La variación en distancias y ángulos de enlace al pasar del vacío a la disolución se reproduce bien con ambos métodos, y es más acusada cuando se emplea una cavidad más pequeña con el método SCRF. Las principales diferencias se observan en las distancias N-H. La simulación DFT/MM predice un alargamiento considerable, de entre 0,012 Å y 0,028 Å, mientras que ASEP/MD da un alargamiento de alrededor de 0,004 Å; con el método SCRF apenas hay variaciones y se produce incluso un ligero acortamiento en la distancia N<sup>1</sup>-H<sup>4</sup>. No obstante, en la simulación

DFT/MM hay una gran variación en estas distancias N–H debido a las vibraciones moleculares, con lo que los valores dados por los otros dos métodos entran dentro del intervalo de mayor frecuencia.

Otra diferencia significativa se da en los ángulos diedros reflejados en la tabla. En vacío, los hidrógenos H<sup>5</sup> y H<sup>6</sup> se hallan fuera del plano que contiene a los nitrógenos, de manera que el N<sup>3</sup> es ligeramente piramidal. En disolución, los resultados dependen del método empleado; así, el método SCRF predice la misma estructura piramidal. El método ASEP/MD, sin embargo, predice una conformación prácticamente plana de toda la molécula de triazeno. La simulación DFT/MM da también, en promedio, una estructura plana, pero con una elevada flexibilidad. En la figura 12.3 se muestra la frecuencia con que se dan los distintos valores de los dos ángulos diedros en la simulación DFT/MM, donde se comprueba que, en efecto, la estructura más frecuente es prácticamente plana, aunque los ángulos alcanzan valores de hasta 50°.



**Figura 12.3:** Distribución de valores de ángulos diedros (en grados) en la simulación DFT/MM del *trans*-triazeno. En negro: ángulo  $N^1=N^2-N^3-H^5$  (desplazado 180°); en rojo: ángulo  $N^1=N^2-N^3-H^6$ .

En cuanto a las cargas atómicas, el método de continuo predice principalmente un aumento de la carga negativa en N<sup>1</sup> y de la carga positiva en los hidrógenos, efecto que se intensifica cuando disminuye el tamaño de la cavidad. El método ASEP/MD predice también un aumento de la carga negativa en N<sup>1</sup>, pero da una disminución de la carga positiva en H<sup>5</sup> y H<sup>6</sup> que, junto con una mayor disminución de la carga negativa en N<sup>3</sup>, produce una «despolarización» del grupo –NH<sub>2</sub>. Hay que destacar que estas cargas atómicas son las calculadas mediante CHELPG, que reproducen el potencial electrostático generado por el soluto; en las simulaciones

DFT/MM, sin embargo, las interacciones electrostáticas se calculan directamente a partir del potencial electrostático molecular, por lo que no se tienen valores de cargas atómicas de referencia.

El momento dipolar del triazeno experimenta un importante aumento al pasar del vacío a la disolución. A partir de la simulación DFT/MM se obtiene para este aumento una magnitud de 1,49 D, el 87 % del valor del momento dipolar en vacío. Con el método ASEP/MD y con el método de continuo y la cavidad menor se obtienen valores similares y próximos entre sí, en torno a 1,27 D (75 %). El cálculo con el método de continuo y la cavidad de mayor tamaño, sin embargo, proporciona un valor más alejado: 0,55 D (32 %), lo que pone de manifiesto una vez más la dependencia de los resultados con el tamaño de la cavidad.

Las variaciones que sufre el triazeno al pasar del vacío a la disolución acuosa, tanto en su geometría como en su distribución electrónica, son coherentes con la estabilización de la estructura *zwitteriónica*:  $\text{HN}=\text{N}-\text{NH}_2 \leftrightarrow \text{HN}^\ominus-\text{N}=\text{N}^\oplus\text{H}_2$ .

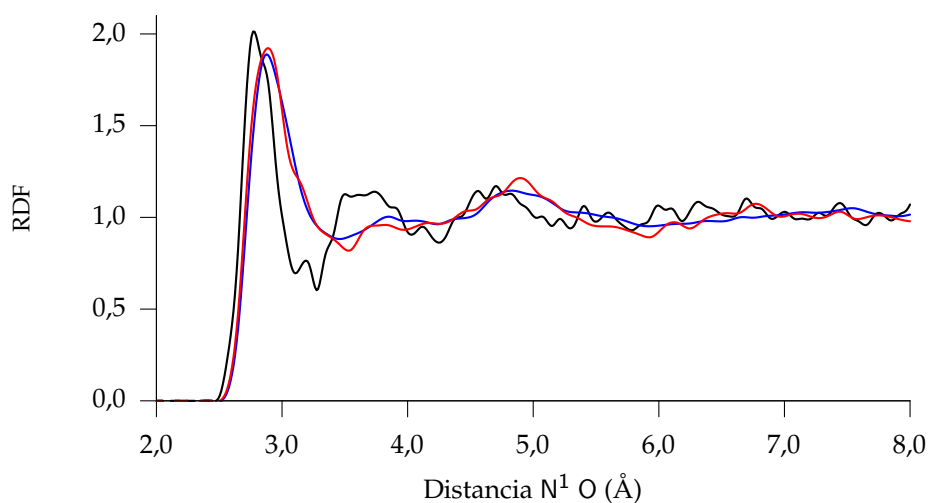
#### – Estructura del disolvente

En las figuras 12.4 y 12.5 se muestran, respectivamente, las RDF del O del agua alrededor del N<sup>1</sup> y del N<sup>3</sup> del triazeno. La RDF obtenida de la simulación DFT/MM se muestra en negro, en rojo aparece la correspondiente a una simulación de dinámica molecular clásica, tomando la geometría y cargas atómicas optimizadas con el método ASEP/MD.

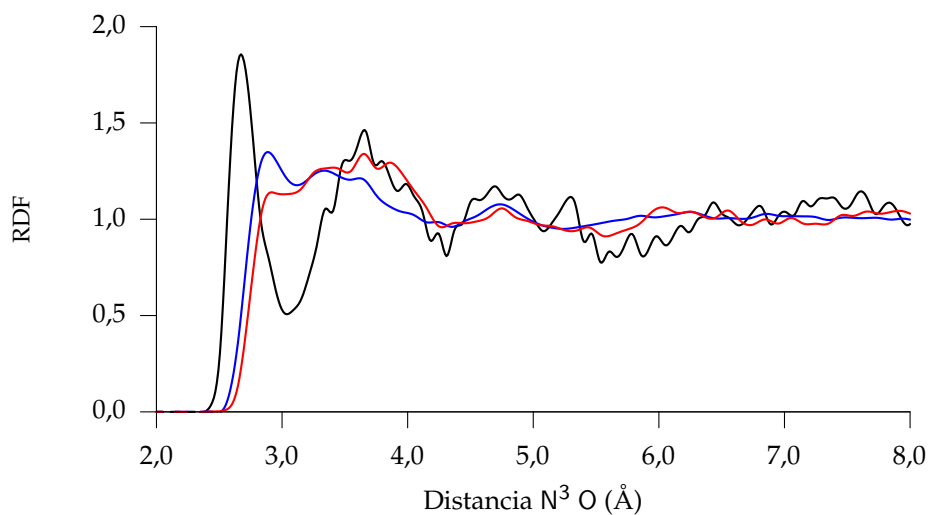
En las RDF del N<sup>1</sup> se aprecia una semejanza significativa entre los resultados DFT/MM y ASEP/MD, la altura del primer pico se reproduce bien, aunque la distancia está algo sobrestimada (2,85 Å frente a 2,75 Å), también se sobrestima su anchura. El segundo y tercer máximos de la RDF están más marcados para la simulación DFT/MM que en el resultado ASEP/MD.

Sin embargo, en las RDF del N<sup>3</sup> hay una diferencia importante: en la simulación DFT/MM aparece un pico muy marcado a 2,7 Å, mientras que en el resultado ASEP/MD sólo se observa un ensanchamiento y un hombro a unos 2,95 Å. El resto de la RDF, a partir de 3,5 Å sí se reproduce bien.

Para intentar determinar la causa de esta diferencia en la RDF del N<sup>3</sup> se realizó una segunda dinámica molecular clásica con moléculas rígidas, pero esta vez tomando la geometría y cargas atómicas optimizadas con el método de continuo SCRF ( $f=0,98$ ). Como ya se ha mencionado, la diferencia más destacada en la estructura del soluto es que, en este caso, el N<sup>3</sup> es piramidal y no plano. Las RDF resultantes se muestran en azul en las figuras 12.4 y 12.5. Se aprecia que apenas hay cambios en la RDF del N<sup>1</sup>, lo cual es lógico teniendo en cuenta que las diferencias en esta zona de la molécula de soluto son poco importantes. En la RDF del N<sup>3</sup>, por otra parte,



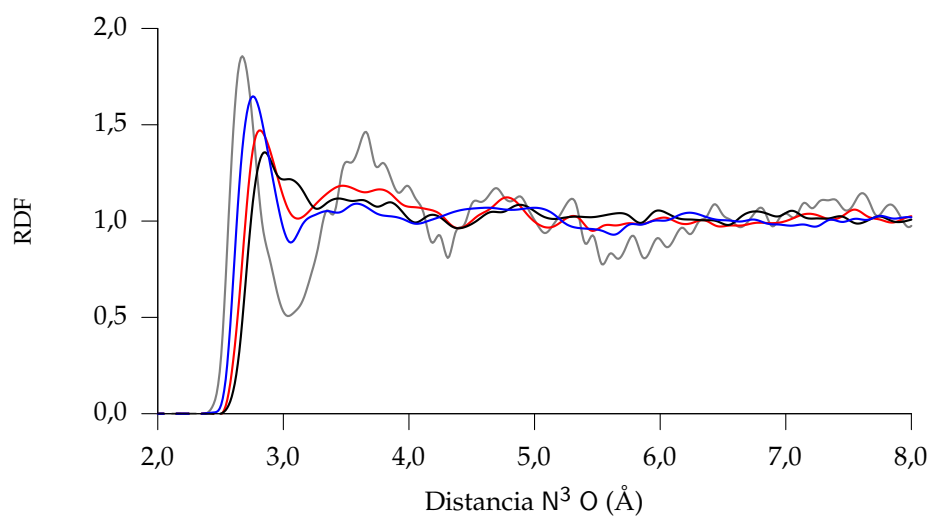
**Figura 12.4:** Funciones de distribución radial entre el N<sup>1</sup> del triazeno y el O del agua, a partir de la simulación DFT/MM (negro) y de simulaciones clásicas con moléculas rígidas, tomando la geometría optimizada con ASEP/MD (rojo) y SCRf (azul).



**Figura 12.5:** Funciones de distribución radial entre el N<sup>3</sup> del triazeno y el O del agua, a partir de la simulación DFT/MM (negro) y de simulaciones clásicas con moléculas rígidas, tomando la geometría y cargas optimizadas con ASEP/MD (rojo) y SCRf (azul).

sí se observa una ligera acentuación del hombro de 2,95 Å, que se convierte en un pico a 2,85 Å, pero también un desplazamiento a menores distancias del segundo pico, originalmente alrededor de 3,7 Å. En general, la RDF es mucho más similar a la obtenida con ASEP/MD que a la DFT/MM. Esto indica que la consideración de una estructura plana para la molécula de triazeno no es la única fuente de la discrepancia observada entre las RDF obtenidas con estos dos métodos.

Se hizo también una serie de cálculos en los que se tomó la geometría del triazeno procedente de algunas configuraciones individuales de la dinámica DFT/MM. Con estas estructuras se realizó un cálculo ASEP/MD parcialmente acoplado, sin optimización de geometría, obteniéndose la distribución de cargas del soluto y la configuración del disolvente a su alrededor. Algunas de las RDF resultantes de estos cálculos se muestran en la figura 12.6. Las funciones dibujadas en negro y azul corresponden a estructuras en las que el N<sup>3</sup> es claramente piramidal, con ángulos diedros del orden de 25°, la diferencia más destacada entre ambas se encuentra en las distancias N<sup>3</sup>-H, que son de unos 1,02 Å para la estructura en negro y de 1,08 Å para la azul. La función en rojo, por su parte, corresponde a una estructura con distancias N<sup>3</sup>-H relativamente largas (de unos 1,10 Å, pero prácticamente plana (diedros de 1° y 10°).



**Figura 12.6:** Funciones de distribución radial entre el N<sup>3</sup> del triazeno y el O del agua, obtenidas con el método ASEP/MD a partir de diferentes geometrías del soluto tomadas de la simulación DFT/MM. En gris, la RDF de la simulación DFT/MM, ver el texto para el resto.

A partir de estos resultados puede deducirse que, efectivamente, la variabilidad en las distancias N<sup>3</sup>-H y en los ángulos diedros son factores que influyen en la forma final de las RDF. Otro factor importante que puede determinar la divergencia entre las RDF de los métodos DFT/MM y ASEP/MD es la doble representación



del soluto en este último. En los cálculos cuánticos, el soluto se introduce como una distribución electrónica continua, mientras que durante las simulaciones de dinámica molecular, se reduce a un conjunto discreto de cargas situadas sobre los núcleos atómicos. En trabajos previos [138] ya se comprobó que una mejora en la representación de la distribución electrónica del soluto aumentando el número de cargas puede conducir a funciones de distribución radial más semejantes a las obtenidas experimentalmente y, en este caso, a las proporcionadas por las simulaciones DFT/MM.

### 12.3.2. Intermedios y estados de transición

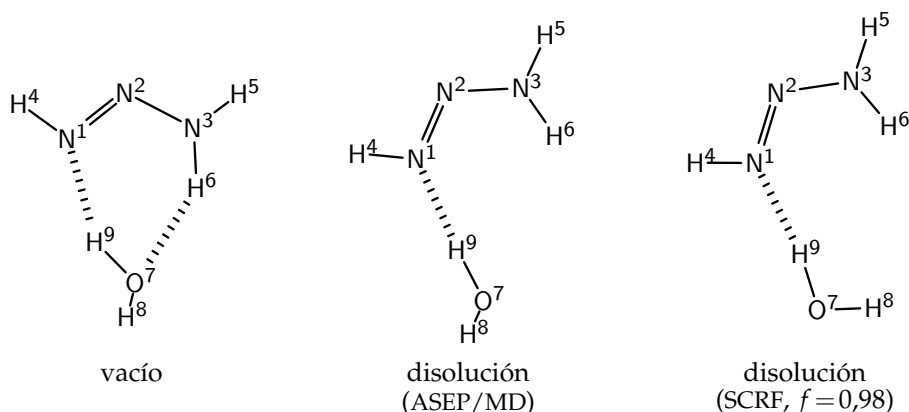
El resto de las especies mostradas en la figura 12.1, agua, complejo, ET1 y ET2, se estudió únicamente con los métodos SCRF y ASEP/MD. Para la molécula de agua no se muestran los resultados, ya que son similares a los reflejados en la sección 9.2.3 (pág. 133). En primer lugar se optimizó la estructura del complejo triazeno-agua, estabilizado en vacío por dos enlaces de hidrógeno entre ambas moléculas. Los parámetros geométricos y cargas atómicas obtenidos se muestran en la tabla 12.7 y en la figura 12.8.

Con respecto a la estructura en vacío, el principal cambio en disolución es la ruptura del enlace de hidrógeno  $O^7-H^6$  y el acortamiento del  $N^1-H^9$ , la molécula de agua pasa a estar ligada únicamente al  $N^1$ . También se aprecia, con el método ASEP/MD, un ligero alargamiento de los enlaces  $N-H$  y  $O-H$  que no participan en la unión intermolecular, todo esto es coherente con el establecimiento de una competencia entre la interacción con el disolvente y la interacción entre las dos moléculas. De nuevo el método ASEP/MD predice una molécula de triazeno plana, mientras que el SCRF da una estructura piramidal para el  $N^3$ , incluso algo más pronunciada que en el vacío. En las cargas atómicas, el efecto más destacado es la polarización de la molécula de agua, sobre todo del enlace  $O^7-H^8$ .

Al contrario de lo que ocurriría con la molécula de triazeno, en este caso la disminución del tamaño de la cavidad no acerca los resultados SCRF a los ASEP/MD. Con la cavidad predeterminada ( $f=1,30$ ) algunos parámetros moleculares, especialmente los correspondientes al triazeno, son muy similares a los obtenidos mediante ASEP/MD, pero no se observa una ruptura tan clara del enlace  $O^7-H^6$ . Con la cavidad más reducida ( $f=0,98$ ) se produce una mayor polarización en la molécula de agua y un alejamiento de  $O^7$  y  $H^6$ , pero la orientación de  $H^8$  cambia completamente, colocándose enfrente a  $H^6$ . Esto pone de manifiesto la principal limitación de los métodos de continuo: el tamaño y la forma de la cavidad escogida afecta a los resultados, y no es posible definir una cavidad que funcione bien en todas las condiciones. Este problema es especialmente patente cuando existen enlaces débiles en el soluto, como ocurre en el caso del complejo triazeno-agua.

**Tabla 12.7:** Geometrías optimizadas y cargas atómicas, en vacío y en disolución, del complejo del *trans*-triazeno con agua. Distancias en Å, ángulos en grados, cargas en *e*.

	Vacío	SCRF		ASEP/MD
		<i>f</i> = 1,30	<i>f</i> = 0,98	
N <sup>1</sup> =N <sup>2</sup>	1,282	1,282	1,282	1,282
N <sup>2</sup> -N <sup>3</sup>	1,333	1,331	1,323	1,321
N <sup>1</sup> -H <sup>4</sup>	1,031	1,030	1,030	1,033
N <sup>3</sup> -H <sup>5</sup>	1,013	1,015	1,016	1,023
N <sup>3</sup> -H <sup>6</sup>	1,037	1,034	1,026	1,033
N <sup>1</sup> -H <sup>9</sup>	1,855	1,805	1,768	1,713
H <sup>6</sup> -O <sup>7</sup>	1,962	2,077	3,425	3,102
O <sup>7</sup> -H <sup>9</sup>	1,002	1,006	1,012	1,017
O <sup>7</sup> -H <sup>8</sup>	0,974	0,975	0,976	0,989
N <sup>1</sup> =N <sup>2</sup> -N <sup>3</sup>	113,9	114,1	115,4	115,2
N <sup>2</sup> =N <sup>1</sup> -H <sup>4</sup>	105,7	106,2	106,4	106,5
N <sup>2</sup> -N <sup>3</sup> -H <sup>5</sup>	114,7	115,2	116,1	116,3
N <sup>2</sup> -N <sup>3</sup> -H <sup>6</sup>	120,3	119,8	121,1	122,7
N <sup>2</sup> =N <sup>1</sup> -H <sup>9</sup>	113,4	114,4	136,6	132,2
N <sup>3</sup> -H <sup>6</sup> -O <sup>7</sup>	144,5	142,0	118,2	120,3
N <sup>1</sup> -H <sup>9</sup> -O <sup>7</sup>	150,4	156,0	170,8	169,8
H <sup>6</sup> -O <sup>7</sup> -H <sup>9</sup>	77,0	72,7	38,8	51,6
H <sup>8</sup> -O <sup>7</sup> -H <sup>9</sup>	105,8	105,4	105,1	106,8
N <sup>1</sup> =N <sup>2</sup> -N <sup>3</sup> -H <sup>5</sup>	171,3	169,7	169,3	178,9
N <sup>1</sup> =N <sup>2</sup> -N <sup>3</sup> -H <sup>6</sup>	7,7	10,2	9,9	0,8
O <sup>7</sup> -H <sup>9</sup> -N <sup>1</sup> =N <sup>2</sup>	2,9	8,0	172,1	59,2
N <sup>1</sup> -H <sup>9</sup> -O <sup>7</sup> -H <sup>8</sup>	127,4	116,0	171,5	172,3
N <sup>3</sup> -H <sup>6</sup> -O <sup>7</sup> -H <sup>8</sup>	103,4	106,6	170,0	103,4
N <sup>1</sup>	-0,158	-0,135	-0,085	-0,130
N <sup>2</sup>	-0,158	-0,204	-0,296	-0,244
N <sup>3</sup>	-0,320	-0,282	-0,333	-0,250
H <sup>4</sup>	0,192	0,202	0,242	0,226
H <sup>5</sup>	0,279	0,288	0,359	0,313
H <sup>6</sup>	0,290	0,301	0,343	0,303
O <sup>7</sup>	-0,682	-0,800	-0,965	-0,958
H <sup>8</sup>	0,355	0,425	0,490	0,479
H <sup>9</sup>	0,202	0,205	0,245	0,259



**Figura 12.8:** Geometrías optimizadas en vacío y en disolución (calculada con el método ASEP/MD y con el método de continuo SCRF) del complejo triazeno-agua.

En el caso del estado de transición ET1, se muestran sus geometrías y cargas en la tabla 12.9. En esta especie apenas se observan cambios entre las estructuras optimizadas en vacío y en disolución, pero los pocos cambios que se producen siguen siendo algo más pronunciados con el método ASEP/MD que con el SCRF. Con respecto a la molécula de triazeno, se igualan las distancias N–N y se acortan las distancias N–H no implicadas en la reacción. El átomo H<sup>6</sup>, que es el que sufre el desplazamiento, se sitúa a medio camino entre N<sup>1</sup> y N<sup>3</sup>, como es lógico, a una distancia sensiblemente mayor que la correspondiente a un enlace N–H normal. El ángulo N–N–N se estrecha al producirse la transferencia del H<sup>6</sup>. También en comparación con el triazeno, las cargas atómicas del ET1 son menores en valor absoluto, especialmente en disolución; esto indica que, durante la transferencia protónica, el reactivo se polariza más que el estado de transición y es previsible que en disolución se establezca en mayor medida que éste, originando un aumento de la energía de activación.

Las geometrías y cargas optimizadas para el estado de transición ET2 se muestran en la tabla 12.10. En este caso sí se aprecian mayores diferencias entre las estructuras en vacío y en disolución. Los cambios más importantes son el acortamiento de los enlaces N–N y, especialmente, de los N–H implicados en la transferencia (una diferencia de 0,125 Å), así como el alargamiento de las distancias entre estos H y el O (un aumento de 0,240 Å). El efecto general es la separación de dos fragmentos correspondientes a N<sub>3</sub>H<sub>4</sub> y a OH, como se observa en la figura 12.11. En los ángulos de enlace no hay variaciones importantes al pasar del vacío a la disolución. El comportamiento de las cargas atómicas también es interesante, el valor absoluto de la carga sobre el O aumenta en casi 0,5 *e* y también aumentan las cargas sobre los átomos de H. Los resultados obtenidos con el método SCRF y con *f* = 0,98 son en este caso muy similares a los del ASEP/MD, pero los aumentos en las longitudes

**Tabla 12.9:** Geometrías optimizadas y cargas atómicas, en vacío y en disolución, del estado de transición unimolecular del *trans*-triazeno, ET1. Los valores que no aparecen en la tabla se pueden obtener por simetría. Distancias en Å, ángulos en grados, cargas en *e*.

	Vacío	SCRF		ASEP/MD
		<i>f</i> = 1,30	<i>f</i> = 0,98	
N <sup>1</sup> –N <sup>2</sup>	1,310	1,311	1,311	1,310
N <sup>1</sup> –H <sup>4</sup>	1,027	1,027	1,029	1,033
N <sup>1</sup> –H <sup>6</sup>	1,365	1,366	1,368	1,369
N <sup>1</sup> –N <sup>2</sup> –N <sup>3</sup>	101,3	101,6	102,1	101,8
N <sup>2</sup> –N <sup>1</sup> –H <sup>4</sup>	115,1	115,3	115,3	116,0
N <sup>2</sup> –N <sup>1</sup> –H <sup>6</sup>	81,5	81,2	80,8	81,2
N <sup>1</sup> –H <sup>6</sup> –N <sup>3</sup>	95,8	96,1	96,4	95,9
H <sup>4</sup> –N <sup>1</sup> –N <sup>2</sup> –N <sup>3</sup>	162,6	161,3	159,1	164,1
H <sup>6</sup> –N <sup>1</sup> –N <sup>2</sup> –N <sup>3</sup>	0,0	0,0	0,0	0,1
N <sup>1</sup>	–0,258	–0,269	–0,285	–0,285
N <sup>2</sup>	–0,130	–0,154	–0,200	–0,172
H <sup>4</sup>	0,260	0,282	0,315	0,306
H <sup>6</sup>	0,125	0,127	0,139	0,129

de los enlaces X–H no implicados en la transferencia siguen siendo mayores con este último.

Considerando las variaciones en la geometría y en las cargas que tienen lugar en el ET2 al introducirse en disolución acuosa, puede decirse que su estructura es prácticamente la de un par iónico de N<sub>3</sub>H<sub>4</sub><sup>⊕</sup> y OH<sup>⊖</sup>. Esta importante separación de cargas hace prever una estabilización de ET2 respecto a los reactivos bastante mayor que en el caso unimolecular, lo que debe ocasionar una disminución en la energía de activación al pasar del vacío a la disolución, como se verá en el siguiente apartado.

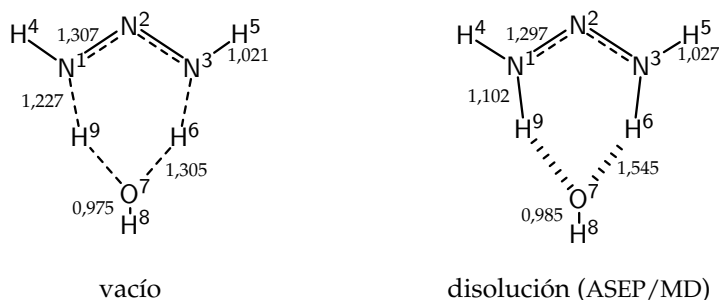
### 12.3.3. Resultados energéticos

Tras obtener las estructuras optimizadas en vacío y en disolución para todas las especies implicadas, se procedió a calcular las diferencias de energía libre entre ellas empleando las ecuaciones (11.1) y (11.2), como se ha indicado. Se calcularon las variaciones de energía libre para los distintos procesos que se muestran en la tabla 12.12.

Para el proceso de desplazamiento unimolecular, sin intervención directa de las moléculas del disolvente, la energía libre de activación resulta en todos los

**Tabla 12.10:** Geometrías optimizadas y cargas atómicas, en vacío y en disolución, del estado de transición bimolecular del *trans*-triazeno con agua, ET2. Los valores que no aparecen en la tabla se pueden obtener por simetría. Distancias en Å, ángulos en grados, cargas en *e*.

	Vacío	SCRF		ASEP/MD
		$f = 1,30$	$f = 0,98$	
N <sup>1</sup> –N <sup>2</sup>	1,307	1,305	1,297	1,297
N <sup>1</sup> –H <sup>4</sup>	1,021	1,020	1,019	1,027
N <sup>1</sup> –H <sup>9</sup>	1,227	1,175	1,101	1,102
O <sup>7</sup> –H <sup>9</sup>	1,305	1,385	1,560	1,545
O <sup>7</sup> –H <sup>8</sup>	0,975	0,976	0,977	0,985
N <sup>1</sup> –N <sup>2</sup> –N <sup>3</sup>	112,8	113,0	114,1	113,9
N <sup>2</sup> –N <sup>1</sup> –H <sup>4</sup>	111,1	112,6	115,1	114,9
N <sup>2</sup> –N <sup>1</sup> –H <sup>9</sup>	112,7	113,5	115,7	116,1
N <sup>1</sup> –H <sup>9</sup> –O <sup>7</sup>	149,4	149,9	149,7	148,4
H <sup>6</sup> –O <sup>7</sup> –H <sup>9</sup>	82,4	79,6	74,9	76,2
N <sup>2</sup> –O <sup>7</sup> –H <sup>8</sup>	122,8	119,5	111,5	127,0
H <sup>4</sup> –N <sup>1</sup> –N <sup>2</sup> –N <sup>3</sup>	178,5	176,9	178,8	179,2
H <sup>9</sup> –N <sup>1</sup> –N <sup>2</sup> –N <sup>3</sup>	4,9	6,8	2,4	1,8
H <sup>4</sup> –N <sup>1</sup> –H <sup>9</sup> –O <sup>7</sup>	178,8	176,2	178,8	172,5
N <sup>1</sup>	–0,108	–0,112	–0,046	–0,079
N <sup>2</sup>	–0,235	–0,236	–0,266	–0,222
H <sup>4</sup>	0,205	0,239	0,267	0,270
H <sup>9</sup>	0,223	0,267	0,308	0,302
O <sup>7</sup>	–0,772	–0,962	–1,238	–1,255
H <sup>8</sup>	0,367	0,411	0,446	0,489



**Figura 12.11:** Geometrías optimizadas en vacío y en disolución (calculada con el método ASEP/MD) del estado de transición bimolecular del *trans*-triazeno con agua, ET2. Distancias en Å.

**Tabla 12.12:** Diferencias de energía libre correspondientes a distintos procesos del triazeno, calculadas en vacío y en disolución con distintos métodos, en kcal/mol. En el caso del SCRF se dan los valores calculados con volumen predeterminado ( $f = 1,30$ ) y escalado ( $f = 0,98$ ). En el caso del ASEP/MD se da también la componente  $\Delta G_{QM/MM}$  (ec. 11.2).

	Vacío	SCRF		ASEP/MD	
		$f = 1,30$	$f = 0,98$	$\Delta G$	$\Delta G_{QM/MM}$
$N_3H_3 \rightarrow ET1$	32,03	33,93	37,39	34,91	6,13
$N_3H_3 + H_2O \rightarrow$ complejo	0,86	4,72	7,95	6,66	0,63
complejo $\rightarrow ET2$	9,21	8,86	11,40	5,46	-8,09
$N_3H_3 + H_2O \rightarrow ET2$	10,07				

casos muy elevada, de más de 30 kcal/mol. Además, todos los métodos empleados predicen un aumento de dicha energía en disolución acuosa de entre 1,9 kcal/mol y 5,4 kcal/mol. Con el método ASEP/MD se observa que la razón de este aumento es, principalmente, la pérdida de interacciones favorables entre soluto y disolvente, que se traduce en un valor de  $\Delta^\ddagger G_{QM/MM}$  positivo y significativo (6,1 kcal/mol).

La intervención de una molécula de disolvente favorece claramente el desplazamiento protónico. En vacío, la energía de activación se reduce a unas 10 kcal/mol, correspondientes casi en su totalidad al paso del complejo triazeno-agua a ET2, ya que la formación del complejo a partir de los reactivos transcurre prácticamente sin variación en la energía libre; existe una disminución de entalpía que se ve compensada por la disminución de entropía.

En disolución la situación es más compleja. Aunque se ha calculado la diferencia de energía libre entre los reactivos por separado y el complejo, no está claro el sentido físico de este resultado en un caso como éste, en el que uno de los reactivos es el propio disolvente. Esto equivale a considerar que el «punto de partida» para la reacción de formación del complejo es la situación en la que los dos reactivos están a una distancia suficientemente grande como para poder considerarlos aislados. Sin embargo, este estado es irreal, pues en la disolución siempre habrá alguna molécula de disolvente (que es uno de los reactivos) en una posición más favorable para formar el complejo con el otro reactivo. Así pues, no es sencillo definir cuál es la situación que podemos llamar «reactivos» en una reacción como esta; parece lógico que no sea una especie tan definida como el complejo triazeno-agua, pero tampoco algo tan inconexo como las dos moléculas aisladas. Más concretamente, no tiene sentido tener en cuenta la disminución de entropía que se produce al aproximar los dos reactivos si, en la disolución, siempre va a haber una molécula de agua cerca

del triazeno. Dada esta situación, probablemente sea más adecuado considerar la energía libre del proceso complejo $\rightarrow$ ET2 como la energía libre de activación en disolución.

Aceptando esta aproximación, la energía de activación resultante para el desplazamiento asistido por el disolvente con el método ASEP/MD es de 5,46 kcal/mol, casi la mitad que en vacío, mientras con el método de continuo SCRF los resultados —y el signo de la variación con respecto al vacío— dependen del tamaño de la cavidad (8,86 kcal/mol para  $f=1,30$  y 11,40 kcal/mol para  $f=0,98$ ). La disminución de la energía de activación observada con ASEP/MD es mayoritariamente debida a  $\Delta^\ddagger G_{QM/MM}$ , que en esta ocasión es de signo negativo. Esto es coherente con el alto grado de separación de carga encontrado para ET2, lo que hace que la interacción con el disolvente sea más favorable para el estado de transición que para el complejo.

Pye y col. [264] hicieron un estudio empleando diversos métodos *ab initio* de distintas conformaciones y estados de transición del triazeno en fase gaseosa. Con el método de mayor calidad de todos los que emplearon (MP2/6-31+G\*) obtuvieron 41,37 kcal/mol y 18,36 kcal/mol para las energías de activación de los desplazamientos unimolecular y bimolecular, respectivamente, en vacío. Ambos valores son entre 8 kcal/mol y 9 kcal/mol mayores que los señalados en la tabla 12.12. Para determinar la causa de estas discrepancias se realizaron otros cálculos adicionales en vacío, variando el método cuántico y el conjunto de base. Los resultados se muestran en la tabla 12.13.

**Tabla 12.13:** Energías libres de activación calculadas en vacío para los desplazamientos protónicos unimolecular y bimolecular en el *trans*-triazeno, en kcal/mol. La base «Gen» es la de tipo NCO(7111/411/1) H(41/1) empleada en este trabajo [185, 186].

	BP86/Gen	BP86/6-31+G*	B3LYP/Gen
N <sub>3</sub> H <sub>3</sub> $\rightarrow$ ET1	32,03	34,03	37,62
N <sub>3</sub> H <sub>3</sub> + H <sub>2</sub> O $\rightarrow$ ET2	10,07	11,46	16,24
	MP2/Gen//BP86/Gen	MP2/6-31+G*[264]	
N <sub>3</sub> H <sub>3</sub> $\rightarrow$ ET1	39,92	41,37	
N <sub>3</sub> H <sub>3</sub> + H <sub>2</sub> O $\rightarrow$ ET2	17,73	18,36	

Al cambiar únicamente la base y emplear la 6-31+G\*, las energías aumentan ligeramente, pero aún están muy lejos de los resultados de Pye y col. Cambiando sólo el funcional por B3LYP, sin embargo, la variación es mucho mayor y los resultados son más similares a los mencionados. La influencia de la geometría es prácticamente despreciable, como se observa en el cálculo MP2/Gen//BP86/Gen,

cuyos resultados difieren tan sólo alrededor de 1 kcal/mol de los MP2/6-31+G\*. Estas pruebas permiten concluir que la razón principal de la discrepancia en la energía de activación es el funcional empleado, BP86, y que otro funcional más adecuado, como B3LYP, hubiera dado resultados más comparables a los de [Pye y col.](#) Desafortunadamente, en el programa utilizado para hacer los cálculos DFT/MM, no se cuenta con la posibilidad de aplicar funcionales híbridos como el B3LYP en los cálculos cuánticos y se consideró más adecuado mantener un mismo método cuántico en todo el estudio, confiando en que los errores introducidos puedan suponerse sistemáticos. De hecho, el efecto catalítico del agua es de unas 22 kcal/mol con BP86/Gen, valor muy similar a las 23 kcal/mol obtenidas con MP2/6-31+G\*.

Por otro lado, existe al menos una medición experimental de las constantes de velocidad del desplazamiento protónico en triazenos. Smith y col. [259] determinaron, a partir de los ensanchamientos de las señales RMN, las constantes de tautomerización de distintos alquiltriazenos en metanol a 295 K, obteniendo valores de entre  $8,5 \text{ s}^{-1}$  para el 1-metil-3-*terc*-butiltriazeno y  $2040 \text{ s}^{-1}$  para el 1-metil-3-etiltriazeno. Estas constantes de velocidad corresponden a energías de activación de entre 16,0 kcal/mol y 12,8 kcal/mol, respectivamente, sustituyendo  $k$  en la fórmula:

$$\Delta^\ddagger G = Nk_B T \left( \ln \frac{k_B}{h} - \ln \frac{k}{\kappa T} \right) \quad (12.1)$$

Para el coeficiente de transmisión,  $\kappa$ , se supone un valor de 1. Más adelante se estimará el valor de  $\kappa$ , pero el resultado para  $\Delta^\ddagger G$  apenas depende de él, siempre que se mantenga en un orden de magnitud cercano a la unidad.

Estos valores no son directamente comparables con los obtenidos teóricamente. En primer lugar, porque están determinados para sistemas diferentes (alquiltriazenos en metanol frente a triazeno en agua) y en segundo lugar, porque en el experimento pueden tener lugar otros mecanismos de reacción diferentes al estudiado teóricamente. Sin embargo, los órdenes de magnitud de las energías de activación sí son similares, sobre todo si se tiene en cuenta que, como se ha apuntado antes, los cálculos teóricos pueden estar subestimados en unas 8 kcal/mol.

#### 12.3.4. Trayectorias de reacción DFT/MM

Tras haber obtenido la estructura optimizada para ET2 en disolución, se realizó un estudio dinámico de la reacción de desplazamiento protónico con el método DFT/MM, como se detalla en el apartado 12.2.1. En total se calcularon 50 trayectorias de 1 ps, en las que las coordenadas de los núcleos del soluto no están sujetas a restricciones. En todos los casos se partió de un soluto con idéntica estructura —la optimizada para ET2 con el método ASEP/MD—, las condiciones iniciales de las trayectorias se diferencian en la configuración del disolvente a su alrededor y en



las velocidades iniciales, tanto de las moléculas del disolvente como de los núcleos del soluto.

Dado que la reacción estudiada ( $\text{N}_3\text{H}_3 + \text{H}_2\text{O} \rightarrow \text{N}_3\text{H}_3 + \text{H}_2\text{O}$ ) es simétrica y los reactivos son indistinguibles de los productos, se consideraron trayectorias reactivas aquellas en las que la molécula de agua formada contiene diferentes átomos de hidrógeno al inicio y al final ( $\text{H}^9$  y  $\text{H}^6$ , ver figura 12.1). Así, del total de 50 trayectorias, se contabilizaron 31 (62%) reactivas y 19 (38%) no reactivas. De estas trayectorias no reactivas, 11 fueron del tipo reactivos→reactivos (la molécula de agua contiene a  $\text{H}^9$  al principio y al final de la trayectoria) y 8 del tipo productos→productos (la molécula de agua contiene a  $\text{H}^6$ ).

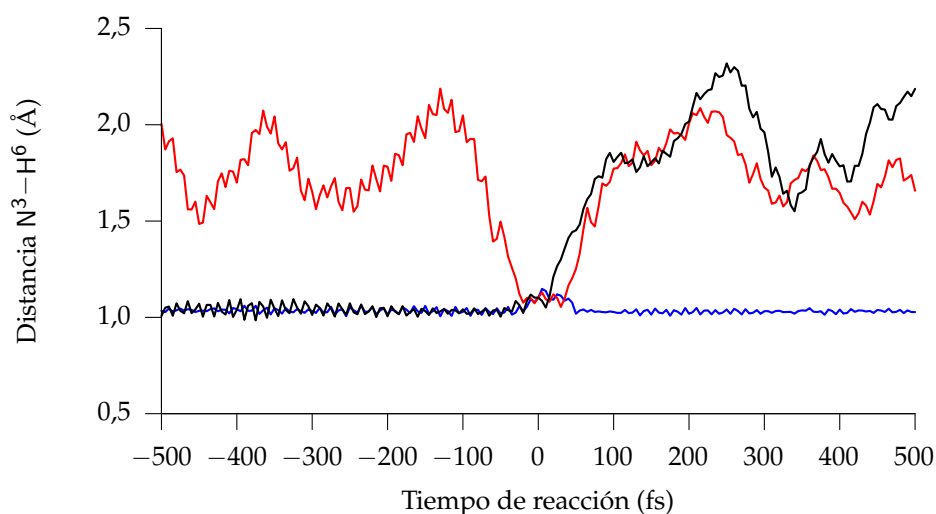
A partir de estos datos puede evaluarse el coeficiente de transmisión,  $\kappa$ , según la expresión [273, 274]:

$$\kappa = \frac{\sum v_i \delta_i}{\sum v_i} \quad (12.2)$$

donde  $v_i$  es la velocidad inicial en el modo de transición de cada trayectoria, proporcional a la raíz cuadrada de la energía cinética asignada en cada caso,  $v_i \propto \sqrt{E_i}$ , y  $\delta_i$  es un factor que vale 1 para las trayectorias reactivas y 0 para las no reactivas. De esta forma se obtiene para  $\kappa$  un valor de 0,73. Este valor, próximo a 1, indica que las condiciones iniciales empleadas para las simulaciones son una buena representación del estado de transición de la reacción. Por otro lado, su desviación con respecto a la unidad es significativa, lo que pone de manifiesto la existencia de recruces y efectos dinámicos que limitan la validez de la teoría del estado de transición.

En la figura 12.14 se muestran tres ejemplos de los distintos tipos de trayectoria, en los que se ha representado la variación de la distancia  $\text{N}^3-\text{H}^6$  con el tiempo. Para una trayectoria reactiva, esta distancia comienza con un valor de aproximadamente 1,0 Å para, tras pasar por el estado de transición (a 0 fs), aumentar hasta valores de 1,5 Å o más, dependiendo del movimiento de la molécula de agua tras la reacción. Para trayectorias no reactivas del tipo reactivos→reactivos, la distancia  $\text{N}^3-\text{H}^6$  se mantiene siempre en torno a 1,0 Å (con un ligero aumento en el estado de transición), mientras que para las del tipo productos→productos, la distancia es mayor al principio y al final de la trayectoria, con una disminución y posterior aumento en la zona central, correspondiente a la formación del estado de transición. Dada la definición de reactivos y productos, el comportamiento inverso se observaría al examinar la distancia  $\text{N}^1-\text{H}^9$ .

Como ya se ha mencionado, en cada trayectoria se introdujo una energía cinética aleatoria en la coordenada de reacción (modo normal con constante de fuerza negativa), siguiendo una distribución de Maxwell-Boltzmann a 298 K, de manera que el promedio de esta energía fue de 0,58 kcal/mol. Se observa que, en general, las trayectorias no reactivas tienden a ser aquellas en las que la energía introducida

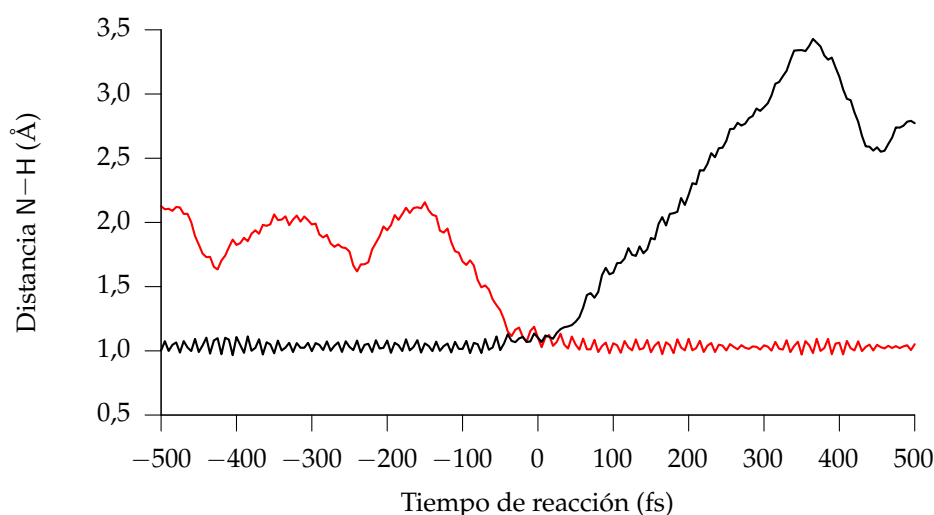


**Figura 12.14:** Evolución de la distancia  $N^3-H^6$  en tres tipos de trayectorias para la reacción estudiada. En negro: trayectoria reactiva reactivos→productos. En rojo: trayectoria no reactiva productos→productos. En azul: trayectoria no reactiva reactivos→reactivos.

en la coordenada de reacción fue menor. Así, el promedio para las trayectorias no reactivas es de 0,27 kcal/mol, mientras que para las trayectorias reactivas es de 0,76 kcal/mol. Sin embargo, se encontraron trayectorias reactivas con energía tan baja como 0,001 kcal/mol y trayectorias no reactivas con hasta 1,20 kcal/mol de energía cinética en la coordenada de reacción. De las 12 trayectorias con energía mayor de 1,0 kcal/mol, 10 resultaron reactivas. Todo esto se puede observar en la figura 12.17, que se comentará más adelante.

Además de estos datos, es interesante también examinar otras características de las trayectorias obtenidas. Se puede, por ejemplo, observar cuál es el destino de la molécula de agua después de la reacción y determinar si ésta permanece unida por enlaces de hidrógeno al triazeno, en una situación similar al complejo triazeno-agua estudiado anteriormente, o si, por el contrario, se aleja rápidamente para introducirse en el grueso del disolvente. Estos dos tipos de comportamiento se pueden discriminar al representar la variación de las distancias  $N^1-H^9$  y  $N^3-H^6$  con el tiempo. Tanto en las trayectorias reactivas como en las no reactivas, una de estas dos distancias estará en torno a 1,0 Å al final del intervalo de 1 ps, mientras que la otra distancia determina el comportamiento de la molécula de agua: si se mantiene entre 1,5 Å y 2,0 Å, el agua permanece unida al triazeno; si se observa una tendencia a aumentar, al menos el enlace de hidrógeno correspondiente se rompe definitivamente. Obviamente, por la simetría de la reacción, todo esto también es aplicable para la parte de la trayectoria anterior a la formación del estado de transición. Un buen ejemplo es la trayectoria mostrada en la figura 12.15, se

trata de una trayectoria reactiva en la que, antes de que la reacción tenga lugar, la molécula de agua forma un enlace de hidrógeno  $N^1-H^9$  relativamente duradero con el triazeno, pero al producirse la transferencia protónica el enlace  $N^3-H^6$  se rompe rápidamente y los reactivos se alejan.

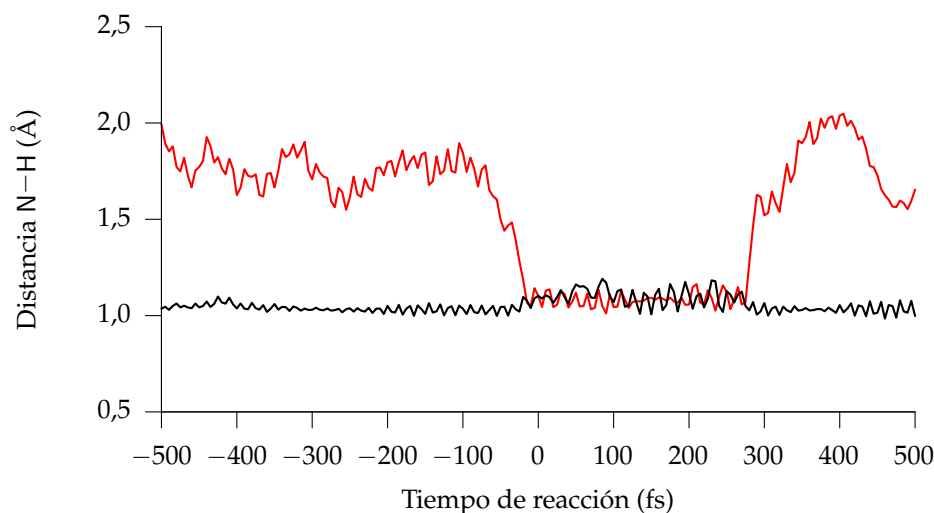


**Figura 12.15:** Ejemplo de trayectoria en la que se observan dos comportamientos de la molécula de agua en la reacción con el triazeno. En negro: distancia  $N^3-H^6$ . En rojo: distancia  $N^1-H^9$ .

En la gran mayoría de las trayectorias calculadas se encontró un comportamiento similar al de la línea roja de la figura 12.15. Considerando tanto la situación de los reactivos en los 500 fs anteriores al estado de transición como en los 500 fs posteriores, sólo en 13 de los 100 casos disponibles se observa una distancia  $N^1-H^9$  o  $N^3-H^6$  claramente mayor que 2,0 Å. Además, el valor absoluto de la diferencia entre las distancias  $O^7-H^6$  y  $O^6-H^9$ , que también puede ser un indicador del grado de separación entre triazeno y agua, raramente supera los 2,5 Å (en la estructura optimizada para el complejo, esta diferencia es de 2,1 Å, ver tabla 12.7). Estos resultados indican que la unión por enlaces de hidrógeno entre el triazeno y el agua da lugar a un complejo estable en la escala de tiempo de estas simulaciones, lo que apoya la idea de que es más apropiado tomar la estructura optimizada del complejo triazeno-agua como el reactivo a la hora de calcular la energía libre de activación.

Otro aspecto que salta a la vista al examinar las trayectorias es la existencia, en muchas de ellas, de un intervalo temporal considerable en el que la estructura del soluto permanece próxima a la del estado de transición. Esto se aprecia, por ejemplo, en la trayectoria mostrada en la figura 12.16, que es un caso bastante extremo. Aquí, durante casi 300 fs, las distancias  $N^1-H^9$  y  $N^3-H^6$  se mantienen prácticamente iguales, con una estructura casi simétrica, sin definirse claramente la

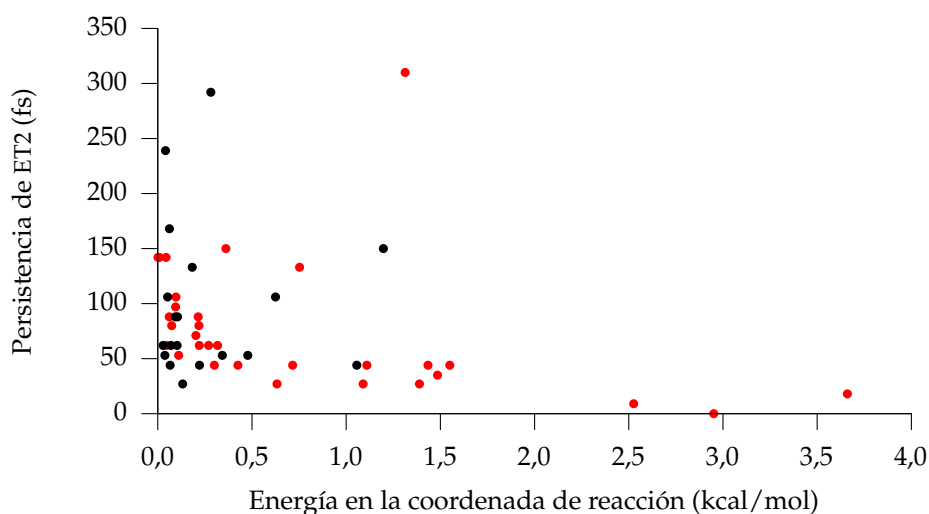
separación entre triazeno y agua. Finalmente, los dos reactivos vuelven a separarse, en este caso la trayectoria es no reactiva.



**Figura 12.16:** Ejemplo de trayectoria en la que se observa un tiempo de persistencia prolongado de la estructura del estado de transición. En negro: distancia  $N^3-H^6$ . En rojo: distancia  $N^1-H^9$ .

En casi todas las trayectorias se observa un intervalo de este tipo, en el que se mantiene la estructura de transición, más o menos prolongado. En la mayoría de los casos la duración de este estado es de unos 50 fs, que corresponde a unas 3 o 4 vibraciones de los enlaces N-H, aunque hay unas pocas trayectorias en las que la transferencia es casi inmediata y algunas más en la que la persistencia del estado de transición es mayor. En promedio, el tiempo de vida de estas estructuras de transición es de alrededor de 85 fs. No parece haber ninguna relación directa entre las energías en la coordenada de reacción y la duración del intervalo de persistencia, representadas en la figura 12.17; no obstante, las trayectorias en las que este intervalo es más corto tienden a ser reactivas y las tres trayectorias con mayor energía (más de 2,0 kcal/mol) presentan todas un intervalo menor de 20 fs.

Esta característica de las trayectorias pone de manifiesto la posible influencia del comportamiento dinámico del disolvente sobre el transcurso de la reacción. La localización del estado de transición se ha hecho sobre una superficie de energía libre en la que se supone, en todo momento, un equilibrio entre soluto y disolvente. Sin embargo, cuando existen importantes efectos dinámicos, el soluto puede quedar «atrapado» con la estructura del estado de transición, a la espera de que los movimientos del disolvente hagan posible su evolución hacia reactivos o productos. De esta manera, el estado de transición en la superficie de energía libre puede convertirse en un mínimo para una determinada trayectoria.



**Figura 12.17:** Representación del tiempo aproximado de persistencia del estado de transición frente a la energía cinética introducida en la coordenada de reacción para las trayectorias simuladas. En negro: trayectorias no reactivas. En rojo: trayectorias reactivas.

## 12.4. Conclusiones

En este trabajo se ha llevado a cabo el estudio de una reacción en disolución combinando dos métodos mixtos: ASEP/MD y DFT/MM. Se ha mostrado cómo puede emplearse uno y otro método para obtener distintos tipos de propiedades. Con ASEP/MD es sencillo calcular propiedades termodinámicas y estructurales promedio, mientras que con DFT/MM puede examinarse con detalle el comportamiento dinámico del sistema. Además, dado que es posible emplear casi la misma combinación de modelos cuánticos y clásicos en ambos métodos, la transferencia de resultados entre uno y otro resulta prácticamente inmediata y representa pocos problemas.

En cuanto al sistema concreto estudiado, la isomerización del triazeno, pueden señalarse las siguientes conclusiones particulares:

- La participación de una molécula de agua en el proceso de desplazamiento protónico reduce de manera muy importante la energía de activación.
- El efecto del disolvente es algo menor, pero también es significativo. Además, en disolución acuosa tiende a favorecerse una estructura con un alto grado de separación de cargas en el estado de transición. Esto puede indicar una preferencia por mecanismos iónicos para la reacción en lugar del estudiado, más concertado.

- El valor obtenido para el coeficiente de transmisión,  $\kappa$ , significativamente menor que la unidad, junto la persistencia de la estructura de transición observada en muchos casos, hace pensar que los efectos dinámicos del disolvente pueden ser importantes en esta reacción.

Finalmente, para completar el estudio sería interesante incluir otros posibles mecanismos de reacción que pueden tener lugar en el sistema real, y un adecuado tratamiento de otros factores como el efecto túnel. Sin embargo, con el trabajo realizado se ha puesto a punto una forma de combinar dos métodos mixtos diferentes para el estudio de reacciones en disolución, que era el objetivo primordial planteado.

## CONCLUSIONES GENERALES

Durante el desarrollo del presente trabajo se ha elaborado un programa informático en el lenguaje Fortran 90. Este programa permite aplicar el método ASEP/MD a los sistemas químicos de interés, incluyendo las ampliaciones y modificaciones que se han propuesto. Las principales características de este método son las siguientes:

- El soluto se representa mediante los métodos cuánticos usuales, lo que permite el estudio de efectos cuánticos, estados excitados, ruptura y formación de enlaces, etc.
- Las moléculas del disolvente se representan con los métodos de la mecánica molecular y el conjunto del sistema se somete a simulaciones de dinámica molecular. De esta manera se permite el estudio de las propiedades termodinámicas de la disolución.
- Gracias a la aproximación del campo medio (ACM), el número de cálculos cuánticos a realizar se reduce considerablemente con respecto a otros métodos que combinan mecánica cuántica y mecánica molecular (QM/MM).
- A lo largo del proceso ASEP/MD, la distribución electrónica del soluto y la estructura del disolvente a su alrededor se equilibran mutuamente. En el caso de que se optimice la geometría del soluto, también ésta estará en equilibrio con aquéllas.

El buen funcionamiento tanto del método como del programa informático ha sido ampliamente comprobado mediante las aplicaciones presentadas en este y en

otros trabajos publicados. Más concretamente, de los resultados expuestos en los capítulos precedentes se pueden destacar las siguientes conclusiones generales:

1. La aproximación del campo medio por sí sola no introduce errores significativos en los resultados. En los diversos sistemas estudiados se observa que el error cometido en la energía de interacción soluto-disolvente y en el gradiente sobre la superficie de energía libre es menor del 5 %. El error en el momento dipolar es prácticamente despreciable, menor del 1 %.
2. El método desarrollado para la obtención de geometrías moleculares optimizadas en disolución proporciona buenos resultados, tanto para estructuras estables como de transición. Estos resultados son comparables a los obtenidos con otros métodos computacionalmente más costosos. Además, el empleo de la ACM permite localizar los puntos estacionarios en la superficie de energía libre sin necesidad de restringir los grados de libertad del soluto.
3. El procedimiento empleado para el cálculo de diferencias de energía libre da resultados satisfactorios. Junto con el método de optimización de geometrías, este procedimiento permite el análisis termodinámico de equilibrios químicos y reacciones en disolución.
4. En todos los sistemas estudiados se ha puesto de manifiesto la importancia de las interacciones electrostáticas entre soluto y disolvente. Esto confirma la necesidad de obtener una adecuada descripción de las distribuciones de carga tanto del soluto como del disolvente.

### 13.1. Desarrollo futuro

El trabajo expuesto en esta tesis se enmarca dentro de la investigación seguida por nuestro grupo y dirigida, fundamentalmente, al estudio de los efectos del disolvente y al desarrollo y aplicación del método ASEP/MD. Durante los últimos años se ha logrado un importante avance en este sentido y se ha obtenido un éxito significativo. Sin embargo, hay algunos aspectos y propiedades que, a pesar de su interés, aún no han sido estudiados en profundidad. No obstante, el trabajo realizado representa un paso importante y facilita el desarrollo futuro de estas aplicaciones, entre las que se incluyen:

- Extender el método ASEP/MD para permitir la representación cuántica simultánea de las moléculas de soluto y disolvente. Esto posibilitaría el tratamiento de disoluciones concentradas.



- Aplicar el método FEG a la determinación de caminos de reacción, lo que permitiría la aplicación de la «Teoría Variacional del Estado de Transición» (*Variational Transition State Theory*, VTST) [275, 276].
- Continuar con el desarrollo de metodologías conjuntas de ASEP/MD y otros métodos QM/MM para el estudio de reacciones en disolución.
- Encontrar métodos y modelos que permitan mejorar la representación del soluto durante las simulaciones de dinámica molecular.
- Extender el método ASEP/MD al estudio de sistemas y procesos enzimáticos.



---

---

# APÉNDICES

---

---

*La física está haciéndose demasiado complicada para los físicos*

*David Hilbert (1862-1943)*



---

---

# A

---

## AJUSTE DE LAS CARGAS AL POTENCIAL ELECTROSTÁTICO

Dado un potencial electrostático numérico en una serie de puntos,  $\mathbf{r}_i = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m\}$ , se desea obtener un conjunto de cargas puntuales situadas en posiciones fijas  $\mathbf{s}_a = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$ , con valores  $q_a = \{q_1, q_2, \dots, q_n\}$ , de manera que el potencial generado por dichas cargas en los puntos  $\mathbf{r}_i$  se ajuste lo más exactamente posible al potencial numérico. Llamamos  $V_i$  al potencial en el punto  $\mathbf{r}_i$  y  $V_i^q$  al potencial generado por el conjunto de cargas  $q_a$  en dicho punto. En general, dado que el número de puntos suele ser mayor que el de cargas ( $m > n$ ), el ajuste no puede ser perfecto y ambos potenciales serán diferentes. Se procede entonces a realizar un ajuste por mínimos cuadrados de las cargas al potencial electrostático, el objetivo es minimizar la diferencia entre el potencial calculado,  $V_i^q$ , y el dado,  $V_i$ , expresada como la suma de las diferencias al cuadrado,  $D$ :

$$D = \sum_{i=1}^m \left( V_i^q - V_i \right)^2 \quad V_i^q = \sum_{a=1}^n \frac{q_a}{d_{ai}} \quad d_{ai} = |\mathbf{r}_i - \mathbf{s}_a| \quad (\text{A.1})$$

donde, para evitar el factor  $1/4\pi\epsilon_0$ , se emplean unidades atómicas. Se incluye también la restricción de que la suma de todas las cargas ajustadas sea igual a  $Q$ , cuyo valor dependerá del sistema concreto, normalmente  $Q = 0$ :

$$\sum_{a=1}^n q_a = Q \quad \Rightarrow \quad C = \sum_{a=1}^n q_a - Q = 0 \quad (\text{A.2})$$

Para minimizar esta diferencia,  $D$ , se emplea el método de los multiplicadores de Lagrange, utilizando  $2\lambda$  como multiplicador. En primer lugar se construye una

función  $F(q, \lambda) = D + 2\lambda C$ :

$$\begin{aligned} F(q, \lambda) &= \sum_{i=1}^m \left( \sum_{a=1}^n \frac{q_a}{d_{ai}} - V_i \right)^2 + 2\lambda \left( \sum_{a=1}^n q_a - Q \right) = \\ &= \sum_{i=1}^m \sum_{a=1}^n \sum_{b=1}^n \frac{q_a q_b}{d_{ai} d_{bi}} - 2 \sum_{i=1}^m \sum_{a=1}^n \frac{q_a V_i}{d_{ai}} + \sum_{i=1}^m V_i^2 + 2\lambda \sum_{a=1}^n q_a - 2\lambda Q \end{aligned} \quad (\text{A.3})$$

para después igualar a cero sus derivadas con respecto a las variables  $q_a$  y  $\lambda$ :

$$\begin{aligned} \frac{\partial F(q, \lambda)}{\partial q_a} &= 2 \sum_{i=1}^m \sum_{b=1}^n \frac{q_b}{d_{ai} d_{bi}} - 2 \sum_{i=1}^m \frac{V_i}{d_{ai}} + 2\lambda = 0 \\ \forall a \quad \sum_{b=1}^n \left( q_b \sum_{i=1}^m \frac{1}{d_{ai} d_{bi}} \right) + \lambda &= \sum_{i=1}^m \frac{V_i}{d_{ai}} \end{aligned} \quad (\text{A.4a})$$

$$\frac{\partial F(q, \lambda)}{\partial \lambda} = 2 \sum_{a=1}^n q_a - 2Q = 0 \quad \sum_{a=1}^n q_a = \sum_{b=1}^n q_b = Q \quad (\text{A.4b})$$

Estas ecuaciones se pueden expresar fácilmente en forma matricial:

$$\begin{pmatrix} e_{11} & e_{12} & \cdots & e_{1n} & 1 \\ e_{21} & e_{22} & \cdots & e_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ e_{n1} & e_{n2} & \cdots & e_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \\ \lambda \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \\ Q \end{pmatrix} \quad \begin{aligned} e_{ab} &= \sum_{i=1}^m \frac{1}{d_{ai} d_{bi}} \\ t_a &= \sum_{i=1}^m \frac{V_i}{d_{ai}} \end{aligned} \quad (\text{A.5})$$

de manera que, resolviendo este sistema de ecuaciones lineales por cualquiera de los métodos al uso, se obtiene el conjunto de cargas  $q_a$ , situadas en los puntos  $s_a$ , que mejor reproduce el potencial numérico dado en los puntos  $r_i$ .

---

---

# B

---

## DISOLVENTE POLARIZABLE

En el caso de realizar un cálculo con disolvente polarizable, es necesario calcular el momento dipolar inducido en cada una de las moléculas de disolvente. Utilizaremos los subíndices  $\{i, j\}$  para designar a las distintas moléculas de disolvente y los subíndices  $\{a, b\}$  para las cargas (generalmente situadas en los núcleos atómicos) dentro de cada molécula, que denominamos cargas permanentes. Las cargas de la molécula de soluto se designan con el subíndice  $s$ .

### B.1. Cálculo de dipolos inducidos

El momento dipolar inducido en la molécula  $i$ ,  $\mu_i$ , se calcula como:

$$\mu_i = \alpha_i E_i = \alpha_i (E_i^q + E_i^\mu) \quad (\text{B.1})$$

donde  $\alpha_i$  es el tensor de polarizabilidad de la molécula  $i$ ,  $E_i$  es el campo eléctrico generado por todas las otras moléculas (de disolvente y de soluto) en el centro de masas de la molécula  $i$ ,  $E_i^q$  es el campo eléctrico generado únicamente por las cargas permanentes, mientras que  $E_i^\mu$  es el generado por los momentos dipolares inducidos.

El campo generado por las cargas permanentes es sencillo de calcular (en unidades atómicas):

$$E_i^q = \sum_{j \neq i} \sum_a \frac{q_{a_j} \mathbf{r}_{a_j i}}{r_{a_j i}^3} + \sum_s \frac{q_s \mathbf{r}_{s i}}{r_{s i}^3} \quad (\text{B.2})$$

donde  $q_{a_j}$  es el valor de la carga  $a$  de la molécula  $j$  y  $\mathbf{r}_{a_j i} = \mathbf{r}_i - \mathbf{r}_{a_j}$  es el vector

que conecta esta carga con el centro de masas de la molécula  $i$ . El primer sumando corresponde al campo eléctrico generado por el disolvente, el segundo corresponde al soluto.

Para calcular el campo generado por los dipolos inducidos, puede emplearse la fórmula:

$$\mathbf{E}_i^\mu = \sum_{j \neq i} T_{ji} \boldsymbol{\mu}_j \quad T_{ji} = \nabla \nabla \frac{1}{r_{ji}} = \frac{1}{r_{ji}^3} \left[ \frac{3 \mathbf{r}_{ji} \mathbf{r}_{ji}^T}{r_{ji}^2} - I \right] \quad (\text{B.3})$$

donde  $\mathbf{r}_{ji} = \mathbf{r}_i - \mathbf{r}_j$  es el vector que conecta los centros de masas de las moléculas  $i$  y  $j$ .  $I$  es la matriz unidad de orden 3. En esta fórmula se considera que el dipolo es puntual, una aproximación que sólo es aplicable cuando la distancia entre las moléculas  $i$  y  $j$  es suficientemente grande. A menores distancias, esta aproximación pierde validez y es necesario corregir los resultados, para lo cual se modifica la expresión del tensor  $T_{ji}$  [277]:

$$T_{ji} = \frac{f_{ji}^3}{r_{ji}^3} \left[ \frac{3 f_{ji} \mathbf{r}_{ji} \mathbf{r}_{ji}^T}{r_{ji}^2} - (4 - 3 f_{ji}) I \right] \quad (\text{B.4})$$

$$f_{ji} = \min \left( 1, \frac{r_{ji}}{3,1407 a_0 \sqrt[3]{\frac{1}{3} \text{tr}(\boldsymbol{\alpha}_i)}} \right)$$

donde la función  $\min(x, y)$  toma el valor del menor de sus dos argumentos,  $x$  e  $y$ ,  $\text{tr}(\boldsymbol{\alpha}_i)$  es la traza (suma de los valores diagonales) del tensor polarizabilidad de la molécula  $i$  y  $a_0$  es el radio de Bohr, tomado como la unidad atómica de longitud. De esta manera, a grandes distancias  $f = 1$  y la expresión de  $T_{ji}$  es equivalente a la de la ecuación (B.3).

Según las fórmulas anteriores, para el cálculo de cada  $\boldsymbol{\mu}_i$  es necesario conocer  $\mathbf{E}^\mu$  (ec. B.1), y para el cálculo de  $\mathbf{E}^\mu$  es necesario conocer el resto de los  $\boldsymbol{\mu}_j$  (ec. B.3), lo que obliga a realizar un proceso iterativo, calculando los  $\boldsymbol{\mu}_i$  repetidamente hasta llegar a la convergencia.

## B.2. Energías de interacción

La energía de interacción de este conjunto de cargas y dipolos puede descomponerse en varios términos:  $E = E_{\text{qq}} + E_{\text{q}\mu} + E_{\mu\mu} + E^{\text{dist}}$ , donde  $E_{\text{qq}}$  es la energía de interacción entre las cargas permanentes,  $E_{\text{q}\mu}$  es la energía de interacción entre los dipolos inducidos y las cargas permanentes,  $E_{\mu\mu}$  es la energía de interacción de los dipolos inducidos entre sí y  $E^{\text{dist}}$  es la energía necesaria para crear los dipolos



inducidos (energía de distorsión). Sus expresiones son:

$$\begin{aligned}
 E_{\text{qq}} &= \frac{1}{2} \sum_i \sum_a q_{a_i} V_{a_i}^{\text{q}} & E_{\text{q}\mu} &= - \sum_i \boldsymbol{\mu}_i^{\text{T}} \mathbf{E}_i^{\text{q}} & E_{\mu\mu} &= -\frac{1}{2} \sum_i \boldsymbol{\mu}_i^{\text{T}} \mathbf{E}_i^{\mu} \\
 E^{\text{dist}} &= \frac{1}{2} \sum_i \boldsymbol{\mu}_i^{\text{T}} \boldsymbol{\alpha}_i^{-1} \boldsymbol{\mu}_i = \frac{1}{2} \sum_i \boldsymbol{\mu}_i^{\text{T}} (\mathbf{E}_i^{\text{q}} + \mathbf{E}_i^{\mu}) = -\frac{1}{2} E_{\text{q}\mu} - E_{\mu\mu} \\
 E &= E_{\text{qq}} + E_{\text{q}\mu} + E_{\mu\mu} + E^{\text{dist}} = E_{\text{qq}} + \frac{1}{2} E_{\text{q}\mu}
 \end{aligned} \tag{B.5}$$

donde  $V_{a_i}^{\text{q}}$  es el potencial electrostático generado en la posición de la carga  $q_{a_i}$  por las cargas permanentes del resto de las moléculas:

$$V_{a_i}^{\text{q}} = \sum_{j \neq i} \sum_b \frac{q_{b_j}}{r_{b_j i}} \tag{B.6}$$

De esta manera, la contribución energética de la polarización del disolvente, una vez simplificados los términos que implican a los dipolos inducidos, se reduce a  $\frac{1}{2} E_{\text{q}\mu}$ .

Hasta aquí se ha considerado el soluto, al igual que el disolvente, como un conjunto de cargas puntuales. Sin embargo, en el método ASEP/MD se representa el soluto mecanocuánticamente, mediante su densidad electrónica  $\rho$ . Por otra parte, los momentos dipolares inducidos de las moléculas de disolvente ya no se representan explícitamente, sino a través de las cargas que reproducen el potencial correspondiente: las cargas  $q_i^0$  representan las cargas permanentes y  $q_i^{\mu}$  los dipolos inducidos. En este caso las energías de interacción vienen dadas por:

$$\begin{aligned}
 E_{\text{qq}} &= \frac{1}{2} \sum_i q_i^0 V_i^{\text{q}} & E_{\text{q}\mu} &= - \sum_i \boldsymbol{\mu}_i^{\text{T}} \mathbf{E}_i^{\text{q}} & E_{\mu\mu} &= -\frac{1}{2} \sum_i \boldsymbol{\mu}_i^{\text{T}} \mathbf{E}_i^{\mu} \\
 E_{\rho\text{q}} &= \sum_i q_i^0 V_i^{\rho} & E_{\rho\mu} &= - \sum_i \boldsymbol{\mu}_i^{\text{T}} \mathbf{E}_i^{\rho} \simeq \sum_i q_i^{\mu} V_i^{\rho} \\
 E_{\mu}^{\text{dist}} &= \frac{1}{2} \sum_i \boldsymbol{\mu}_i^{\text{T}} (\mathbf{E}_i^{\text{q}} + \mathbf{E}_i^{\mu} + \mathbf{E}_i^{\rho}) = -\frac{1}{2} E_{\text{q}\mu} - E_{\mu\mu} - \frac{1}{2} E_{\rho\mu} \\
 E &= E_{\text{qq}} + E_{\text{q}\mu} + E_{\mu\mu} + E_{\rho\text{q}} + E_{\rho\mu} + E_{\mu}^{\text{dist}} + E_{\rho}^{\text{dist}} = \\
 &= E_{\text{qq}} + E_{\rho\text{q}} + E_{\rho}^{\text{dist}} + \frac{1}{2} E_{\text{q}\mu} + \frac{1}{2} E_{\rho\mu}
 \end{aligned} \tag{B.7}$$

En estas ecuaciones,  $V_i^{\rho}$  es el potencial electrostático creado por el soluto en el punto  $i$ ,  $\mathbf{E}_i^{\rho}$  sería el campo eléctrico generado por el soluto, aunque en la práctica  $E_{\rho\mu}$  se calcula a partir del potencial  $V_i^{\rho}$ . Finalmente,  $E_{\rho}^{\text{dist}}$  es la energía de distorsión o de polarización del soluto, obtenida como:

$$E_{\rho}^{\text{dist}} = \langle \Psi | \hat{H}_{\text{QM}} | \Psi \rangle - \langle \Psi^0 | \hat{H}_{\text{QM}} | \Psi^0 \rangle \tag{B.8}$$



---

---

# C

---

## ENERGÍA LIBRE EN LOS MÉTODOS QM/MM

Cuando se estudia un sistema mediante un método combinado de mecánica cuántica y mecánica molecular, ya sea un método QM/MM tradicional o el método ASEP/MD, puede observarse una división clara entre los grados de libertad pertenecientes a la parte del sistema descrita mediante mecánica cuántica,  $\mathbf{r}_{\text{QM}}$ , y la descrita mediante mecánica molecular,  $\mathbf{r}_{\text{MM}}$ . Para la energía potencial del sistema se puede establecer una partición sencilla, de la siguiente manera:

$$E(\mathbf{r}_{\text{QM}}, \mathbf{r}_{\text{MM}}) = E_{\text{QM}}(\mathbf{r}_{\text{QM}}) + E_{\text{MM}}(\mathbf{r}_{\text{MM}}) + E_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}, \mathbf{r}_{\text{MM}}) \quad (\text{C.1})$$

Para la energía libre, sin embargo, se sigue el desarrollo propuesto por Zhang y col. [155]. Se considera una «coordenada de reacción» más o menos arbitraria, que puede expresarse en función de las coordenadas del soluto,  $R(\mathbf{r}_{\text{QM}})$ , lo que permite trabajar con integrales unidimensionales. La integral configuracional puede escribirse como:

$$\begin{aligned} Z &= \int Z(r_{\text{R}}) dr_{\text{R}} \\ Z(r_{\text{R}}) &= \int \delta[R(\mathbf{r}_{\text{QM}}) - r_{\text{R}}] e^{-\beta E(\mathbf{r}_{\text{QM}}, \mathbf{r}_{\text{MM}})} d\mathbf{r}_{\text{QM}} d\mathbf{r}_{\text{MM}} \end{aligned} \quad (\text{C.2})$$

siendo  $r_{\text{R}}$  el valor de la coordenada de reacción y  $\delta(x - x_0)$  la función delta de Dirac.  $Z(r_{\text{R}})$  es, pues, la integral configuracional restringida a los casos en los que la coordenada de reacción toma el valor  $r_{\text{R}}$ .

Para cada posible estructura del soluto,  $\mathbf{r}_{\text{QM}}$ , puede definirse un «potencial de fuerza media» (*Potential of Mean Force*, PMF) sobre la porción QM por integración

de las fluctuaciones en la porción MM del sistema:

$$A_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}) = -k_{\text{B}}T \ln \int e^{-\beta[E_{\text{MM}}(\mathbf{r}_{\text{MM}}) + E_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}, \mathbf{r}_{\text{MM}})]} d\mathbf{r}_{\text{MM}} \quad (\text{C.3})$$

En estas condiciones, la integral configuracional  $Z(r_{\text{R}})$  puede expresarse como:

$$Z(r_{\text{R}}) = \int \delta[r_{\text{R}} - R(\mathbf{r}_{\text{QM}})] e^{-\beta[E_{\text{QM}}(\mathbf{r}_{\text{QM}}) + A_{\text{QM/MM}}(\mathbf{r}_{\text{QM}})]} d\mathbf{r}_{\text{QM}} \quad (\text{C.4})$$

A continuación, de todos los puntos en los que  $R(\mathbf{r}_{\text{QM}}) = r_{\text{R}}$ , se considera uno determinado,  $\mathbf{r}_{\text{QM}}^*$ , que se tomará como referencia. La integral configuracional puede expresarse en función de las energías en este punto:

$$\begin{aligned} F(\mathbf{r}_{\text{QM}}) &= E_{\text{QM}}(\mathbf{r}_{\text{QM}}) + A_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}) \\ Z(r_{\text{R}}) &= e^{-\beta F(\mathbf{r}_{\text{QM}}^*)} \int \delta[r_{\text{R}} - R(\mathbf{r}_{\text{QM}})] e^{-\beta[F(\mathbf{r}_{\text{QM}}) - F(\mathbf{r}_{\text{QM}}^*)]} d\mathbf{r}_{\text{QM}} \end{aligned} \quad (\text{C.5})$$

y la energía libre para  $r_{\text{R}}$  vendrá dada por:

$$A(r_{\text{R}}) = F(\mathbf{r}_{\text{QM}}^*) - k_{\text{B}}T \ln \int \delta[r_{\text{R}} - R(\mathbf{r}_{\text{QM}})] e^{-\beta[F(\mathbf{r}_{\text{QM}}) - F(\mathbf{r}_{\text{QM}}^*)]} d\mathbf{r}_{\text{QM}} \quad (\text{C.6})$$

El último término de la expresión anterior contiene únicamente las fluctuaciones que tienen lugar en la porción QM del sistema, alrededor del punto  $\mathbf{r}_{\text{QM}}^*$ . Si  $\mathbf{r}_{\text{QM}}^*$  minimiza la energía del sistema para un determinado valor de  $r_{\text{R}}$ , a temperatura cero dicho término se anularía. En cualquier caso puede hacerse una aproximación importante, como es suponer que, entre dos puntos  $\mathbf{r}_{\text{QM}}^{\text{A}}$  y  $\mathbf{r}_{\text{QM}}^{\text{B}}$ , el término de las fluctuaciones se mantiene constante y que, por lo tanto, la diferencia de energía libre puede calcularse como:

$$\begin{aligned} \Delta A_{\text{A} \rightarrow \text{B}} &\simeq F(\mathbf{r}_{\text{QM}}^{\text{B}}) - F(\mathbf{r}_{\text{QM}}^{\text{A}}) = \\ &= [E_{\text{QM}}(\mathbf{r}_{\text{QM}}^{\text{B}}) + A_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}^{\text{B}})] - [E_{\text{QM}}(\mathbf{r}_{\text{QM}}^{\text{A}}) + A_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}^{\text{A}})] = \\ &= \Delta E_{\text{QM}} + \Delta A_{\text{QM/MM}} \end{aligned} \quad (\text{C.7})$$

El término  $\Delta A_{\text{QM/MM}}$ , por su parte puede calcularse por FEP:

$$\Delta A_{\text{QM/MM}} = -k_{\text{B}}T \ln \left\langle e^{-\beta[E_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}^{\text{B}}) - E_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}^{\text{A}})]} \right\rangle_{\text{A}} \quad (\text{C.8})$$

donde  $\langle \dots \rangle_{\text{A}}$  representa un promedio de configuraciones en las que la porción QM se mantiene fija en  $\mathbf{r}_{\text{QM}}^{\text{A}}$ .

La expresión (C.7) se puede mejorar evaluando de manera aproximada el término de las fluctuaciones en (C.6). Para ello se calculan las frecuencias armónicas de la parte cuántica del sistema y se obtiene la aproximación armónica para la diferencia

de dicho término,  $\Delta V_{\text{QM}}$ :

$$\Delta A_{\text{A} \rightarrow \text{B}} \simeq \Delta E_{\text{QM}} + \Delta A_{\text{QM/MM}} + \Delta V_{\text{QM}} \quad (\text{C.9})$$

Por último, dado que para sistemas poco compresibles  $\Delta A$  y  $\Delta G$  —la energía libre de Gibbs— pueden considerarse equivalentes, estas fórmulas también son aplicables cuando se sustituye la primera por la segunda.



---

---

## D

---

# ACTUALIZACIÓN DE LA HESSIANA

Como se ha mencionado en el apartado 6.2.2 (pág 105), en los métodos cuasi-Newton se emplea una matriz  $H$  que sirve de aproximación para la inversa de la matriz hessiana,  $H \simeq G^{-1}$ . En cada iteración del proceso de optimización, a partir de la variación experimentada por el gradiente y del desplazamiento de la geometría, se actualiza esta matriz. Existen varias fórmulas para actualizar  $H$ , pero probablemente las dos más conocidas son la de Davidon-Fletcher-Powell (DFP) [165-167] y la Broyden-Fletcher-Goldfarb-Shanno (BFGS) [168-172], cuyas expresiones son:

$$\delta = \mathbf{r}_{k+1} - \mathbf{r}_k \quad \gamma = \mathbf{g}_{k+1} - \mathbf{g}_k$$

$$H_{k+1}^{\text{BFGS}} = H_k + \left(1 + \frac{\gamma^T H_k \gamma}{\delta^T \gamma}\right) \frac{\delta \delta^T}{\delta^T \gamma} - \left(\frac{\delta \gamma^T H_k + H_k \gamma \delta^T}{\delta^T \gamma}\right) \quad (\text{D.1a})$$

$$H_{k+1}^{\text{DFP}} = H_k + \frac{\delta \delta^T}{\delta^T \gamma} - \frac{H_k \gamma \gamma^T H_k}{\gamma^T H_k \gamma} \quad (\text{D.1b})$$

Estas fórmulas, en la mayoría de los casos prácticos, tienen la propiedad de que si  $H_k$  es «definida» (es decir, si todos sus valores propios tienen el mismo signo y ninguno es nulo),  $H_{k+1}$  mantiene esta propiedad. En concreto, si la matriz inicial,  $H_0$ , es definida positiva, sucesivas iteraciones de cualquiera de las fórmulas anteriores generarán únicamente matrices definidas positivas, por lo que se asegura la estabilidad de los métodos de optimización basados en dichas fórmulas cuando se emplean para la localización de estructuras estables. De manera similar, se pueden obtener fórmulas parecidas para la matriz  $B = H^{-1}$ , que sirve de aproximación

para la hessiana,  $G$ :

$$B_{k+1}^{\text{BFGS}} = B_k + \frac{\gamma\gamma^T}{\gamma^T\delta} - \frac{B_k\delta\delta^TB_k}{\delta^TB_k\delta} \quad (\text{D.2a})$$

$$B_{k+1}^{\text{DFP}} = B_k + \left(1 + \frac{\delta^TB_k\delta}{\gamma^T\delta}\right) \frac{\gamma\gamma^T}{\gamma^T\delta} - \left(\frac{\gamma\delta^TB_k + B_k\delta\gamma^T}{\gamma^T\delta}\right) \quad (\text{D.2b})$$

Puede observarse que existe una simetría entre las dos fórmulas, de hecho, si se intercambian  $H \leftrightarrow B$  y  $\delta \leftrightarrow \gamma$ , las fórmulas BFGS y DFP resultan iguales, se dice que son fórmulas duales o complementarias. En la práctica, la fórmula BFGS (y el método cuasi-Newton derivado de ella) suele funcionar mejor que la DFP.

Puede desarrollarse una fórmula más general para la actualización de la matriz hessiana:

$$B_{k+1} = B_k + ju^T + uj^T - (j^T\delta)uu^T$$

$$j = \gamma - B_k\delta \quad u = \frac{M\delta}{\delta^TM\delta} \quad M\delta = a\gamma + bB_k\delta \quad (\text{D.3})$$

Se obtienen diferentes fórmulas de actualización según el valor que se dé a  $a$  y  $b$  (manteniendo  $a + b = 1$ ). Tomando  $a = 1$  resulta la fórmula DFP, la fórmula BFGS se obtiene con:

$$a = \frac{(\delta^TB_k\delta)^{1/2}}{(\delta^T\gamma)^{1/2} + (\delta^TB_k\delta)^{1/2}} \quad b = \frac{(\delta^T\gamma)^{1/2}}{(\delta^T\gamma)^{1/2} + (\delta^TB_k\delta)^{1/2}} \quad (\text{D.4})$$

Otras dos fórmulas útiles son la de Powell ( $M = I$ ) y la fórmula BFGS modificada propuesta por Prat-Resina y col. [231], en la que:

$$a = \frac{(\delta^TB_k\gamma)^2}{(\delta^TB_kB_k\delta)(\gamma^T\gamma)} \quad b = 1 - a \quad (\text{D.5})$$

En general, todas estas fórmulas de actualización de la hessiana son convergentes: durante el curso de la optimización, la matriz  $B$  (o  $H$ ) se acerca cada vez más al valor correcto,  $G$  (o  $G^{-1}$ ).



---

---

# E

---

## DERIVADAS DE LA ENERGÍA LIBRE

Para obtener las derivadas sobre la superficie de energía libre es necesario calcular las derivadas de  $A_{\text{QM/MM}}$  con respecto a las coordenadas  $\mathbf{r}_{\text{QM}}$ . Se parte de la ecuación:

$$\begin{aligned} A_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}) &= -k_{\text{B}}T \ln Z_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}) \\ Z_{\text{QM/MM}}(\mathbf{r}_{\text{QM}}) &= \int e^{-\beta[E_{\text{MM}}(\mathbf{r}_{\text{MM}})+E_{\text{QM/MM}}(\mathbf{r}_{\text{MM}},\mathbf{r}_{\text{QM}})]} d\mathbf{r}_{\text{MM}} \end{aligned} \quad (\text{E.1})$$

En lo sucesivo, por sencillez, no se explicitan las dependencias respecto a  $\mathbf{r}_{\text{MM}}$  o  $\mathbf{r}_{\text{QM}}$ .

Para la primera derivada:

$$\frac{\partial A_{\text{QM/MM}}}{\partial \mathbf{r}_{\text{QM}}} = -k_{\text{B}}T \frac{Z'_{\text{QM/MM}}}{Z_{\text{QM/MM}}} = \frac{\int E'_{\text{QM/MM}} e^{-\beta(E_{\text{MM}}+E_{\text{QM/MM}})} d\mathbf{r}_{\text{MM}}}{\int e^{-\beta(E_{\text{MM}}+E_{\text{QM/MM}})} d\mathbf{r}_{\text{MM}}} \quad (\text{E.2})$$

Se multiplica el numerador y el denominador por  $e^{-\beta E_{\text{QM}}}$  y, dado que  $E_{\text{QM}}$  no depende de  $\mathbf{r}_{\text{MM}}$ , se introduce en la integral. Haciendo  $E = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QM/MM}}$ , queda la expresión:

$$\frac{\partial A_{\text{QM/MM}}}{\partial \mathbf{r}_{\text{QM}}} = \frac{\int E'_{\text{QM/MM}} e^{-\beta E} d\mathbf{r}_{\text{MM}}}{\int e^{-\beta E} d\mathbf{r}_{\text{MM}}} = \langle E'_{\text{QM/MM}} \rangle = \left\langle \frac{\partial E_{\text{QM/MM}}}{\partial \mathbf{r}_{\text{QM}}} \right\rangle \quad (\text{E.3})$$

es decir, la contribución QM/MM al gradiente de la energía libre puede calcularse como el promedio del gradiente de la interacción soluto-disolvente, promedio

realizado sobre todas las configuraciones del disolvente consideradas.

La segunda derivada, por su parte, se obtiene a partir de la anterior:

$$\begin{aligned}
\frac{\partial^2 A_{\text{QM/MM}}}{\partial r_{\text{QM}}^2} &= \frac{\partial}{\partial r_{\text{QM}}} \frac{\partial A_{\text{QM/MM}}}{\partial r_{\text{QM}}} = \frac{\partial}{\partial r_{\text{QM}}} \left( \frac{\int E'_{\text{QM/MM}} e^{-\beta E} \mathbf{d}r_{\text{MM}}}{\int e^{-\beta E} \mathbf{d}r_{\text{MM}}} \right) = \\
&= \frac{1}{\int e^{-\beta E} \mathbf{d}r_{\text{MM}}} \frac{\partial}{\partial r_{\text{QM}}} \left( \int E'_{\text{QM/MM}} e^{-\beta E} \mathbf{d}r_{\text{MM}} \right) + \\
&\quad + \int E'_{\text{QM/MM}} e^{-\beta E} \mathbf{d}r_{\text{MM}} \frac{\partial}{\partial r_{\text{QM}}} \left( \frac{1}{\int e^{-\beta E} \mathbf{d}r_{\text{MM}}} \right) = \\
&= \frac{\int E''_{\text{QM/MM}} e^{-\beta E} \mathbf{d}r_{\text{MM}} - \beta \int E'_{\text{QM/MM}} (E'_{\text{QM/MM}} + E'_{\text{QM}}) e^{-\beta E} \mathbf{d}r_{\text{MM}}}{\int e^{-\beta E} \mathbf{d}r_{\text{MM}}} + \\
&\quad + \frac{\beta \int E'_{\text{QM/MM}} e^{-\beta E} \mathbf{d}r_{\text{MM}} \int (E'_{\text{QM/MM}} + E'_{\text{QM}}) e^{-\beta E} \mathbf{d}r_{\text{MM}}}{\int e^{-\beta E} \mathbf{d}r_{\text{MM}} \int e^{-\beta E} \mathbf{d}r_{\text{MM}}} = \\
&= \frac{\int E''_{\text{QM/MM}} e^{-\beta E} \mathbf{d}r_{\text{MM}}}{\int e^{-\beta E} \mathbf{d}r_{\text{MM}}} - \beta \frac{\int E'_{\text{QM/MM}} (E'_{\text{QM/MM}} + E'_{\text{QM}}) e^{-\beta E} \mathbf{d}r_{\text{MM}}}{\int e^{-\beta E} \mathbf{d}r_{\text{MM}}} + \\
&\quad + \beta \frac{\int E'_{\text{QM/MM}} e^{-\beta E} \mathbf{d}r_{\text{MM}} \int (E'_{\text{QM/MM}} + E'_{\text{QM}}) e^{-\beta E} \mathbf{d}r_{\text{MM}}}{\int e^{-\beta E} \mathbf{d}r_{\text{MM}} \int e^{-\beta E} \mathbf{d}r_{\text{MM}}} = \\
&= \langle E''_{\text{QM/MM}} \rangle - \beta \langle E'_{\text{QM/MM}} E'_{\text{QM/MM}} \rangle - \beta \langle E'_{\text{QM/MM}} E'_{\text{QM}} \rangle + \\
&\quad + \beta \langle E'_{\text{QM/MM}} \rangle \langle E'_{\text{QM/MM}} \rangle + \beta \langle E'_{\text{QM/MM}} \rangle \langle E'_{\text{QM}} \rangle = \\
&= \langle E''_{\text{QM/MM}} \rangle - \beta \langle E'_{\text{QM/MM}} E'_{\text{QM/MM}} \rangle - \beta E'_{\text{QM}} \langle E'_{\text{QM/MM}} \rangle + \\
&\quad + \beta \langle E'_{\text{QM/MM}} \rangle \langle E'_{\text{QM/MM}} \rangle + \beta E'_{\text{QM}} \langle E'_{\text{QM/MM}} \rangle \tag{E.4}
\end{aligned}$$

donde  $E_{\text{QM}}$  se escribe fuera de los promedios porque no depende de  $r_{\text{MM}}$ ; por lo tanto, los términos en los que aparece se anulan:

$$\frac{\partial^2 A_{\text{QM/MM}}}{\partial r_{\text{QM}}^2} = \langle E''_{\text{QM/MM}} \rangle + \beta [\langle E'_{\text{QM/MM}} \rangle \langle E'_{\text{QM/MM}} \rangle - \langle E'_{\text{QM/MM}} E'_{\text{QM/MM}} \rangle] \tag{E.5}$$

El segundo sumando de la ecuación (E.5) representa las fluctuaciones térmicas del gradiente. Aplicando la aproximación del campo medio, pueden despreciarse

estas fluctuaciones, en cuyo caso se obtiene como segunda derivada:

$$\frac{\partial^2 A_{\text{QM/MM}}}{\partial r_{\text{QM}}^2} \simeq \langle E''_{\text{QM/MM}} \rangle = \left\langle \frac{\partial^2 E_{\text{QM/MM}}}{\partial r_{\text{QM}}^2} \right\rangle \quad (\text{E.6})$$

Al igual que ocurre en otros casos, estas fórmulas son también válidas para  $G_{\text{QM/MM}}$  en el caso de sistemas poco compresibles.



---

---

# F

---

## CÓDIGO DEL PROGRAMA ASEP/MD

### F.1. Principal.f90

```
!ASEP-MD 1.0.5-es

!08-10-2004: 1.0.6 - Corregir paso de Newton y permitir Masa=0:
!           CalcularIncremento
!04-06-2004: 1.0.5 - Corregir "bug" al restar el campo del soluto:
!           CalcularPotencial
!07-05-2003: 1.0.4 - Escribir la hessiana en cada ciclo:
!           OptimizarGeometria
!13-03-2003: 1.0.3 - No leer cuaterniones para monoatomicas:
!           CalcularPotencial
!20-02-2003: 1.0.2 - Corregir ActualizarHessiana (sobra -1 en bucle DO)
!
!23-01-2003: 1.0.2 - Annadir valor por defecto para MaxDifCargas
!           Permitir ConvEnergia y ConvCargas = 0: LeerEntrada
!22-01-2003: 1.0.2 - Admitir 'CHG' (CHELPG) como TipoCargas:
!           LeerEntrada,ModificarGaussian,LeerSalidaGaussian
!14-01-2003: 1.0.1 - Mantener la quiralidad en GirarMolecula

!CONTIENE:
!ASEP_MD (Programa principal)
!LeerEntrada
!LeerContinuacion
!EscribePrin(Iter,Cont)

!-----
! En este modulo se definen las variables globales del calculo
!-----

MODULE Parametros
  CHARACTER*100 :: DirDinamica,ProgAbInitio,FicheroDinamica,EntradaMoldy, &
    ProgDinamica,ContinuaMoldy,SalidaDinamica,FicheroDump,FicheroAbInitio, &
```

```

        SalidaAbInicio,EntradaGeneral,SalidaGeneral,FicheroCont,Extension, &
        DinamicaInicial,CargasFijas
CHARACTER :: AbInicio*5,Dinamica*5,TipoCargas*3,Acoplamiento*3,DipInd*3, &
        Polarizacion*2,HacerDinamica*2,PrimeraCapa*2,BusquedaLineal*2, &
        EstTransicion*2
INTEGER :: MaxIter,MaxIterPol,Inicio,MaxDumps,NumConfig,NumConfigIni, &
        NumConfigTot,DivMalla,Ciclo,CicloFin,MaxIterOpt,CalcHessiana
DOUBLE PRECISION :: ConvEnergia,ConvCargas,MaxDifCargas,FactorRadios, &
        RadioCapa,Proximidad,ConvGradiente,MinPaso,MaxPaso,Temperatura

NAMELIST /Entrada/ DirDinamica,ProgDinamica,ProgAbInicio,AbInicio, &
        Dinamica,FicheroCont,SalidaGeneral,FicheroAbInicio,SalidaAbInicio, &
        FicheroDinamica,SalidaDinamica,DinamicaInicial,HacerDinamica, &
        ConvEnergia,ConvCargas,MaxDifCargas,Inicio,MaxIter,MaxIterPol, &
        NumConfig,NumConfigIni,FactorRadios,DivMalla,PrimeraCapa,RadioCapa, &
        Proximidad,Acoplamiento,TipoCargas,Polarizacion,DipInd,CargasFijas, &
        MaxIterOpt,ConvGradiente,MinPaso,MaxPaso,CalcHessiana,BusquedaLineal, &
        EstTransicion

END MODULE Parametros

!-----
! En este modulo se define el sistema que se calculara
!-----
MODULE Sistema
    TYPE Atomo
        INTEGER :: id
        DOUBLE PRECISION :: x,y,z,q
        INTEGER :: NumAt
        DOUBLE PRECISION :: Masa
        CHARACTER :: Nombre*7
    END TYPE Atomo

    TYPE (Atomo), DIMENSION(:), ALLOCATABLE :: Solute,Disolvente,SoluteAnt
    INTEGER :: Carga,Multiplicidad,MoleculasSolute,MoleculasDisolvente
    DOUBLE PRECISION, DIMENSION(:,:,:), ALLOCATABLE :: Potencial
    DOUBLE PRECISION, DIMENSION(3,3) :: PolDisolvente
    CHARACTER :: TipoPotencial*13,NombreSolute*20,NombreDisolvente*20
END MODULE Sistema

!-----
! En este modulo aparecen las variables que solo se utilizan en la optimizacion
! de geometria
!-----
MODULE Optimizacion
    DOUBLE PRECISION, DIMENSION(:), ALLOCATABLE :: Gradiente,GradAnt,GradPot,&
        Coordenadas,Incremento
    DOUBLE PRECISION, DIMENSION(:,:), ALLOCATABLE :: Hessiana,HessPot
    DOUBLE PRECISION :: EnergPot,Paso
END MODULE Optimizacion

!-----
! En este modulo se definen las variables que almacenan los resultados de los
! calculos
!-----
MODULE Resultados
    DOUBLE PRECISION, DIMENSION(:), ALLOCATABLE :: Energia,EPol,DifCargas, &
        DipoloExc

```

```

DOUBLE PRECISION, DIMENSION(:,,:), ALLOCATABLE :: Dipolo,CargasDisolvente,&
  Puntos,ConfigDisolv,CargasCapa,IntQMMM,IntMM
DOUBLE PRECISION, DIMENSION(:,,:,:), ALLOCATABLE :: CampoMedio
DOUBLE PRECISION :: EnergiaQ0,EnergiaExc
DOUBLE PRECISION, PARAMETER :: DebyeAtomica=0.393430063965D0, &
  AngstromAtomica=1.8897259582D0,KcalmolAtomica=1.5936012687D-3, &
  KBoltzmann=3.16682968D-6
INTEGER, DIMENSION(:), ALLOCATABLE :: ConfigNum
END MODULE Resultados

```

```

!-----
! Programa principal
!-----
! 33:      Fichero con las cargas fijas y polarizabilidad del disolvente
! 40:      Fichero con los datos para continuar un calculo
! Error:   Variable para controlar los errores
! Fin:     Variable para controlar cuando debe finalizar el ciclo
! i,j:     Contadores
!-----

```

```

PROGRAM ASEP_MD
  USE Parametros
  USE Sistema, ONLY: Carga,Multiplicidad,Soluto,Disolvente,SolutoAnt, &
    PolDisolvente
  USE Resultados
  IMPLICIT NONE
  INTERFACE
    SUBROUTINE EscribePrin(Iter,Cont)
      INTEGER, INTENT(IN) :: Iter
      CHARACTER, INTENT(IN) :: Cont*2
    END SUBROUTINE EscribePrin
    SUBROUTINE GirarMolecula(Tipo,Ordenar)
      CHARACTER, INTENT(IN) :: Tipo*3,Ordenar*2
    END SUBROUTINE GirarMolecula
    SUBROUTINE EjecutarAbInicio(Tipo)
      CHARACTER*4, INTENT(IN) :: Tipo
    END SUBROUTINE EjecutarAbInicio
    SUBROUTINE LeerAbInicio(Tipo)
      CHARACTER*4, INTENT(IN) :: Tipo
    END SUBROUTINE LeerAbInicio
  END INTERFACE
  INTEGER :: Error,Fin,i,j

```

```

!Se lee de la linea de comandos o se le pide al usuario el nombre del
!fichero donde estan definidos los parametros del calculo
!(se le pone la extension ".dat" si no la tiene)

```

```

CALL GETARG(1,EntradaGeneral)
IF (TRIM(EntradaGeneral) == '') THEN
  WRITE(6,101) 'Fichero_de_entrada?'
  READ(5,101) EntradaGeneral
END IF
IF (INDEX(EntradaGeneral,'.') == 0) THEN
  EntradaGeneral = TRIM(EntradaGeneral)//'.dat'
END IF

```

```

!Se leen y se comprueban los valores de los parametros
CALL LeerEntrada

```

```

!Se leen los parametros de la dinamica y el sistema que se estudiara

```

```

CALL LeerSistema
CALL GirarMolecula('Dis','Si')
CALL GirarMolecula('Sol','Si')

!Se preparan los datos iniciales
ALLOCATE (Energia(0:MaxIter+MaxIterPol),Dipolo(0:MaxIter+MaxIterPol,3), &
  IntQMMM(MaxIter+MaxIterPol,2),IntMM(MaxIter+MaxIterPol,4), &
  CampoMedio(MaxIter+MaxIterPol,4,3),EPol(0:MaxIter+MaxIterPol), &
  DifCargas(MaxIter+MaxIterPol),DipoloExc(3))
Dipolo=0.0D0
IntMM=0.0D0
CampoMedio=0.0D0
EPol=0.0D0
IF (NumConfigIni > NumConfigTot) NumConfigIni=1
NumConfig=MIN(NumConfig,NumConfigTot-NumConfigIni+1)

WRITE(21,201) TRIM(DirDinamica),TRIM(ProgAbInicio),AbInicio,Dinamica, &
  TRIM(SalidaGeneral),TRIM(FicheroCont),TRIM(FicheroAbInicio), &
  TRIM(SalidaAbInicio),TRIM(FicheroDinamica),TRIM(SalidaDinamica), &
  TRIM(DinamicaInicial),TRIM(ProgDinamica),ConvEnergia,ConvCargas, &
  MaxDifCargas,TipoCargas,MaxIter,MaxIterPol,Inicio,Acoplamiento, &
  Polarizacion,HacerDinamica,NumConfig,NumConfigIni,FactorRadios, &
  DivMalla,RadioCapa,PrimeraCapa,Proximidad,DipInd,TRIM(CargasFijas), &
  MaxIterOpt,ConvGradiente,MinPaso,MaxPaso,CalcHessiana,BusquedaLineal, &
  EstTransicion

!Se calculan las posiciones donde se colocaran las cargas que representan
!al disolvente
CALL PosicionesCargas

Ciclo=0
Extension='.cic.0'

IF (Inicio == 0) THEN
  IF (Polarizacion == 'NO') THEN
    !Se abre el fichero donde se escribieran los datos para la continuacion
    OPEN(40,STATUS='REPLACE',ACTION='WRITE',FILE=FicheroCont,IOSTAT=Error)
    IF (Error /= 0) THEN
      WRITE(21,666) 'Error_al_crear_el_fichero_de_continuacion.'
      STOP
    END IF

    !Se ejecuta el programa ab-inicio en vacio y se lee su salida
    WRITE(21,669) 'Calculo_en_vacio'
    WRITE(6,669) 'Calculo_en_vacio'
    SolutAnt(:)=Solut(:)
    CALL EjecutarAbInicio('Nada')
    CALL LeerAbInicio('Nada')

    CALL EscribePrin(Ciclo,'Si')

    !Se hace una dinamica previa si asi esta dicho
    IF (Acoplamiento == 'TOT') Disolvente%q=Soluto%q
    IF (TRIM(DinamicaInicial) /= '') THEN
      CALL EjecutarDinamica
    END IF
    EPol(0)=Energia(0)
  ELSE

```



```

      !Se ejecuta el programa ab-inicio en vacio y se lee su salida
      WRITE (21,669) 'Calculo_en_vacio_del_estado_excitado'
      WRITE (6,669) 'Calculo_en_vacio_del_estado_excitado'
      CALL EjecutarAbInicio('Nada')
      CALL LeerAbInicio('Nada')
      EnergiaExc=Energia(0)
      DipoloExc(:)=Dipolo(0,:)

      !Se pone al dia leyendo los resultados de calculos anteriores
      WRITE (21,669) 'Lectura_de_un_calculo_anterior'
      WRITE (6,669) 'Lectura_de_un_calculo_anterior'
      CALL LeerContinuacion !Al final Ciclo=Inicio-1 si Inicio /= 0
      CALL EscribePrin(Ciclo,'No')
    END IF
  ELSE
    !Se pone al dia leyendo los resultados de calculos anteriores
    WRITE (21,669) 'Lectura_de_un_calculo_anterior'
    WRITE (6,669) 'Lectura_de_un_calculo_anterior'
    CALL LeerContinuacion !Al final Ciclo=Inicio-1 si Inicio /= 0
    CALL EscribePrin(Ciclo,'No')
  END IF

  IF (Polarizacion == 'SI') THEN
    IF (Inicio == 0) THEN
      Energia(0)=EnergiaExc
      Dipolo(0,:)=DipoloExc
    END IF
    !Se leen las cargas fijas del disolvente y la polarizabilidad
    OPEN (33,STATUS='OLD',ACTION='READ',FILE=CargasFijas,IOSTAT=Error)
    IF (Error /= 0) THEN
      WRITE (21,666) 'Error_al_abrir_el_fichero_con_las_cargas_fijas'
      STOP
    END IF
    READ (33,*) Disolvente(:) %q
    READ (33,*) PolDisolvente(1,:)
    READ (33,*) PolDisolvente(2,:)
    READ (33,*) PolDisolvente(3,:)
    CLOSE(33)

    !Se preparan algunos datos
    Ciclo=MaxIter
    Energia(MaxIter)=Energia(CicloFin)
  END IF

  !Se efectua el ciclo dinamica + ab-inicio o los ciclos de polarizacion
  !del disolvente
  Fin=0
  DO WHILE (Fin == 0 .AND. Ciclo < MaxIter+MaxIterPol)
    Ciclo=Ciclo+1
    IF (Polarizacion == 'NO') THEN
      WRITE (Extension,301) Ciclo
      Extension='.cic.'//ADJUSTL(Extension)
      WRITE (21,667) Ciclo
      WRITE (6,667) Ciclo
    ELSE
      WRITE (Extension,301) Ciclo-MaxIter
      Extension='.cic.pol.'//ADJUSTL(Extension)
      WRITE (21,668) Ciclo-MaxIter
    END IF
  END DO

```

```

WRITE(6,668) Ciclo-MaxIter
END IF

!Si se optimiza la geometria hay que recentrar el soluto
!Copia el soluto en el disolvente si el acoplamiento es total
IF (MaxIterOpt /= 0) CALL GirarMolecula('Sol','No')
IF (Acoplamiento == 'TOT') THEN
  Disolvente%x=Soluto%x
  Disolvente%y=Soluto%y
  Disolvente%z=Soluto%z
  Disolvente%q=Soluto%q
END IF

!Se modifica la entrada para la dinamica y se ejecuta
IF (Polarizacion == 'NO') CALL EjecutarDinamica

!Se calculan las cargas que representan al disolvente
WRITE(6,101) 'Calculo_de_las_cargas_que_representan_al_disolvente...'
CALL CalcularCargas

!Se hace el calculo ab-initio con las cargas del disolvente
SolutAnt(:)=Solut(:)
IF (MaxIterOpt /= 0) THEN
  CALL OptimizarGeometria
  CALL CalcularCampo
ELSE
  CALL EjecutarAbInitio('Nada')
  CALL LeerAbInitio('Nada')
END IF

!Se calcula la energia de interaccion soluto-disolvente
IntQMMM(Ciclo,1)=SUM(CargasDisolvente(:,4)*CargasDisolvente(:,6))
IntQMMM(Ciclo,2)=SUM(CargasDisolvente(:,5)*CargasDisolvente(:,6))
IF (PrimeraCapa == 'SI') THEN
  IntQMMM(Ciclo,:)=IntQMMM(Ciclo,:)+SUM(CargasCapa(:,4)*CargasCapa(:,5))
END IF
IF (MaxIterOpt == 0) EPol(Ciclo)=Energia(Ciclo)-IntQMMM(Ciclo,1)

!Se escriben los resultados
DifCargas(Ciclo)=MAXVAL(ABS(SolutAnt(:)%q-Solut(:)%q))
WRITE(21,670)
IF (Polarizacion == 'NO') THEN
  CALL EscribirPrin(Ciclo,'Si')
ELSE
  CALL EscribirPrin(Ciclo,'No')
END IF

!Se comprueba la convergencia (no en el primer ciclo)
IF (Ciclo /= 1 .AND. Ciclo /= MaxIter+1) THEN
  IF (ABS(Energia(Ciclo)-Energia(Ciclo-1)) < ConvEnergia .AND. &
    DifCargas(Ciclo) < ConvCargas) Fin=1
END IF
END DO

!Si se quiere hacer un calculo no acoplado se supone que ha convergido
IF (Acoplamiento == 'NIN') THEN
  Fin=1
  WRITE(21,101) 'Fin_del_calculo_no_acoplado.'

```

```

WRITE(6,101) 'Fin_del_calculo_no_acoplado.'
ELSE IF (Fin == 1) THEN
  WRITE(21,101) 'Se_alcanzo_la_convergencia.'
  WRITE(6,101) 'Se_alcanzo_la_convergencia.'
END IF

!Se hace un calculo con las cargas fijas si ha convergido
IF (Polarizacion == 'SI' .AND. Fin == 1) THEN
  Extension=' .cic.Q0'
  CALL EjecutarAbInicio('Nada')
  CALL LeerAbInicio('Nada')
  WRITE(21,306) EnergiaQ0,EnergiaQ0/KcalmolAtomica
  WRITE(6,306) EnergiaQ0,EnergiaQ0/KcalmolAtomica
END IF

!Se escriben los resultados
WRITE(21,669) 'Resultados_finales:'
IF (Polarizacion == 'SI') j=MaxIter+1
DO i=j,Ciclo
  WRITE(21,670)
  CALL EscribePrin(i,'No')
END DO

SELECT CASE (Fin)
CASE (0)
  WRITE(21,101) 'Se_alcanzo_el_llimite_maximo_de_iteraciones.'
  WRITE(6,101) 'Se_alcanzo_el_llimite_maximo_de_iteraciones.'
CASE (1)
  IF (Acoplamiento == 'NIN') THEN
    WRITE(21,101) 'Fin_del_calculo_no_acoplado.'
    WRITE(6,101) 'Fin_del_calculo_no_acoplado.'
  ELSE
    IF (Polarizacion == 'SI') WRITE(21,306) EnergiaQ0, &
      EnergiaQ0/KcalmolAtomica
    WRITE(21,101) 'Se_alcanzo_la_convergencia.'
    WRITE(6,101) 'Se_alcanzo_la_convergencia.'
  END IF
END SELECT
WRITE(21,670)

CLOSE(21)
CLOSE(40)

101 FORMAT (A)
201 FORMAT ( &
' _DirDinamica_=_____ ',A/' _ProgAbInicio_=_____ ',A/ &
' _AbInicio_=_____ ',A/' _Dinamica_=_____ ',A/ &
' _SalidaGeneral_=____ ',A/' _FicheroCont_=_____ ',A/ &
' _FicheroAbInicio_=____ ',A/' _SalidaAbInicio_=____ ',A/ &
' _FicheroDinamica_=____ ',A/' _SalidaDinamica_=____ ',A/ &
' _DinamicaInicial_=____ ',A/' _ProgDinamica_=_____ ',A/ &
' _ConvEnergia_=_____ ',F16.8/' _ConvCargas_=_____ ',F16.8/ &
' _MaxDifCargas_=_____ ',F16.8/' _TipoCargas_=_____ ',A/ &
' _MaxIter_=_____ ',I6/' _MaxIterPol_=_____ ',I6/ &
' _Inicio_=_____ ',I6/' _Acoplamiento_=_____ ',A/ &
' _Polarizacion_=_____ ',A/' _HacerDinamica_=_____ ',A/ &
' _NumConfig_=_____ ',I6/' _NumConfigIni_=_____ ',I6/ &
' _FactorRadios_=_____ ',F16.8/' _DivMalla_=_____ ',I6/ &

```

```

' _RadioCapa_=_',F16.8/' _PrimeraCapa_=_',A/ &
' _Proximidad_=_',F16.8/ &
' _DipInd_=_',A/' _CargasFijas_=_',A/ &
' _MaxIterOpt_=_',I6/' _ConvGradiente_=_',F16.8/ &
' _MinPaso_=_',F16.8/' _MaxPaso_=_',F16.8/ &
' _CalcHessiana_=_',I6/' _BusquedaLineal_=_',A/ &
' _EstTransicion_=_',A/)
301  FORMAT (I4)
306  FORMAT ( &
' _**Energia_en_disolucion_(cargas_perm.)_=_',F16.8,' _Eh_||' /&
' _E(Q0)_=<Y|H(0)+V(0)|Y>_',F16.8,' _Kcal/mol_||' /&
' _',76(' '))
666  FORMAT (80('-')/A/'El programa se detiene'/80('-'))
667  FORMAT (80('-')/'Ciclo_',I3,' :'/80('-'))
668  FORMAT (80('-')/'Ciclo de Polarizacion_',I3,' :'/80('-'))
669  FORMAT (/80('-')/A/80('-'))
670  FORMAT (80('-'))

END PROGRAM ASEP_MD

!-----
! Lee y comprueba los valores dados en la entrada
!-----
! 20,21: Ficheros de entrada y salida generales, respectivamente
! Error: Variable para controlar los errores
!-----

SUBROUTINE LeerEntrada
  USE Parametros
  IMPLICIT NONE
  INTEGER :: Error

  !Asigna valores por defecto a las variables
  DirDinamica='./'
  ProgAbInitio='./gaussian'
  ProgDinamica=' '
  AbInitio='GAUSS'
  Dinamica='MOLDY'
  SalidaGeneral='asepmd.out'
  FicheroCont='asepmd.con'
  FicheroAbInitio='asepmd.g'
  SalidaAbInitio='asepmd.sgau'
  FicheroDinamica='asepmd.ctr'
  SalidaDinamica='asepmd.smol'
  DinamicaInicial=' '
  ConvEnergia=0.01D0
  ConvCargas=0.01D0
  MaxDifCargas=10.0D0
  TipoCargas='MUL'
  MaxIter=10
  MaxIterPol=5
  Inicio=0
  Acoplamiento='PAR'
  Polarizacion='NO'
  HacerDinamica='SI'
  NumConfig=100
  NumConfigIni=1
  FactorRadios=0.7D0
  DivMalla=12

```

```

RadioCapa=7.5
PrimeraCapa='SI'
Proximidad=1.0D0
DipInd='NOR'
CargasFijas='          '
MaxIterOpt=0
ConvGradiente=1.0D-5
MinPaso=1.0D-7
MaxPaso=0.1
CalcHessiana=10
BusquedaLineal='NO'
EstTransicion='NO'

!Lee las variables del fichero de entrada
OPEN(20,STATUS='OLD',ACTION='READ',FILE=EntradaGeneral,IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(6,666) 'Error_al_abrir_el_fichero_de_entrada_general.'
  STOP
END IF
READ(20,NML=Entrada)
CLOSE(20)

!Se abre el fichero donde se escribirá la salida
OPEN(21,STATUS='REPLACE',ACTION='WRITE',FILE=SalidaGeneral,IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(6,666) 'Error_al_crear_el_fichero_de_salida_general.'
  STOP
END IF

WRITE(6,668) '          ASEP/MD'
WRITE(21,668) '          ASEP/MD'
WRITE(21,200)

!Comprueba la validez de los valores y hace algunos cambios
IF (AbInicio /= 'GAUSS' .AND. AbInicio /= 'HONDO') THEN
  WRITE(21,667) 'Programa_de_calculo_ab-inicio_no_reconocido.', &
    "  AbInicio='GAUSS'/'HONDO' "
  STOP
END IF
IF (Dinamica /= 'MOLDY') THEN
  WRITE(21,667) 'Programa_de_dinamica_molecular_no_reconocido.', &
    "  Dinamica='MOLDY' "
  STOP
END IF
IF (ConvEnergia < 0.0D0) THEN
  WRITE(21,666) 'El_valor_de_ConvEnergia_no_puede_ser_negativo.'
  STOP
END IF
IF (ConvCargas < 0.0D0) THEN
  WRITE(21,666) 'El_valor_de_ConvCargas_no_puede_ser_negativo.'
  STOP
END IF
IF (MaxDifCargas < 0.0D0) THEN
  WRITE(21,666) 'El_valor_de_MaxDifCargas_no_puede_ser_negativo.'
  STOP
END IF
IF (TipoCargas /= 'MUL' .AND. TipoCargas /= 'CHE' .AND. &
  TipoCargas /= 'CHG' .AND. TipoCargas /= 'POT') THEN

```

```

WRITE(21,667) 'Calculo_de_las_cargas_del_soluto_no_reconocido.', &
              "___TipoCargas_=_MUL'/'CHE'/'CHG'/'POT' "
STOP
END IF
IF (MaxIter < 0) THEN
WRITE(21,666) 'El_valor_de_MaxIter_no_puede_ser_negativo.'
STOP
END IF
IF (MaxIterPol < 0) THEN
WRITE(21,666) 'El_valor_de_MaxIterPol_no_puede_ser_negativo.'
STOP
END IF
IF (Inicio < 0 .OR. Inicio > MaxIter) THEN
WRITE(21,666) 'El_valor_de_Inicio_no_puede_ser_negativo_ni_mayor_que_
              MaxIter.'
STOP
END IF
IF (Acoplamiento /= 'NIN' .AND. Acoplamiento /= 'PAR' .AND. &
    Acoplamiento /= 'TOT') THEN
WRITE(21,667) 'Tipo_de_acoplamiento_no_reconocido.', &
              "___Acoplamiento_=_NIN'/'PAR'/'TOT' "
STOP
ELSE IF (Acoplamiento == 'NIN' .AND. MaxIter /= 1) THEN
MaxIter=1
WRITE(21,668) 'Calculo_no_acoplado,_MaxIter_sera_igual_a_1.'
WRITE(6,668) 'Calculo_no_acoplado,_MaxIter_sera_igual_a_1.'
ELSE IF (Acoplamiento == 'TOT' .AND. Polarizacion == 'SI') THEN
WRITE(21,666) 'No_es_compatible_el_acoplamiento_total_con_el_disolvente_
              polarizable.'
STOP
END IF
IF (Polarizacion /= 'SI' .AND. Polarizacion /= 'NO') THEN
WRITE(21,667) 'Valor_de_Polarizacion_no_valido.', &
              "___Polarizacion_=_SI'/'NO' "
STOP
ELSE IF (Polarizacion == 'SI' .AND. MaxIterOpt > 0) THEN
MaxIterOpt=0
WRITE(21,668) 'Disolvente_polarizable,_MaxIterOpt_sera_igual_a_0.'
WRITE(6,668) 'Disolvente_polarizable,_MaxIterOpt_sera_igual_a_0.'
ELSE IF (Polarizacion == 'NO') THEN
MaxIterPol=0
END IF
IF (HacerDinamica /= 'SI' .AND. HacerDinamica /= 'NO') THEN
WRITE(21,667) 'Valor_de_HacerDinamica_no_valido.', &
              "___HacerDinamica_=_SI'/'NO' "
STOP
END IF
IF (NumConfig < 1) THEN
WRITE(21,666) 'El_valor_de_NumConfig_debe_ser_mayor_que_0.'
STOP
END IF
IF (NumConfigIni < 1) THEN
WRITE(21,666) 'El_valor_de_NumConfigIni_debe_ser_mayor_que_0.'
STOP
END IF
IF (FactorRadios <= 0.0D0) THEN
WRITE(21,666) 'El_valor_de_FactorRadios_debe_ser_mayor_que_0.0.'
STOP

```

```

END IF
IF (DivMalla < 2) THEN
  WRITE(21,666) 'El_valor_de_DivMalla_debe_ser_mayor_que_1.'
  STOP
END IF
IF (RadioCapa <= 0.0D0) THEN
  WRITE(21,666) 'El_valor_de_RadioCapa_debe_ser_mayor_que_0.0.'
  STOP
END IF
IF (PrimeraCapa /= 'SI' .AND. PrimeraCapa /= 'NO') THEN
  WRITE(21,667) 'Valor_de_PrimeraCapa_no_valido.', &
    "___PrimeraCapa_='SI'/'NO' "
  STOP
END IF
IF (Proximidad <= 0.0D0) THEN
  WRITE(21,666) 'El_valor_de_Proximidad_debe_ser_mayor_que_0.0.'
  STOP
END IF
IF (DipInd /= 'NOR' .AND. DipInd /= 'CON') THEN
  WRITE(21,667) 'Calculo_de_los_dipolos_inducidos_no_reconocido.', &
    "___DipInd_='NOR'/'CON' "
  STOP
END IF
IF (MaxIterOpt < 0) THEN
  WRITE(21,666) 'El_valor_de_MaxIterOpt_no_puede_ser_negativo.'
  STOP
END IF
IF (ConvGradiente <= 0.0D0) THEN
  WRITE(21,666) 'El_valor_de_ConvGradiente_debe_ser_mayor_que_0.0.'
  STOP
END IF
IF (MinPaso < 0.0D0) THEN
  WRITE(21,666) 'El_valor_de_MinPaso_no_puede_ser_negativo.'
  STOP
END IF
IF (MaxPaso <= 0.0D0) THEN
  WRITE(21,666) 'El_valor_de_MaxPaso_debe_ser_mayor_que_0.0.'
  STOP
END IF
IF (CalcHessiana < 0) THEN
  WRITE(21,666) 'El_valor_de_CalcHessiana_debe_ser_mayor_o_igual_que_0.'
  STOP
END IF
IF (BusquedaLineal /= 'SI' .AND. BusquedaLineal /= 'NO') THEN
  WRITE(21,667) 'Valor_de_BusquedaLineal_no_valido.', &
    "___BusquedaLineal_='SI'/'NO' "
  STOP
END IF
IF (EstTransicion /= 'SI' .AND. EstTransicion /= 'NO') THEN
  WRITE(21,667) 'Valor_de_EstTransicion_no_valido.', &
    "___EstTransicion_='SI'/'NO' "
  STOP
END IF

```

```

200 FORMAT (/ &
  '___ASEP-MD_Copyright_(C)_2003_I.F.G.,_M.E.M.,_M.L.S.,_M.A.A.' / &
  '___ASEP-MD_no_tiene_NINGUNA_GARANTIA:' / &
  '___Este_programa_es_software_libre._Puede' / &

```

```

' _ _ _ _ _ redistribuirlo_bajo_ciertas_condiciones.' / &
' _ _ _ _ _ Para_mas_detalles_vea_el_fichero_COPYING.' / &
' _ _ _ _ _ incluido_con_las_fuentes.' / )
666 FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))
667 FORMAT (80('-')/A/'Posibilidades:'/A/'El_programa_se_detiene'/80('-'))
668 FORMAT (/80('-')/A/80('-'))

END SUBROUTINE LeerEntrada

!-----
!Con esta subrutina se leen los resultados de una ejecucion anterior
!-----
!40: Fichero con los datos para continuar un calculo
!Error: Variable para controlar los errores
!i: Contador
!-----
SUBROUTINE LeerContinuacion
  USE Parametros
  USE Sistema, ONLY: Carga,Multiplicidad,Soluto,Disolvente,SolutoAnt
  USE Resultados, ONLY: Energia,EPol,Dipolo,IntQMMM,IntMM,DifCargas, &
    CampoMedio
  IMPLICIT NONE
  INTEGER :: Error,i

  !Se abre el fichero de continuacion y se van leyendo los datos necesarios
  OPEN(40,STATUS='OLD',ACTION='READWRITE',FILE=FicheroCont,IOSTAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_abrir_el_fichero_de_continuacion.'
    STOP
  END IF
  READ(40,*) Carga,Multiplicidad
  READ(40,*) Soluto(:) %NumAt
  READ(40,*) Energia(0),Dipolo(0,:)
  READ(40,*) (Soluto(i) %x,Soluto(i) %y,Soluto(i) %z,i=1,SIZE(Soluto))
  READ(40,*) Soluto(:) %q
  READ(40,*) (Disolvente(i) %x,Disolvente(i) %y,Disolvente(i) %z, &
    i=1,SIZE(Disolvente))
  READ(40,*) Disolvente(:) %q
  EPol(0)=Energia(0)
  SolutoAnt(:)=Soluto(:)
  READ(40,*,IOSTAT=Error) Ciclo
  DO WHILE(Error == 0 .AND. (Ciclo /= Inicio .OR. Inicio == 0))
    READ(40,*) Energia(Ciclo),Dipolo(Ciclo,:),EPol(Ciclo)
    READ(40,*) IntMM(Ciclo,:)
    READ(40,*) IntQMMM(Ciclo,:)
    READ(40,*) (CampoMedio(Ciclo,i,:),i=1,4)
    READ(40,*) (Soluto(i) %x,Soluto(i) %y,Soluto(i) %z,i=1,SIZE(Soluto))
    READ(40,*) (Soluto(i) %q,i=1,SIZE(Soluto))
    READ(40,*) (Disolvente(i) %x,Disolvente(i) %y,Disolvente(i) %z, &
      i=1,SIZE(Disolvente))
    READ(40,*) (Disolvente(i) %q,i=1,SIZE(Disolvente))
    READ(40,*) (SolutoAnt(i) %x,SolutoAnt(i) %y,SolutoAnt(i) %z, &
      i=1,SIZE(SolutoAnt))
    READ(40,*) (SolutoAnt(i) %q,i=1,SIZE(SolutoAnt))
    DifCargas(Ciclo)=MAXVAL(ABS(SolutoAnt(:) %q-Soluto(:) %q))
    READ(40,*,IOSTAT=Error) Ciclo
  END DO

```



```

        IF (Error == 0) THEN
            Ciclo=Ciclo-1
        ELSE IF (Ciclo < Inicio-1) THEN
            WRITE(21,666) 'No_se_puede_iniciar_el_calculo_en_el_ciclo_indicado.'
            STOP
        END IF
        BACKSPACE(40)
        IF (Polarizacion == 'SI') CicloFin=Ciclo

666   FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE LeerContinuacion

!-----
! Escribe los resultados en los ficheros de salida o continuacion
!-----
! Iter: Numero de ciclo cuyos datos se escriben
! Cont: 'Si' o 'No', segun se escriba o no en el fichero de continuacion
! i: Contador
!-----
SUBROUTINE EscribePrin(Iter,Cont)
    USE Parametros, ONLY: MaxIter
    USE Sistema, ONLY: Soluto,Disolvente,SolutoAnt,Carga,Multiplicidad
    USE Resultados, ONLY: Energia,EPol,Dipolo,IntMM,IntQMM,DifCargas, &
        CampoMedio,DebyeAtomica,KcalmolAtomica
    IMPLICIT NONE
    INTEGER, INTENT(IN) :: Iter
    CHARACTER, INTENT(IN) :: Cont*2
    INTEGER :: i

    IF (Cont == 'Si') OPEN(41,STATUS='REPLACE',ACTION='WRITE',FILE='fort.41')

    IF (Iter == 0) THEN
        WRITE(21,302) 'Vacio'
        WRITE(21,303) Energia(Iter),SQRT(SUM(Dipolo(Iter,)**2))/DebyeAtomica
        WRITE(6,302) 'Vacio'
        WRITE(6,303) Energia(Iter),SQRT(SUM(Dipolo(Iter,)**2))/DebyeAtomica
        IF (Cont == 'Si') THEN
            WRITE(40,*) Carga, Multiplicidad
            WRITE(40,*) Soluto(:) %NumAt
            WRITE(40,*) Energia(Iter),Dipolo(Iter,:)
            WRITE(40,*) (Soluto(i) %x,Soluto(i) %y,Soluto(i) %z,i=1,SIZE(Soluto))
            WRITE(40,*) Soluto(:) %q
            WRITE(40,*) (Disolvente(i) %x,Disolvente(i) %y,Disolvente(i) %z, &
                i=1,SIZE(Disolvente))
            WRITE(40,*) Disolvente(:) %q
            WRITE(41,*) Carga, Multiplicidad
            WRITE(41,*) Soluto(:) %NumAt
            WRITE(41,*) Energia(Iter),Dipolo(Iter,:)
            WRITE(41,*) (Soluto(i) %x,Soluto(i) %y,Soluto(i) %z,i=1,SIZE(Soluto))
            WRITE(41,*) Soluto(:) %q
            WRITE(41,*) (Disolvente(i) %x,Disolvente(i) %y,Disolvente(i) %z, &
                i=1,SIZE(Disolvente))
            WRITE(41,*) Disolvente(:) %q
        END IF
    ELSE
        IF (Iter <= MaxIter) THEN
            WRITE(21,302) 'Ciclo',Iter

```

```

ELSE
  WRITE(21,302) 'Ciclo_de_polarizacion',Iter-MaxIter
END IF
WRITE(21,304) Energia(Iter),SQRT(SUM(Dipolo(Iter,):**2))/DebyeAtomica, &
  Energia(Iter)-Energia(Iter-1),DifCargas(Iter)
WRITE(21,305) IntMM(Iter,1),IntMM(Iter,1)/KcalmolAtomica, &
  IntMM(Iter,2),IntMM(Iter,2)/KcalmolAtomica, &
  IntMM(Iter,3),IntMM(Iter,3)/KcalmolAtomica, &
  CampoMedio(Iter,2:)-CampoMedio(Iter,1,1)*CampoMedio(Iter,1,:), &
  CampoMedio(Iter,3:)-CampoMedio(Iter,1,2)*CampoMedio(Iter,1,:), &
  CampoMedio(Iter,4:)-CampoMedio(Iter,1,3)*CampoMedio(Iter,1,:), &
  SUM(CampoMedio(Iter,1:)*Dipolo(Iter,:)), &
  SUM(CampoMedio(Iter,1:)*Dipolo(Iter:))/KcalmolAtomica, &
  Energia(0),Energia(0)/KcalmolAtomica, &
  Energia(Iter),Energia(Iter)/KcalmolAtomica, &
  Energia(Iter)-Energia(0),(Energia(Iter)-Energia(0))/KcalmolAtomica, &
  IntQMMM(Iter,1),IntQMMM(Iter,1)/KcalmolAtomica, &
  IntQMMM(Iter,2),IntQMMM(Iter,2)/KcalmolAtomica, &
  IntQMMM(Iter,1)-IntQMMM(Iter,2), &
  (IntQMMM(Iter,1)-IntQMMM(Iter,2))/KcalmolAtomica, &
  Energia(Iter)-IntQMMM(Iter,1), &
  (Energia(Iter)-IntQMMM(Iter,1))/KcalmolAtomica, &
  Energia(Iter)-Energia(0)-IntQMMM(Iter,1), &
  (Energia(Iter)-Energia(0)-IntQMMM(Iter,1))/KcalmolAtomica, &
  0.5D0*(IntQMMM(Iter,1)+IntQMMM(Iter,2)+IntMM(Iter,2)), &
  0.5D0*(IntQMMM(Iter,1)+IntQMMM(Iter,2)+IntMM(Iter,2))/KcalmolAtomica,&
  EPol(Iter)-EPol(Iter-1)+IntMM(Iter,4), &
  (EPol(Iter)-EPol(Iter-1)+IntMM(Iter,4))/KcalmolAtomica
IF (Iter <= MaxIter) THEN
  WRITE(6,302) 'Ciclo',Iter
ELSE
  WRITE(6,302) 'Ciclo_de_polarizacion',Iter-MaxIter
END IF
WRITE(6,304) Energia(Iter),SQRT(SUM(Dipolo(Iter,):**2))/DebyeAtomica, &
  Energia(Iter)-Energia(Iter-1),DifCargas(Iter)
IF (Cont == 'Si') THEN
  WRITE(40,*) Iter
  WRITE(40,*) Energia(Iter),Dipolo(Iter,:),EPol(Iter)
  WRITE(40,*) IntMM(Iter,:)
  WRITE(40,*) IntQMMM(Iter,:)
  WRITE(40,*) (CampoMedio(Iter,i,:),i=1,4)
  WRITE(40,*) (Solut(i) %x,Solut(i) %y,Solut(i) %z,i=1,SIZE(Solut))
  WRITE(40,*) Solut(:) %q
  WRITE(40,*) (Disolvente(i) %x,Disolvente(i) %y,Disolvente(i) %z, &
    i=1,SIZE(Disolvente))
  WRITE(40,*) Disolvente(:) %q
  WRITE(40,*) (SolutAnt(i) %x,SolutAnt(i) %y,SolutAnt(i) %z, &
    i=1,SIZE(SolutAnt))
  WRITE(40,*) SolutAnt(:) %q
  WRITE(41,*) Iter
  WRITE(41,*) Energia(Iter),Dipolo(Iter,:),EPol(Iter)
  WRITE(41,*) IntMM(Iter,:)
  WRITE(41,*) IntQMMM(Iter,:)
  WRITE(41,*) (CampoMedio(Iter,i,:),i=1,4)
  WRITE(41,*) (Solut(i) %x,Solut(i) %y,Solut(i) %z,i=1,SIZE(Solut))
  WRITE(41,*) Solut(:) %q
  WRITE(41,*) (Disolvente(i) %x,Disolvente(i) %y,Disolvente(i) %z, &
    i=1,SIZE(Disolvente))

```

```

WRITE(41,*) Disolvente(:) %q
WRITE(41,*) (SolutoAnt(i) %x, SolutoAnt(i) %y, SolutoAnt(i) %z, &
i=1, SIZE(SolutoAnt))
WRITE(41,*) SolutoAnt(:) %q
END IF
END IF

IF (Cont == 'Si') THEN
CLOSE(41)
CALL SYSTEM('cat_fort.41_>>_fort.42')
END IF

302 FORMAT (A, 1X, I2)
303 FORMAT ( &
'          Energia=', F16.8, ' Eh      Momento_dipolar=', F8.4, ' D' /&
304 FORMAT ( &
'          Energia=', F16.8, ' Eh      Momento_dipolar=', F8.4, ' D' /&
'          Dif.energia=', F16.8, ' Eh      Max. dif. cargas=', F12.8, ' e-' /&
305 FORMAT ( &
'          ,76(' ') /&
'          ,16(' '), ' ENERGIAS_MM (Promedio_de_N_configuraciones) ', 17(' '), ' |' /&
'          **Interaccion_cargas(sol)-cargas(dis) =', F16.8, ' Eh' /&
'          U(q,s) =', SUM[s*V(q)]', F16.8, ' Kcal/mol' /&
'          **Interaccion_dipolos(dis)-cargas(dis) =', F16.8, ' Eh' /&
'          U(p,q) =', -SUM[p*E(q)]', F16.8, ' Kcal/mol' /&
'          **Interaccion_de_Van_der_Waals =', F16.8, ' Eh' /&
'          U(vdw) =', F16.8, ' Kcal/mol' /&
'          **Fluctuacion_del_campo_electrico =', 2(E10.5E1, ' '), E10.5E1, ' ' /&
'          (forma_tensorial, u.a.) =', 2(E10.5E1, ' '), E10.5E1, ' ' /&
'          <E^2> - <E>^2 =', 2(E10.5E1, ' '), E10.5E1, ' ' /&
'          **Interaccion_dipolo-campo_electrico =', F16.8, ' Eh' /&
'          =', F16.8, ' Kcal/mol' /&
'          ,76(' ') /&
'          ,76(' ') /&
'          ,32(' '), ' ENERGIAS_QM', 33(' '), ' |' /&
'          **Energia_en_vacio =', F16.8, ' Eh' /&
'          E(0) =', <Y(0) | H(0) | Y(0)>', F16.8, ' Kcal/mol' /&
'          **Energia_total_en_disolucion =', F16.8, ' Eh' /&
'          E =', <Y | H(0) + V | Y>', F16.8, ' Kcal/mol' /&
'          **Diferencia_de_energia =', F16.8, ' Eh' /&
'          D(E) =', E - E(0)', F16.8, ' Kcal/mol' /&
'          ,76(' ') /&
'          ,76(' ') /&
'          ,31(' '), ' ENERGIAS_QM/MM', 31(' '), ' |' /&
'          **Interaccion_total =', F16.8, ' Eh' /&
'          U(qt,r) =', SUM[qt*V(r)]', F16.8, ' Kcal/mol' /&
'          **Interaccion_con_cargas_permanentes =', F16.8, ' Eh' /&
'          U(q,r) =', SUM[q*V(r)]', F16.8, ' Kcal/mol' /&
'          **Interaccion_con_cargas_inducidas =', F16.8, ' Eh' /&
'          U(p,r) =', SUM[(qt-q)*V(r)]', F16.8, ' Kcal/mol' /&
'          **Energia_interna_en_disolucion =', F16.8, ' Eh' /&
'          E(int) =', <Y | H(0) | Y> =', E - U(qt,r)', F16.8, ' Kcal/mol' /&
'          **Energia_de_polarizacion_del_soluto =', F16.8, ' Eh' /&
'          E(pol) sol =', E - E(0) - U(qt,r)', F16.8, ' Kcal/mol' /&
'          **Energia_de_interaccion_electrostatica =', F16.8, ' Eh' /&
'          U(elec) =', U(qt,r) + 1/2*U(p,q) - 1/2*U(p,r)', F16.8, ' Kcal/mol' /&
'          **Variacion_de_energia_libre =', F16.8, ' Eh' /&
'          D(G), D(A) =', F16.8, ' Kcal/mol' /&

```

```
      '...',76(' '))
```

```
END SUBROUTINE EscribePrin
```

## F.2. Ejecutar.f90

```

!CONTIENE:
!LeerSistema
!LeerAbInicio(Tipo)
!EjecutarAbInicio(Tipo)
!EjecutarDinamica
!LeerConfig(Config)

!-----
! Lee la definicion del sistema (disolvente, soluto, potencial...) del fichero
! adecuado
!-----
SUBROUTINE LeerSistema
  USE Parametros
  IMPLICIT NONE

  SELECT CASE(Dinamica)
  CASE ('MOLDY')
    CALL LeerControlMoldy
    CALL LeerSistemaMoldy
  END SELECT

END SUBROUTINE LeerSistema

!-----
! Lee la salida del calculo ab-initio (energia, momento dipolar, cargas...)
!-----
! Tipo: Tipo de calculo que se ha realizado ('Nada','Grad','Hess')
!-----
SUBROUTINE LeerAbInicio(Tipo)
  USE Parametros
  IMPLICIT NONE
  INTERFACE
    SUBROUTINE LeerSalidaGaussian(Tipo)
      CHARACTER, INTENT(IN) :: Tipo*4
    END SUBROUTINE LeerSalidaGaussian
  END INTERFACE
  CHARACTER*4, INTENT(IN) :: Tipo

  SELECT CASE(AbInicio)
  CASE ('GAUSS')
    CALL LeerSalidaGaussian(Tipo)
  CASE ('HONDO')
    CALL LeerSalidaHondo(Tipo)
  END SELECT

END SUBROUTINE LeerAbInicio

!-----
! Ejecuta el programa de calculo ab-initio adecuado
! Modifica previamente la entrada si no se trata de un calculo en vacio
!-----
! Tipo: El tipo de calculo que se realizara ('Nada','Grad','Hess')
! Linea: Cadena de texto que contiene la sentencia ejecutable
!-----
SUBROUTINE EjecutarAbInicio(Tipo)

```

```

USE Parametros
IMPLICIT NONE
INTERFACE
  SUBROUTINE ModificarGaussian(Tipo)
    CHARACTER, INTENT(IN) :: Tipo*4
  END SUBROUTINE ModificarGaussian
  SUBROUTINE ModificarHondo(Tipo)
    CHARACTER, INTENT(IN) :: Tipo*4
  END SUBROUTINE ModificarHondo
END INTERFACE
CHARACTER*4, INTENT(IN) :: Tipo
CHARACTER*200 :: Linea

SELECT CASE (AbInicio)
CASE ('GAUSS')
  CALL ModificarGaussian(Tipo)
  Linea=TRIM(ProgAbInicio)//'_'//TRIM(FicheroAbInicio)//TRIM(Extension)//&
    '_'//TRIM(SalidaAbInicio)//TRIM(Extension)
  WRITE(6,101) TRIM(Linea)
  IF (MaxIterOpt == 0) THEN
    IF (Ciclo == 0) THEN
      WRITE(21,101) 'Calculo_con_Gaussian_en_vacio...'
    ELSE IF (Extension == '.cic.Q0') THEN
      WRITE(21,101) 'Calculo_con_Gaussian_con_las_cargas_permanentes...'
    ELSE
      WRITE(21,101) 'Calculo_con_Gaussian...'
    END IF
  END IF
  CALL SYSTEM(Linea)
  CALL SYSTEM('mv_Test.FChk_fchk'//Extension)
CASE ('HONDO')
  CALL ModificarHondo(Tipo)
  Linea=TRIM(ProgAbInicio)//'_'//TRIM(FicheroAbInicio)//TRIM(Extension)//&
    '_'//TRIM(SalidaAbInicio)//TRIM(Extension)
  WRITE(6,101) TRIM(Linea)
  IF (MaxIterOpt == 0) THEN
    IF (Ciclo == 0) THEN
      WRITE(21,101) 'Calculo_con_Hondo_en_vacio...'
    ELSE IF (Extension == '.cic.Q0') THEN
      WRITE(21,101) 'Calculo_con_Hondo_con_las_cargas_permanentes...'
    ELSE
      WRITE(21,101) 'Calculo_con_Hondo...'
    END IF
  END IF
  CALL SYSTEM(TRIM(Linea))
END SELECT

101  FORMAT (A)

END SUBROUTINE EjecutarAbInicio

!-----
! Ejecuta el programa de dinamica molecular adecuado
! Modifica previamente la entrada
!-----
! Linea: Cadena de texto que contiene la sentencia ejecutable
!-----
SUBROUTINE EjecutarDinamica

```

```

USE Parametros
IMPLICIT NONE
CHARACTER*200 :: Linea

SELECT CASE(Dinamica)
CASE ('MOLDY')
CALL ModificarMoldy
IF (HacerDinamica == 'SI' .OR. (Ciclo > Inicio .AND. (Inicio > 0 .OR. TRIM
(DinamicaInicial) /= '')) THEN
Linea=TRIM(ProgDinamica)//'_'//TRIM(FicheroDinamica)//TRIM(Extension)&
//'_'//TRIM(SalidaDinamica)//TRIM(Extension)
WRITE(6,101) TRIM(Linea)
IF (Ciclo == 0) THEN
WRITE(21,101) 'Calculo_previo_con_Moldy...'
ELSE
WRITE(21,101) 'Calculo_con_Moldy...'
END IF
CALL SYSTEM(Linea)
ELSE
WRITE(21,101) 'No_se_efectua_la_dinamica_molecular_(HacerDinamica=NO)'
WRITE(6,101) 'No_se_efectua_la_dinamica_molecular_(HacerDinamica=NO)'
END IF
END SELECT

101 FORMAT (A)

END SUBROUTINE EjecutarDinamica

!-----
!Extrae la informacion necesaria para calcular la configuracion del sistema
!-----
!Config:          Numero de configuracion que se esta leyendo
!Dump:           Fichero dump que contiene la configuracion (MOLDY)
!NumConfigDump: Numero de configuracion dentro del fichero dump (MOLDY)
!MoleculasPoli:  Numero de moleculas poliatomicas (mas de 2 atomos)
!Ext:           Extension de los ficheros
!Linea:         Cadena de texto que contiene la sentencia ejecutable
!-----
SUBROUTINE LeerConfig(Config)
USE Parametros
USE Sistema, ONLY: Disolvente,Soluto,MoleculasDisolvente,MoleculasSoluto
IMPLICIT NONE
INTEGER, INTENT(IN) :: Config
CHARACTER :: Dump*50,Linea*200,Ext*50
INTEGER :: NumConfigDump,MoleculasPoli

!Si ya se esta haciendo la polarizacion del disolvente, se toman los dumps
!del ultimo ciclo completo
IF (Ciclo > MaxIter) THEN
WRITE(Ext,302) CicloFin
Ext='.cic.'//ADJUSTL(Ext)
ELSE
Ext=Extension
END IF

SELECT CASE(Dinamica)
CASE ('MOLDY')
MoleculasPoli=0

```

```
IF (SIZE(Soluto) > 1) MoléculasPoli=MoléculasPoli+MoléculasSoluto
IF (SIZE(Disolvente) > 1) MoléculasPoli=MoléculasPoli+ &
  MoléculasDisolvente
WRITE(Dump,*) Config/MaxDumps
Dump=TRIM(FicheroDump)//TRIM(Ext)//'- '//ADJUSTL(Dump)
NumConfigDump=MOD(Config,MaxDumps)
WRITE(Linea,301) TRIM(DirDinamica)//'dumpext_-' &
  MoléculasSoluto+MoléculasDisolvente,'_-' &
  NumConfigDump,'_-'//TRIM(Dump)//'_>_999.tmp'
WRITE(6,101) TRIM(Linea)
CALL SYSTEM(Linea)
WRITE(Linea,301) TRIM(DirDinamica)//'dumpext_-' &
  MoléculasSoluto+MoléculasDisolvente,'_-' &
  NumConfigDump,'_-'//TRIM(Dump)//'_>_999.tmp'
WRITE(6,101) TRIM(Linea)
CALL SYSTEM(Linea)
WRITE(Linea,301) TRIM(DirDinamica)//'dumpext_-' &
  MoléculasSoluto+MoléculasDisolvente,'_-' &
  NumConfigDump,'_-'//TRIM(Dump)//'_>_999.tmp'
WRITE(6,101) TRIM(Linea)
CALL SYSTEM(Linea)
END SELECT

101  FORMAT (A)
301  FORMAT (A,I6,A,I6,A,I4,A)
302  FORMAT (I4)

END SUBROUTINE LeerConfig
```



### F.3. SubrutinasMoldy.f90

```

!CONTIENE:
!LeerControlMoldy
!LeerSistemaMoldy
!ModificarMoldy

!-----
! Esta subrutina lee los datos de moldy a partir del fichero de control
!-----
! 22: El fichero de control de moldy
! Linea: La linea que se va leyendo
! Error: Variable para controlar los errores
! Var: La variable que se esta leyendo
! Emp: Configuracion en la que empiezan a almacenarse los dumps
! Int: Intervalo entre los dumps
!-----
SUBROUTINE LeerControlMoldy
  USE Parametros
  IMPLICIT NONE
  CHARACTER :: Linea*200,Var*50
  INTEGER :: Error,Emp,Int

  !Se abre el fichero de control para buscar los datos
  OPEN(22,STATUS='OLD',ACTION='READ',FILE=FicheroDinamica,IOSTAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_abrir_el_fichero_de_control_de_Moldy.'
    STOP
  END IF

  !Va asignando los valores a las variables segun las encuentra
  READ(22,101,IOSTAT=Error) Linea
  DO WHILE (Error == 0)
    IF (INDEX(Linea,'=') > 0) THEN
      Var=ADJUSTL(Linea(1:INDEX(Linea,'=')-1))
      Linea=ADJUSTL(Linea(INDEX(Linea,'=')+1:))
      SELECT CASE (TRIM(Var))
        CASE ('sys-spec-file')
          READ(Linea,101) EntradaMoldy
        CASE ('save-file')
          READ(Linea,101) ContinuaMoldy
        CASE ('nsteps')
          READ(Linea,*) NumConfigTot
        CASE ('dump-file')
          READ(Linea,101) FicheroDump
        CASE ('begin-dump')
          READ(Linea,*) Emp
        CASE ('dump-interval')
          READ(Linea,*) Int
        CASE ('ndumps')
          READ(Linea,*) MaxDumps
        CASE ('temperature')
          READ(Linea,*) Temperatura
      END SELECT
    END IF
    READ(22,101,IOSTAT=Error) Linea
  END DO

```

```

!Comprueba si ha habido algun error
IF (Error > 0) THEN
  WRITE(21,666) 'Error_leyendo_el_fichero_de_control_de_Moldy.'
  STOP
END IF
CLOSE(22)

NumConfigTot=(NumConfigTot-Emp)/Int+1

101  FORMAT (A)
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE LeerControlMoldy

!-----
! Esta subrutina lee el sistema que se va a estudiar a partir de los ficheros
! de entrada de moldy
!-----
! 23:      El fichero de entrada de moldy
! Linea:   La linea que se va leyendo
! Atomos:  Numero de atomos del disolvente, del soluto o total (segun)
! Num:     Numero del atomo cuyos datos se estan leyendo
! Repetido: Numero de atomo que es equivalente al considerado
! Error:   Variable para controlar los errores
! Fin:     Variable para comprobar cuando hay que dejar de leer
! i,j,k:   Contadores
!-----
SUBROUTINE LeerSistemaMoldy
  USE Parametros
  USE Sistema
  USE Resultados, ONLY: AngstromAtomica,KcalmolAtomica
  IMPLICIT NONE
  CHARACTER :: Linea*200
  INTEGER :: Fin,Atomos,Num,Repetido,Error,i,j,k

  !Se abre el fichero de entrada de moldy donde se buscaran los datos del
  !sistema
  OPEN(23,STATUS='OLD',ACTION='READ',FILE=EntradaMoldy,IOSTAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_abrir_el_fichero_de_entrada_de_Moldy.'
    STOP
  END IF

  !Lee el nombre y numero de moleculas del disolvente, luego sus atomos
  Fin=0
  DO WHILE (Error == 0 .AND. Fin == 0)
    READ(23,101,IOSTAT=Error) Linea
    Linea=ADJUSTL(Linea)
    IF (Linea(1:1) /= '#' .AND. TRIM(Linea) /= '') Fin=1
  END DO
  NombreDisolvente=Linea(1:INDEX(Linea,'_')-1)
  Linea=Linea(INDEX(Linea,'_')+1:)
  READ(Linea,*) MoleculasDisolvente
  !Cuenta los atomos del disolvente
  j=0
  Atomos=0
  DO WHILE(Error == 0 .AND. Fin == 1)

```

```

READ(23,101,IOSTAT=Error) Linea
Linea=ADJUSTL(Linea)
SELECT CASE(Linea(1:1))
  CASE ('0','1','2','3','4','5','6','7','8','9')
    Atomos=Atomos+1
  CASE ('#','_')
    j=j+1
  CASE DEFAULT
    Fin=2
END SELECT
END DO
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_leyendo_el_fichero_de_entrada_de_Moldy.'
  STOP
END IF
ALLOCATE(Disolvente(Atomos),STAT=Error)
IF (Error /= 0 .OR. Atomos == 0) THEN
  WRITE(21,666) 'Error_al_crear_la_matriz_del_disolvente_o_disolvente_vacio.'
  STOP
END IF
!Vuelve para atras y lee los atomos del disolvente
DO i=1,Atomos+j+1
  BACKSPACE(23)
END DO
Num=0
DO i=1,Atomos+j
  READ(23,101) Linea
  Linea=ADJUSTL(Linea)
  IF (Linea(1:1) /= '#' .AND. TRIM(Linea) /= '') THEN
    Num=Num+1
    READ(Linea,*) Disolvente(Num) %id
    Repetido=0
    DO k=1,Num-1
      IF (Disolvente(k) %id == Disolvente(Num) %id) Repetido=k
    END DO
    IF (Repetido == 0) THEN
      READ(Linea,*) Disolvente(Num) %id,Disolvente(Num) %x, &
        Disolvente(Num) %y,Disolvente(Num) %z,Disolvente(Num) %Masa, &
        Disolvente(Num) %q,Disolvente(Num) %Nombre
    ELSE
      READ(Linea,*) Disolvente(Num) %id,Disolvente(Num) %x, &
        Disolvente(Num) %y,Disolvente(Num) %z
      Disolvente(Num) %Masa=Disolvente(Repetido) %Masa
      Disolvente(Num) %q=Disolvente(Repetido) %q
      Disolvente(Num) %Nombre=Disolvente(Repetido) %Nombre
    END IF
    Disolvente(Num) %NumAt=0
  END IF
END DO
Disolvente %x=Disolvente %x*AngstromAtomica
Disolvente %y=Disolvente %y*AngstromAtomica
Disolvente %z=Disolvente %z*AngstromAtomica

!Lee el nombre y numero de moleculas del soluto, luego sus atomos
READ(23,101,IOSTAT=Error) Linea
NombreSoluto=Linea(1:INDEX(Linea,'_')-1)
Linea=Linea(INDEX(Linea,'_')+1:)

```

```

READ(Linea,*) MoleculasSoluto
IF (MoleculasSoluto /= 1) THEN
  WRITE(21,666) 'Debe_haber_solo_una_molecula_de_soluto.'
  STOP
END IF
!Cuenta los atomos del soluto
j=0
Atomos=0
DO WHILE(Error == 0 .AND. Fin == 2)
  READ(23,101,IOSTAT=Error) Linea
  Linea=ADJUSTL(Linea)
  SELECT CASE (Linea(1:1))
    CASE ('0','1','2','3','4','5','6','7','8','9')
      Atomos=Atomos+1
    CASE ('#','_')
      j=j+1
    CASE DEFAULT
      Fin=3
  END SELECT
END DO
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_leyendo_el_fichero_de_entrada_de_Moldy.'
  STOP
END IF
ALLOCATE(Soluto(Atomos),SolutoAnt(Atomos),STAT=Error)
IF (Error /= 0 .OR. Atomos == 0) THEN
  WRITE(21,666) 'Error_al_crear_la_matriz_del_soluto_o_solutovacio.'
  STOP
END IF
!Vuelve para atras y lee los atomos del soluto
DO i=1,Atomos+j+1
  BACKSPACE(23)
END DO
Num=0
DO i=1,Atomos+j
  READ(23,101) Linea
  Linea=ADJUSTL(Linea)
  IF (Linea(1:1) /= '#' .AND. TRIM(Linea) /= '') THEN
    Num=Num+1
    READ(Linea,*) Soluto(Num) %id
    Repetido=0
    DO k=1,Num-1
      IF (Soluto(k) %id == Soluto(Num) %id) Repetido=k
    END DO
    IF (Repetido == 0) THEN
      READ(Linea,*) Soluto(Num) %id,Soluto(Num) %x,Soluto(Num) %y, &
        Soluto(Num) %z,Soluto(Num) %Masa,Soluto(Num) %q,Soluto(Num) %Nombre
    ELSE
      READ(Linea,*) Soluto(Num) %id,Soluto(Num) %x,Soluto(Num) %y, &
        Soluto(Num) %z
      Soluto(Num) %Masa=Soluto(Repetido) %Masa
      Soluto(Num) %q=Soluto(Repetido) %q
      Soluto(Num) %Nombre=Soluto(Repetido) %Nombre
    END IF
    Soluto(Num) %NumAt=0
  END IF
END DO
END DO
Soluto %x=Soluto %x*AngstromAtomica

```

```

Solutoy=Solutoy*AngstromAtomica
Solutoz=Solutoz*AngstromAtomica

!Hace algunas comprobaciones
IF (NombreSolutoy == NombreDisolvente) THEN
  WRITE(21,666) 'El_solutoy_el_disolvente_no_deberian_tener_el_mismo_nombre
  .'
  STOP
END IF
IF (Acoplamiento == 'TOT' .AND. SIZE(Disolvente) /= SIZE(Solutoy)) THEN
  WRITE(21,666) 'Solutoy_disolvente_son_diferentes,_no_es_posible_el_
  acoplamiento_total.'
  STOP
END IF
Atomos=SIZE(Disolvente)+SIZE(Solutoy)
IF (MAX(MAXVAL(Solutoy(:,%id)),MAXVAL(Disolvente(:,%id))) > Atomos) THEN
  WRITE(21,666) 'El_numero_de_identificacion_de_los_atomos_no_debe_ser_mayor
  _que_el_numero_total_de_atomos.'
  STOP
END IF

!Lee y dimensiona el potencial interatomico
DO WHILE (Error == 0 .AND. Fin == 3)
  READ(23,101) Linea
  Linea=ADJUSTL(Linea)
  IF (Linea(1:1) /= '#' .AND. TRIM(Linea) /= 'end' .AND. &
    TRIM(Linea) /= '') Fin=4
END DO
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_leyendo_el_fichero_de_entrada_de_Moldy.'
  STOP
END IF
READ(Linea,*) TipoPotencial
SELECT CASE (TRIM(TipoPotencial))
  CASE ('lennard-jones')
    k=3
  CASE ('generic')
    k=7
  CASE DEFAULT
    WRITE(21,666) 'Tipo_de_potencial_no_reconocido,_debe_ser_"lennard-jones"_o
    _"generic".'
    STOP
END SELECT
ALLOCATE(Potencial(Atomos,Atomos,k),STAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_crear_la_matriz_de_los_parametros_del_potencial.'
  STOP
END IF
Potencial=0.0D0
READ(23,101,IOSTAT=Error) Linea
Linea=ADJUSTL(Linea)
DO WHILE (Error == 0 .AND. TRIM(Linea) /= 'end')
  IF (Linea(1:1) /= '#' .AND. TRIM(Linea) /= '') THEN
    READ(Linea,*) i,j,(Potencial(i,j,k),k=2,SIZE(Potencial,3))
    Potencial(i,j,1)=1.0D0
    DO k=1,SIZE(Potencial,3)
      Potencial(j,i,k)=Potencial(i,j,k)
    END DO
  END DO

```

```

        END IF
        READ (23,101,IOSTAT=Error) Linea
        Linea=ADJUSTL(Linea)
    END DO
    IF (Error /= 0) THEN
        WRITE (21,666) 'Error_leyendo_el_fichero_de_entrada_de_Moldy.'
        STOP
    END IF

    !Se pasan los parametros del potencial a unidades atomicas
    SELECT CASE (TRIM(TipoPotencial))
    CASE ('lennard-jones')
        Potencial(:, :,2)=Potencial(:, :,2)*KcalmolAtomica
        Potencial(:, :,3)=Potencial(:, :,3)*AngstromAtomica
    CASE ('generic')
        Potencial(:, :,2)=Potencial(:, :,2)*KcalmolAtomica
        Potencial(:, :,3)=Potencial(:, :,3)/AngstromAtomica
        Potencial(:, :,4)=Potencial(:, :,4)*(KcalmolAtomica*AngstromAtomica**12)
        Potencial(:, :,5)=Potencial(:, :,5)*(KcalmolAtomica*AngstromAtomica**4)
        Potencial(:, :,6)=Potencial(:, :,6)*(KcalmolAtomica*AngstromAtomica**6)
        Potencial(:, :,7)=Potencial(:, :,7)*(KcalmolAtomica*AngstromAtomica**8)
    END SELECT

    CLOSE (23)

101  FORMAT (A)
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE LeerSistemaMoldy

!-----
! Esta subrutina modifica el fichero de control de moldy y la entrada con los
! datos del soluto y del disolvente
!-----
! 22,24: Fichero de control original y modificado, respectivamente
! 23,25: Fichero de entrada original (o 'save') y modificado, respectivamente
! Linea: Es la linea que se va leyendo
! Var: Variable que se va leyendo en el control de Moldy
! Ext: Extension correspondiente al ciclo anterior
! Repetido: Numero de atomo que es equivalente al considerado
! Cont: Variable para saber si se continua una dinamica o no
! Fin: Variable para controla cuando se debe dejar de leer
! Error: Variable para controlar los errores
! i,j: Contadores
!-----
SUBROUTINE ModificarMoldy
    USE Parametros
    USE Sistema
    USE Resultados, ONLY: AngstromAtomica,KcalmolAtomica
    IMPLICIT NONE
    CHARACTER :: Linea*200,Var*50,Ext*50
    INTEGER :: Repetido,Cont,Fin,Error,i,j

    !Se abre el fichero de control original y el modificado
    IF (TRIM(DinamicaInicial) /= '' .AND. Ciclo == 0) THEN
        OPEN (22,STATUS='OLD',ACTION='READ',FILE=DinamicaInicial,IOSTAT=Error)
    ELSE
        OPEN (22,STATUS='OLD',ACTION='READ',FILE=FicheroDinamica,IOSTAT=Error)
    
```

```

END IF
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_abrir_el_fichero_de_control_de_Moldy.'
  STOP
END IF
OPEN(24,STATUS='REPLACE',ACTION='WRITE',FILE=TRIM(FicheroDinamica)// &
  Extension,IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_crear_el_fichero_modificado_para_Moldy.'
  STOP
END IF

!Se sustituye algunos datos del fichero de control
Cont=0
READ(22,101,IOSTAT=Error) Linea
DO WHILE (Error == 0)
  IF (INDEX(Linea,'=') > 0) THEN
    Var=ADJUSTL(Linea(1:INDEX(Linea,'=')-1))
    SELECT CASE (TRIM(Var))
      CASE ('sys-spec-file','save-file')
        WRITE(24,101) TRIM(Linea)//TRIM(Extension)
      CASE ('nsteps')
        WRITE(24,101) TRIM(Linea)
      CASE ('lattice-start')
        Cont=1
        IF ((Ciclo == 1 .AND. TRIM(DinamicaInicial) == '') .OR. Ciclo == 0)
          THEN
            Linea='#'//Linea
            Cont=0
          END IF
        WRITE(24,101) TRIM(Linea)
      CASE ('dump-file')
        WRITE(24,101) TRIM(Linea)//TRIM(Extension)//'- '
      CASE DEFAULT
        WRITE(24,101) TRIM(Linea)
    END SELECT
  END IF
  READ(22,101,IOSTAT=Error) Linea
END DO

CLOSE(22)
CLOSE(24)

!Se abre el fichero de entrada original y el modificado
OPEN(23,STATUS='OLD',ACTION='READ',FILE=EntradaMoldy,IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_abrir_el_fichero_de_entrada_de_Moldy.'
  STOP
END IF
OPEN(25,STATUS='REPLACE',ACTION='WRITE',FILE=TRIM(EntradaMoldy)// &
  Extension,IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_crear_el_fichero_modificado_para_Moldy.'
  STOP
END IF

!Se escriben los comentarios iniciales, los datos del disolvente, los
!datos del soluto y los datos del potencial

```

```
READ(23,101,IOSTAT=Error) Linea
Linea=ADJUSTL(Linea)
DO WHILE (Error == 0 .AND. (Linea(1:1) == '#' .OR. TRIM(Linea) == ''))
  IF (TRIM(Linea) /= '') WRITE(25,101) TRIM(Linea)
  READ(23,101,IOSTAT=Error) Linea
  Linea=ADJUSTL(Linea)
END DO
CLOSE(23)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_leyendo_el_fichero_de_entrada_de_Moldy.'
  STOP
END IF
WRITE(25,303) TRIM(NombreDisolvente)//'_ ',MoleculasDisolvente
DO i=1,SIZE(Disolvente)
  Repetido=0
  DO j=1,i-1
    IF (Disolvente(j) %id == Disolvente(i) %id) Repetido=j
  END DO
  IF (Repetido == 0) THEN
    WRITE(25,301) Disolvente(i) %id,Disolvente(i) %x/AngstromAtomica, &
      Disolvente(i) %y/AngstromAtomica,Disolvente(i) %z/AngstromAtomica, &
      Disolvente(i) %Masa,Disolvente(i) %q,TRIM(Disolvente(i) %Nombre)
  ELSE
    WRITE(25,301) Disolvente(i) %id,Disolvente(i) %x/AngstromAtomica, &
      Disolvente(i) %y/AngstromAtomica,Disolvente(i) %z/AngstromAtomica
  END IF
END DO
WRITE(25,303) TRIM(NombreSolutivo)//'_ ',MoleculasSolutivo
DO i=1,SIZE(Solutivo)
  Repetido=0
  DO j=1,i-1
    IF (Solutivo(j) %id == Solutivo(i) %id) Repetido=j
  END DO
  IF (Repetido == 0) THEN
    WRITE(25,301) Solutivo(i) %id,Solutivo(i) %x/AngstromAtomica, &
      Solutivo(i) %y/AngstromAtomica,Solutivo(i) %z/AngstromAtomica, &
      Solutivo(i) %Masa,Solutivo(i) %q,TRIM(Solutivo(i) %Nombre)
  ELSE
    WRITE(25,301) Solutivo(i) %id,Solutivo(i) %x/AngstromAtomica, &
      Solutivo(i) %y/AngstromAtomica,Solutivo(i) %z/AngstromAtomica
  END IF
END DO
WRITE(25,101) 'end'
WRITE(25,101) TRIM(TipoPotencial)
DO i=1,SIZE(Potencial,1)
  DO j=i,SIZE(Potencial,2)
    IF (Potencial(i,j,1) < 0.1) CYCLE
    SELECT CASE (TRIM(TipoPotencial))
      CASE ('lennard-jones')
        WRITE(25,302) i,j,Potencial(i,j,2)/KcalmolAtomica, &
          Potencial(i,j,3)/AngstromAtomica
      CASE ('generic')
        WRITE(25,302) i,j,Potencial(i,j,2)/KcalmolAtomica, &
          Potencial(i,j,3)*AngstromAtomica, &
          Potencial(i,j,4)/(KcalmolAtomica*AngstromAtomica**12), &
          Potencial(i,j,5)/(KcalmolAtomica*AngstromAtomica**4), &
          Potencial(i,j,6)/(KcalmolAtomica*AngstromAtomica**6), &
          Potencial(i,j,7)/(KcalmolAtomica*AngstromAtomica**8)
```



```

        END SELECT
    END DO
END DO
WRITE(25,101) 'end'

!Al final se copian las posiciones de la simulacion anterior
IF (Cont == 1) THEN
    WRITE(Ext,*) Ciclo-1
    Ext='.cic.'//ADJUSTL(Ext)
    OPEN(23,STATUS='OLD',ACTION='READ',FILE=TRIM(ContinuaMoldy)//Ext, &
        IOSTAT=Error)
    IF (Error /= 0) THEN
        WRITE(21,666) 'Error_al_abrir_el_fichero_"save"_de_Moldy.'
        STOP
    END IF
    Fin=0
    READ(23,101,IOSTAT=Error) Linea
    Linea=ADJUSTL(Linea)
    DO WHILE (Fin < 3 .AND. Error == 0)
        IF (TRIM(Linea) == 'end') Fin=Fin+1
        READ(23,101,IOSTAT=Error) Linea
        Linea=ADJUSTL(Linea)
    END DO
    DO WHILE (Error == 0)
        WRITE(25,101) TRIM(Linea)
        READ(23,101,IOSTAT=Error) Linea
        Linea=ADJUSTL(Linea)
    END DO
    IF (Error > 0) THEN
        WRITE(21,666) 'Error_generando_la_nueva_entrada_de_Moldy.'
        STOP
    END IF
    CLOSE(23)
END IF

CLOSE(25)

101  FORMAT (A)
301  FORMAT (1X,I3,1X,3(F12.7,1X),F4.1,1X,F12.7,1X,A)
302  FORMAT (2(I4),6(F11.7,1X))
303  FORMAT (A,I6)
666  FORMAT (80('-',)/A/'El_programa_se_detiene'/80('-',))

END SUBROUTINE ModificarMoldy

```

## F.4. SubrutinasGaussian.f90

```

!CONTIENE:
!ModificarGaussian(Tipo)
!LeerSalidaGaussian(Tipo)

!-----
! Esta subrutina modifica la entrada de Gaussian para incluir las cargas que
! representan al disolvente y calcular el potencial electrostatico producido
! por el soluto
!-----
! 26,27: Fichero de entrada original y modificado, respectivamente
! 63: Fichero con los puntos donde debe calcularse el potencial
! Tipo: Especifica si se necesitan ademas los gradientes o la hessiana
! Linea: La linea que se va leyendo
! Simb: Simbolo del atomo que se va escribiendo
! Fin: Variable para saber cuando dejar de leer
! Error: Variable para controlar los errores
! i: Contador
!-----
SUBROUTINE ModificarGaussian(Tipo)
  USE Parametros
  USE Sistema, ONLY: Carga,Multiplicidad,Soluto
  USE Resultados, ONLY: CargasDisolvente,CargasCapa,AngstromAtomica
  IMPLICIT NONE
  CHARACTER, INTENT(IN) :: Tipo*4
  CHARACTER :: Linea*200,Simb*2
  INTEGER :: Fin,Error,i

  !Abre el fichero de entrada original y modificado
  OPEN(26,STATUS='OLD',ACTION='READ',FILE=FicheroAbInicio,IOSTAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_abrir_el_fichero_de_entrada_de_Gaussian.'
    STOP
  END IF
  OPEN(27,STATUS='REPLACE',ACTION='WRITE',FILE=TRIM(FicheroAbInicio)// &
    Extension,IOSTAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_crear_la_entrada_modificada_para_Gaussian.'
    STOP
  END IF

  !Lee los datos del fichero de entrada, cuando encuentra la primera linea
  !en blanco introduce una nueva, copia el titulo, escribe el soluto y
  !copia las lineas posteriores
  Fin=0
  DO WHILE (Error == 0 .AND. Fin == 0)
    READ(26,101,IOSTAT=Error) Linea
    IF (TRIM(Linea) == '') THEN
      IF (TipoCargas == 'CHG') THEN
        WRITE(27,101) 'Prop=(Grid,Potential)_Pop=CHelpG_FCheck=All_NoSymm'
      ELSE
        WRITE(27,101) 'Prop=(Grid,Potential)_Pop=CHelp_FCheck=All_NoSymm'
      END IF
      SELECT CASE (Tipo)
        CASE ('Grad')
          WRITE(27,101) 'Force'

```

```

        CASE ('Hess')
            WRITE(27,101) 'Freq'
        END SELECT
        IF (Ciclo /= 0) WRITE(27,101) 'Charge=Bohr'
        Fin=1
    END IF
    WRITE(27,101) TRIM(Linea)
END DO
DO WHILE (Error == 0 .AND. Fin == 1)
    READ(26,101,IOSTAT=Error) Linea
    IF (TRIM(Linea) == '') Fin=2
    WRITE(27,101) TRIM(Linea)
END DO
IF (Error /= 0) THEN
    WRITE(21,666) 'Error_leyendo_el_fichero_de_entrada_de_Gaussian.'
    STOP
END IF
IF (Ciclo == 0) THEN
    READ(26,101) Linea
    WRITE(27,101) TRIM(Linea)
ELSE
    WRITE(27,302) Carga,Multiplicidad
END IF
DO i=1,SIZE(Soluto)
    Simb=Soluto(i) %Nombre(1:2)
    IF (ICHAR(Simb(2:2)) < 65 .OR. ICHAR(Simb(2:2)) > 122) Simb(2:2)='_'
    WRITE(27,304) Simb,Soluto(i) %x/AngstromAtomica, &
        Soluto(i) %y/AngstromAtomica, Soluto(i) %z/AngstromAtomica
END DO
DO WHILE (Error == 0)
    READ(26,101,IOSTAT=Error) Linea
    IF (TRIM(Linea) == '') Fin=Fin+1
    IF (Fin > 2 .AND. Error == 0) WRITE(27,101) TRIM(Linea)
END DO
IF (Error > 0) THEN
    WRITE(21,666) 'Error_leyendo_el_fichero_de_entrada_de_Gaussian.'
    STOP
END IF
CLOSE(26)

!Finalmente escribe las cargas que representan al disolvente y los datos
!necesarios para obtener el potencial generado por el soluto
IF (Ciclo > 0) THEN
    IF (TRIM(Extension) == '.cic.Q0') THEN
        DO i=1,SIZE(CargasDisolvente,1)
            WRITE(27,301) CargasDisolvente(i,1:3),CargasDisolvente(i,5)
        END DO
    ELSE
        DO i=1,SIZE(CargasDisolvente,1)
            WRITE(27,301) CargasDisolvente(i,1:4)
        END DO
    END IF
    IF (PrimeraCapa == 'SI') THEN
        DO i=1,SIZE(CargasCapa,1)
            WRITE(27,301) CargasCapa(i,1:4)
        END DO
    END IF
    WRITE(27,101) ''

```

```

        WRITE(27,303) SIZE(CargasDisolvente,1)+SIZE(CargasCapa,1),3,63,64
    ELSE
        WRITE(27,303) SIZE(CargasDisolvente,1),3,63,64
    END IF
    WRITE(27,101) ''
    CLOSE(27)

    !Crea el fichero "fort.63" con las posiciones donde calcular el potencial
    OPEN(63,STATUS='REPLACE',ACTION='WRITE',FILE='fort.63',IOSTAT=Error)
    IF (Error /= 0) THEN
        WRITE(21,666) 'Error_al_abrir_el_fichero_"fort.63".'
        STOP
    END IF
    DO i=1,SIZE(CargasDisolvente,1)
        WRITE(63,301) CargasDisolvente(i,1:3)/AngstromAtomica
    END DO
    IF (Ciclo > 0 .AND. PrimeraCapa == 'SI') THEN
        DO i=1,SIZE(CargasCapa,1)
            WRITE(63,301) CargasCapa(i,1:3)/AngstromAtomica
        END DO
    END IF
    CLOSE(63)

101  FORMAT (A)
301  FORMAT (4(F19.12,1X))
302  FORMAT (I3,I3)
303  FORMAT (I7,3(I4))
304  FORMAT (A2,3(1X,F19.12))
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE ModificarGaussian

!-----
! Esta subrutina lee la salida del programa Gaussian
!-----
! 28:          Fichero "FChk" o de salida de gaussian
! 64:          Fichero con el potencial calculado en los puntos de 'fort.63'
! Tipo:        El tipo de resultados que ha de leerse
! EnergiaCargas: La energia propia de las cargas que representan al disolvente
! Aux1,Aux2:   Puntos para calcular la energia de las cargas
! Dist:        Distancia entre dos puntos
! Linea:       La linea que se va leyendo
! Repetido:    Atomo equivalente al considerado o numero de ellos
! Error:       Variable para controlar los errores
! i,j:         Contadores
!-----
SUBROUTINE LeerSalidaGaussian(Tipo)
    USE Parametros
    USE Sistema, ONLY: Carga,Multiplicidad,Soluto,SolutoAnt
    USE Resultados, ONLY: Energia,EnergiaQ0,Dipolo,CargasCapa,CargasDisolvente
    USE Optimizacion
    IMPLICIT NONE
    INTERFACE
        SUBROUTINE AjusteLagrange(Tipo)
            CHARACTER, INTENT(IN) :: Tipo*3
        END SUBROUTINE AjusteLagrange
    END INTERFACE
    CHARACTER*4, INTENT(IN) :: Tipo

```

```

DOUBLE PRECISION, DIMENSION(4) :: Aux1,Aux2
DOUBLE PRECISION :: EnergiaCargas,Dist
CHARACTER :: Linea*200
INTEGER :: Repetido,Error,i,j

!Se abre el fichero "Fchk" de gaussian y se comprueba si hay error
OPEN(28,STATUS='OLD',ACTION='READ',FILE='fchk'//Extension,IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_abrir_el_fichero_"Fchk"_de_Gaussian.'
  STOP
END IF

Linea='_'
DO WHILE (Error == 0)
  SELECT CASE (TRIM(Linea(1:40)))
    CASE ('Charge')
      READ(Linea(45:),*) Carga
    CASE ('Multiplicity')
      READ(Linea(45:),*) Multiplicidad
    CASE ('Total_Energy')
      IF (TRIM(Extension) == '.cic.Q0') THEN
        READ(Linea(45:),*) EnergiaQ0
      ELSE
        READ(Linea(45:),*) Energia(Ciclo)
      END IF
    CASE ('Atomic_numbers')
      READ(28,*) (Soluto(i) %NumAt,i=1,SIZE(Soluto))
    CASE ('Cartesian_Gradient')
      IF (Tipo /= 'Nada') READ(28,*) Gradiente(:)
    CASE ('Cartesian_Force_Constants')
      IF (Tipo == 'Hess') THEN
        READ(28,*) ((Hessiana(i,j),j=1,i),i=1,SIZE(Hessiana,1))
        DO i=2,SIZE(Hessiana,1)
          Hessiana(1:i-1,i)=Hessiana(i,1:i-1)
        END DO
      END IF
    CASE ('Dipole_Moment')
      IF (TRIM(Extension) /= '.cic.Q0') READ(28,*) Dipolo(Ciclo,:)
  END SELECT
  READ(28,101,IOSTAT=Error) Linea
END DO
CLOSE(28)

!Se corrige la energia total restandole la de las cargas
IF (Ciclo > 0) THEN
  EnergiaCargas=0.0D0
  DO i=1,SIZE(CargasDisolvente,1)
    Aux1(:)=CargasDisolvente(i,1:4)
    IF (TRIM(Extension) == '.cic.Q0') Aux1(4)=CargasDisolvente(i,5)
    DO j=i+1,SIZE(CargasDisolvente,1)
      Aux2(:)=CargasDisolvente(j,1:4)
      IF (TRIM(Extension) == '.cic.Q0') Aux2(4)=CargasDisolvente(j,5)
      Dist=SQRT(SUM((Aux1(1:3)-Aux2(1:3))**2))
      EnergiaCargas=EnergiaCargas+Aux1(4)*Aux2(4)/Dist
    END DO
  END DO
  IF (PrimeraCapa == 'SI') THEN
    DO j=1,SIZE(CargasCapa,1)
      Aux2(:)=CargasCapa(j,1:4)

```

```

        Dist=SQRT(SUM((Aux1(1:3)-Aux2(1:3))**2))
        EnergiaCargas=EnergiaCargas+Aux1(4)*Aux2(4)/Dist
    END DO
END IF
END DO
IF (PrimeraCapa == 'SI') THEN
    DO i=1,SIZE(CargasCapa,1)
        Aux1(:)=CargasCapa(i,1:4)
        DO j=i+1,SIZE(CargasCapa,1)
            Aux2(:)=CargasCapa(j,1:4)
            Dist=SQRT(SUM((Aux1(1:3)-Aux2(1:3))**2))
            EnergiaCargas=EnergiaCargas+Aux1(4)*Aux2(4)/Dist
        END DO
    END DO
END IF
IF (TRIM(Extension) /= '.cic.Q0') THEN
    Energia(Ciclo)=Energia(Ciclo)-EnergiaCargas
ELSE
    EnergiaQ0=EnergiaQ0-EnergiaCargas
END IF
END IF

!Se lee el potencial generado por el soluto
OPEN(64,STATUS='OLD',ACTION='READ',FILE='fort.64',IOSTAT=Error)
IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_abrir_el_fichero_"fort.64".'
    STOP
END IF
DO i=1,SIZE(CargasDisolvente,1)
    READ(64,302) CargasDisolvente(i,6)
END DO
IF (Ciclo > 0 .AND. PrimeraCapa == 'SI') THEN
    DO i=1,SIZE(CargasCapa,1)
        READ(64,302) CargasCapa(i,5)
    END DO
END IF
CLOSE(64)

!Se leen o se calculan las cargas del soluto segun el tipo
IF (TRIM(Extension) /= '.cic.Q0') THEN
    SELECT CASE(TipoCargas)
        !Si hace falta, se leen las cargas de la salida de gaussian
        CASE ('MUL','CHE','CHG')
            OPEN(28,STATUS='OLD',ACTION='READ',FILE=TRIM(SalidaAbInitio)// &
                Extension,IOSTAT=Error)
            IF (TipoCargas == 'MUL') THEN
                DO WHILE (ADJUSTL(Linea) /= 'Total_atomic_charges:' .AND. &
                    Error == 0)
                    READ(28,101,IOSTAT=Error) Linea
                END DO
            ELSE
                DO WHILE (INDEX(Linea,'Charges_from_ESP_fit') == 0 .AND. Error == 0)
                    READ(28,101,IOSTAT=Error) Linea
                END DO
                READ(28,101,IOSTAT=Error) Linea
            END IF
            IF (Error /= 0) THEN
                WRITE(21,666) 'Error_al_leer_el_fichero_de_salida_de_Gaussian.'
```

```

        STOP
    END IF
    READ (28,101,IOSTAT=Error) Linea
    DO i=1,SIZE(Soluto)
        READ (28,301) Soluto(i) %q
    END DO
    CLOSE (28)
    !Si es necesario se ajustan las cargas del soluto al potencial generado
    CASE ('POT')
        CALL AjusteLagrange('Sol')
    END SELECT
    !Se igualan las cargas de los atomos equivalentes, haciendo la media
    DO i=1,SIZE(Soluto)
        Repetido=0
        DO j=1,i-1
            IF (Soluto(j) %id == Soluto(i) %id) Repetido=j
        END DO
        IF (Repetido == 0) THEN
            Repetido=1
            DO j=i+1,SIZE(Soluto)
                IF (Soluto(j) %id == Soluto(i) %id) THEN
                    Soluto(i) %q=Soluto(i) %q+Soluto(j) %q
                    Repetido=Repetido+1
                END IF
            END DO
            Soluto(i) %q=Soluto(i) %q/Repetido
        ELSE
            Soluto(i) %q=Soluto(Repetido) %q
        END IF
    END DO
    !Se evita que haya mucha variacion en las cargas
    Dist=MaxDifCargas/MAXVAL(ABS(Soluto(:) %q-SolutoAnt(:) %q))
    IF (Dist < 1.0D0) THEN
        Soluto(:) %q=SolutoAnt(:) %q+Dist*(Soluto(:) %q-SolutoAnt(:) %q)
    END IF
END IF

101  FORMAT (A)
301  FORMAT (7X,F12.0)
302  FORMAT (60X,F20.12)
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE LeerSalidaGaussian

```

## F.5. SubrutinasHondo.f90

```

!CONTIENE:
!ModificarHondo(Tipo)
!LeerSalidaHondo(Tipo)

!-----
! Esta subrutina modifica la entrada de Hondo para incluir las cargas que
! representan al disolvente y calcular el potencial electrostatico producido
! por el soluto
!-----
! 26,27: Fichero de entrada original y modificado, respectivamente
! Tipo: Especifica si se necesitan ademas los gradientes o las fuerzas
! Linea: La linea que se va leyendo
! Simb: Simbolo del atomo que se esta leyendo
! Error: Variable para controlar los errores
! Fin: Variable para controlar lo que se ha hecho
! i: Contador
!-----
SUBROUTINE ModificarHondo(Tipo)
  USE Parametros
  USE Sistema, ONLY: Solute,NombreSolute
  USE Resultados, ONLY: CargasDisolvente,CargasCapa,AngstromAtomica
  IMPLICIT NONE
  CHARACTER, INTENT(IN) :: Tipo*4
  CHARACTER :: Linea*200,Simb*2
  INTEGER :: Error,i,Fin

  !Abre el fichero de entrada original y modificado
  OPEN(26,STATUS='OLD',ACTION='READ',FILE=FicheroAbInicio,IOSTAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_abrir_el_fichero_de_entrada_de_Hondo.'
    STOP
  END IF
  OPEN(27,STATUS='REPLACE',ACTION='WRITE',FILE=TRIM(FicheroAbInicio)// &
    Extension,IOSTAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_crear_la_entrada_modificada_para_Hondo.'
    STOP
  END IF

  !Escribe la geometria del soluto
  WRITE(27,101) '_$_GEO'
  WRITE(27,301) '$_$_'//TRIM(NombreSolute)//'_Ciclo_',Ciclo
  DO i=1,SIZE(Solute)
    Simb=Solute(i)%Nombre(1:2)
    IF (ICHAR(Simb(2:2)) < 65 .OR. ICHAR(Simb(2:2)) > 122) Simb(2:2)='__'
    WRITE(27,302) Simb,Solute(i)%x/AngstromAtomica, &
      Solute(i)%y/AngstromAtomica,Solute(i)%z/AngstromAtomica
  END DO
  WRITE(27,101) '_$_END'

  !Lee los datos del fichero de entrada y los copia en el nuevo, añadiendo
  !algunas variables
  Fin=0
  READ(26,101,IOSTAT=Error) Linea
  DO WHILE (Error == 0)

```



```

SELECT CASE (TRIM(ADJUSTL(Linea)))
CASE ('$GEO')
  DO WHILE (ADJUSTL(TRIM(Linea)) /= '$END' .AND. Error == 0)
    READ(26,101,IOSTAT=Error) Linea
  END DO
CASE ('$CNTRL')
  WRITE(27,101) '_$CNTRL'
  READ(26,101,IOSTAT=Error) Linea
  DO WHILE (Error == 0 .AND. MOD(Fin,2) < 1)
    WRITE(27,101) TRIM(Linea)
    READ(26,101,IOSTAT=Error) Linea
    IF (TRIM(ADJUSTL(Linea)) == '$END') Fin=Fin+1
  END DO
  WRITE(27,101) '_$NOSYM=1'
  IF (Tipo == 'Nada') WRITE(27,101) '_$RUNFLG=0'
  IF (Tipo == 'Grad' .OR. Tipo == 'Hess') WRITE(27,101) '_$RUNFLG=2'
  WRITE(27,101) '_$END'
CASE ('$PRP')
  WRITE(27,101) '_$PRP'
  READ(26,101,IOSTAT=Error) Linea
  DO WHILE (Error == 0 .AND. MOD(Fin,4) < 2)
    WRITE(27,101) TRIM(Linea)
    READ(26,101,IOSTAT=Error) Linea
    IF (TRIM(ADJUSTL(Linea)) == '$END') Fin=Fin+2
  END DO
  WRITE(27,101) '_$NOPDC=0'
  WRITE(27,101) '_$NOELP=0'
  WRITE(27,101) '_$END'
CASE DEFAULT
  WRITE(27,101) TRIM(Linea)
END SELECT
READ(26,101,IOSTAT=Error) Linea
END DO

!Si no hay $CNTRL ni $PRP, se annaden
IF (MOD(Fin,2) < 1) THEN
  WRITE(27,101) '_$CNTRL'
  WRITE(27,101) '_$NOSYM=1'
  IF (Tipo == 'Nada') WRITE(27,101) '_$RUNFLG=0'
  IF (Tipo == 'Grad' .OR. Tipo == 'Hess') WRITE(27,101) '_$RUNFLG=2'
  WRITE(27,101) '_$END'
END IF
IF (MOD(Fin,4) < 2) THEN
  WRITE(27,101) '_$PRP'
  WRITE(27,101) '_$NOPDC=0'
  WRITE(27,101) '_$NOELP=0'
  WRITE(27,101) '_$END'
END IF

!Se crea el fichero con las cargas que representan al disolvente
IF (Ciclo > 0) THEN
  WRITE(27,101) '_$EFC2'
  WRITE(27,101) "$_EFCFIL='cargas'//TRIM(Extension)//'"
  WRITE(27,101) '_$END'
  OPEN(29,STATUS='REPLACE',ACTION='WRITE',FILE='cargas'//Extension, &
    FORM='UNFORMATTED',ACCESS='SEQUENTIAL')
  WRITE(29) 'Disolvente_'
  WRITE(29) '
'

```

```

        IF (Extension == '.cic.Q0') THEN
            DO i=1,SIZE(CargasDisolvente,1)
                WRITE(29) 'carga_...',CargasDisolvente(i,5),CargasDisolvente(i,1:3)
            END DO
        ELSE
            DO i=1,SIZE(CargasDisolvente,1)
                WRITE(29) 'carga_...',CargasDisolvente(i,4),CargasDisolvente(i,1:3)
            END DO
        END IF
        IF (PrimeraCapa == 'SI') THEN
            DO i=1,SIZE(CargasCapa,1)
                WRITE(29) 'carga_...',CargasCapa(i,4),CargasCapa(i,1:3)
            END DO
        END IF
        CLOSE(29)
    END IF

    !Se escriben las posiciones donde se calculara el potencial electrostatico
    WRITE(27,101) '$_$MAP'
    WRITE(27,101) '$_$XYZMAP='
    DO i=1,SIZE(CargasDisolvente,1)
        WRITE(27,303) CargasDisolvente(i,1:3)
    END DO
    IF (Ciclo > 0 .AND. PrimeraCapa == 'SI') THEN
        DO i=1,SIZE(CargasCapa,1)
            WRITE(27,303) CargasCapa(i,1:3)
        END DO
    END IF
    WRITE(27,101) '$_$END'

    CLOSE(26)
    CLOSE(27)

101  FORMAT (A)
301  FORMAT (A,I2)
302  FORMAT (3X,A2,3(1X,F19.12))
303  FORMAT (2X,3(F20.12))
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE ModificarHondo

!-----
! Esta subrutina lee la salida del programa Hondo
!-----
! 28:          Fichero "FChk" o de salida de gaussian
! 64:          Fichero con el potencial calculado en los puntos de 'fort.63'
! Tipo:       El tipo de resultados que ha de leerse
! EnergiaCargas: La energia propia de las cargas que representan al disolvente
! Dif:       Distancia entre dos puntos
! Linea:     La linea que se va leyendo
! Repetido:  Atomo equivalente al considerado o numero de ellos
! Fin:      Variable para controlar cuando dejar de leer
! Error:    Variable para controlar los errores
! i,j,k,l:  Contadores
!-----

SUBROUTINE LeerSalidaHondo(Tipo)
    USE Parametros
    USE Sistema, ONLY: Carga,Multiplicidad,Soluto,SolutoAnt

```

```

USE Resultados, ONLY: Energia,EnergiaQ0,Dipolo,CargasCapa,CargasDisolvente
USE Optimizacion
IMPLICIT NONE
INTERFACE
  SUBROUTINE AjusteLagrange(Tipo)
    CHARACTER, INTENT(IN) :: Tipo*3
  END SUBROUTINE AjusteLagrange
END INTERFACE
CHARACTER*4, INTENT(IN) :: Tipo
DOUBLE PRECISION :: Dif
CHARACTER :: Linea*200
INTEGER :: Repetido,Fin,Error,i,j,k,l

!Se abre el fichero de salida de hondo y se comprueba si hay error
OPEN(28,STATUS='OLD',ACTION='READ',FILE=TRIM(SalidaAbInitio)//Extension, &
  IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_abrir_el_fichero_de_salida_de_Hondo.'
  STOP
END IF

!Se leen los datos del fichero de salida de hondo
Fin=0
READ(28,101,IOSTAT=Error) Linea
DO WHILE (Error == 0 .AND. MOD(Fin,4) < 2)
  SELECT CASE (TRIM(ADJUSTL(Linea)))
    CASE ('WAVEFUNCTION_DATA')
      READ(28,301) Carga
      READ(28,301) Multiplicidad
    CASE ('DENSITY_CONVERGED')
      Fin=Fin+4
      DO i=1,3
        READ(28,101) Linea
      END DO
      IF (Extension == '.cic.Q0') THEN
        READ(Linea(65:81),*) EnergiaQ0
      ELSE
        READ(Linea(65:81),*) Energia(Ciclo)
      END IF
    CASE ('-MP2-_ENERGY_CALCULATION_ PROGRAM_WRITTEN_BY_A.MARQUEZ,_C.
      DANIEL_AND_M.DUPIUS')
      DO i=1,16
        READ(28,101) Linea
      END DO
      IF (Extension == '.cic.Q0') THEN
        READ(28,309) EnergiaQ0
      ELSE
        READ(28,309) Energia(Ciclo)
      END IF
    CASE ('ATOMIC_CHARGES_(A.U.)')
      DO i=1,2
        READ(28,101) Linea
      END DO
      DO i=1,SIZE(Soluto)
        READ(28,302) Soluto(i) %NumAt
      END DO
    CASE ('DIPOLE_MOMENT')
      DO i=1,5

```

```

        READ(28,101) Linea
    END DO
    IF (Extension /= 'cic.Q0') THEN
        DO i=1,3
            READ(28,303) Dipolo(Ciclo,i)
        END DO
    END IF
    CASE ('ELECTROSTATIC_POTENTIAL/DIAMAGNETIC_SHIELDING')
        DO i=1,6
            READ(28,101) Linea
        END DO
        DO i=1,SIZE(CargasDisolvente,1)
            READ(28,304) CargasDisolvente(i,6)
        END DO
        IF (Ciclo > 0 .AND. PrimeraCapa == 'SI') THEN
            DO i=1,SIZE(CargasCapa,1)
                READ(28,304) CargasCapa(i,5)
            END DO
        END IF
    CASE ('POTENTIAL-DERIVED_CHARGES')
        IF (Extension /= 'cic.Q0') THEN
            DO i=1,9
                READ(28,101) Linea
            END DO
            SELECT CASE (TipoCargas)
                CASE ('CHE')
                    DO i=1,SIZE(Soluto)
                        READ(28,305) Soluto(i) %q
                    END DO
                CASE ('MUL')
                    DO i=1,SIZE(Soluto)
                        READ(28,306) Soluto(i) %q
                    END DO
            END SELECT
        END IF
    CASE ('GRADIENT_OF_THE_ENERGY')
        IF (Tipo == 'Grad' .OR. Tipo == 'Hess') THEN
            DO i=1,6
                READ(28,101) Linea
            END DO
            DO i=1,SIZE(Soluto),8
                j=MIN(i+7,SIZE(Soluto))
                READ(28,307) (Gradiente(k),k=3*i-2,3*j-2,3)
                READ(28,307) (Gradiente(k),k=3*i-1,3*j-1,3)
                READ(28,307) (Gradiente(k),k=3*i,3*j,3)
            END DO
        END IF
    END SELECT
    READ(28,101,IOSTAT=Error) Linea
    IF (TRIM(ADJUSTL(Linea(15:))) == 'NUCLEAR_COORDINATES') Fin=Fin+1
END DO
IF (Error > 0 .OR. Fin < 4) THEN
    WRITE(21,666) 'Error_leyendo_el_fichero_de_salida_de_Hondo.'
    STOP
END IF
!Se lee la hessiana, que esta al final
IF (Tipo == 'Hess') THEN
    REWIND(28)

```

```

Fin=0
DO WHILE(Error == 0 .AND. Fin == 0)
  READ(28,101,IOSTAT=Error) Linea
  IF (TRIM(ADJUSTL(Linea)) == 'CARTESIAN_FORCE_CONSTANT_MATRIX_(HARTREE/
    BOHR**2)') Fin=1
END DO
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_leyendo_el_fichero_de_salida_de_Hondo.'
  STOP
END IF
READ(28,101) Linea
DO i=1,SIZE(Soluto),4
  DO j=1,11
    READ(28,101) Linea
  END DO
  j=MIN(i+3,SIZE(Soluto))
  DO k=1,SIZE(Soluto)
    READ(28,308) (Hessiana(1,3*k-2),l=3*i-2,3*j)
    READ(28,308) (Hessiana(1,3*k-1),l=3*i-2,3*j)
    READ(28,308) (Hessiana(1,3*k),l=3*i-2,3*j)
  END DO
END DO
!La hessiana debe ser simetrica, se hace la media
DO i=1,SIZE(Hessiana,1)-1
  DO j=i+1,SIZE(Hessiana,1)
    Hessiana(i,j)=(Hessiana(i,j)+Hessiana(j,i))*0.5D0
    Hessiana(j,i)=Hessiana(i,j)
  END DO
END DO
END IF
CLOSE(28)

IF (Extension /= '.cic.Q0') THEN
  !Se ajustan las cargas al potencial si es necesario
  IF (TipoCargas == 'POT') CALL AjusteLagrange('Sol')
  !Se igualan las cargas de los atomos equivalentes, haciendo la media
  DO i=1,SIZE(Soluto)
    Repetido=0
    DO j=1,i-1
      IF (Soluto(j)%id == Soluto(i)%id) Repetido=j
    END DO
    IF (Repetido == 0) THEN
      Repetido=1
      DO j=i+1,SIZE(Soluto)
        IF (Soluto(j)%id == Soluto(i)%id) THEN
          Soluto(i)%q=Soluto(i)%q+Soluto(j)%q
          Repetido=Repetido+1
        END IF
      END DO
      Soluto(i)%q=Soluto(i)%q/Repetido
    ELSE
      Soluto(i)%q=Soluto(Repetido)%q
    END IF
  END DO
  !Se evita que haya mucha variacion en las cargas
  Dif=MaxDifCargas/MAXVAL(ABS(Soluto(:)%q-SolutoAnt(:)%q))
  IF (Dif < 1.0D0) THEN
    Soluto(:)%q=SolutoAnt(:)%q+Dif*(Soluto(:)%q-SolutoAnt(:)%q)
  END IF
END IF

```

```
        END IF
    END IF

101  FORMAT (A)
301  FORMAT (27X,I6)
302  FORMAT (20X,I9)
303  FORMAT (16X,F20.0)
304  FORMAT (38X,F15.0)
305  FORMAT (22X,F10.0)
306  FORMAT (32X,F10.0)
307  FORMAT (9X,8(F15.0))
308  FORMAT (20X,12(F9.0))
309  FORMAT (14X,F17.0)
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE LeerSalidaHondo
```

## F.6. Calculos.f90

```

!CONTIENE:
!PosicionesCargas
!CalcularCargas
!PuntosPotencial
!CalcularPotencial
!ReducirCapa
!CalcularCampo
!AjusteLagrange (Tipo)
!ResolverLU (Coef, Term, Num, Error)
!GirarMolecula (Tipo, Ordenar)
!Diagonalizar (Mat, Vec, Val, Ord, Error)

!-----
! Calcula las posiciones donde se situaran las cargas que representan al
! disolvente
!-----
! Angulos: Coordenadas angulares de las cargas
! Zeta:      Constantes para los angulos zeta
! Fi:       Constantes para los angulos fi
! Radio:    Radios para la primera y la segunda capa
! Pi:       El numero pi
! Pi5:      El numero pi/5
! i, j, k:  Contadores
!-----
SUBROUTINE PosicionesCargas
  USE Parametros
  USE Sistema, ONLY: Disolvente
  USE Resultados, ONLY: CargasDisolvente
  IMPLICIT NONE
  DOUBLE PRECISION, DIMENSION(60,2) :: Angulos
  DOUBLE PRECISION, DIMENSION(4) :: Zeta
  DOUBLE PRECISION, DIMENSION(2) :: Fi, Radio
  DOUBLE PRECISION, PARAMETER :: Pi=3.14159265359D0, Pi5=0.628318530718D0
  INTEGER :: i, j, k

  !Se definen las coordenadas angulares de los vertices de un icosaedro
  !truncado (C60)
  ! Zeta (1)=ASIN(2*SQRT((7-3*SQRT(5))/(115-47*SQRT(5))))
  ! Zeta (2)=ASIN(4*SQRT((7-3*SQRT(5))/(115-47*SQRT(5))))
  ! Zeta (3)=ASIN(2*SQRT((7-SQRT(5))/(55-13*SQRT(5))))
  ! Zeta (4)=ASIN(2*SQRT((163-71*SQRT(5))/(695-303*SQRT(5))))
  ! Fi (1)=ASIN(SQRT((5+2*SQRT(5))/(34+14*SQRT(5))))
  ! Fi (2)=ASIN(SQRT((25-9*SQRT(5))/(38-4*SQRT(5)))/2)
  Zeta(1)=0.350405412839D0
  Zeta(2)=0.756743304925D0
  Zeta(3)=1.02988443003D0
  Zeta(4)=1.39830289777D0
  Fi(1)=0.390712505345D0
  Fi(2)=0.206273234033D0

  !Los 5 primeros puntos tienen Zeta=Zeta(1), los 5 siguientes Zeta=Zeta(2),
  !los 10 siguientes Zeta=Zeta(3) y los diez siguientes Zeta=Zeta(4). Los 30
  !ultimos son simetricos a los 30 primeros. Los valores de Fi dan una
  !vuelta en cada grupo, en pasos de 2*Pi/5; quitando los 10 primeros y los
  !10 ultimos, se suma y se resta Fi(1) o Fi(2), el orden es importante

```

```

!para la division en dos capas.
DO i=1,10
  Angulos(i,1)=Zeta(1+INT((i-1)/5))
  Angulos(i,2)=2*MOD(i-1,5)*Pi5
  Angulos(i+10,1)=Zeta(3)
  Angulos(i+10,2)=2*INT((i-1)/2)*Pi5+(2*MOD(i-1,2)-1)*Fi(1)
  Angulos(i+20,1)=Zeta(4)
  Angulos(i+20,2)=2*INT(i/2)*Pi5+(2*MOD(i,2)-1)*Fi(2)
  Angulos(i+30,1)=Pi-Zeta(4)
  Angulos(i+30,2)=(2*INT((i-1)/2)+1)*Pi5+(2*MOD(i-1,2)-1)*Fi(2)
  Angulos(i+40,1)=Pi-Zeta(3)
  Angulos(i+40,2)=(2*INT((i-1)/2)+1)*Pi5+(2*MOD(i-1,2)-1)*Fi(1)
  Angulos(i+50,1)=Pi-Zeta(2-INT((i-1)/5))
  Angulos(i+50,2)=(2*MOD(i+1,5)+1)*Pi5
END DO

!Se calculan los radios de las dos capas
ALLOCATE(CargasDisolvente(SIZE(Angulos,1),6))
IF (PrimeraCapa == 'SI') THEN
  Radio(2)=SQRT(MAXVAL(Disolvente(:)%x*Disolvente(:)%x+ &
    Disolvente(:)%y*Disolvente(:)%y+ &
    Disolvente(:)%z*Disolvente(:)%z))
  Radio(1)=RadioCapa+Radio(2)+Proximidad
ELSE
  Radio(1)=RadioCapa
END IF
Radio(2)=Radio(1)+MAX(2.5D0,Proximidad)
!Se asignan las coordenadas cartesianas de las cargas en dos capas. La
!primera mitad corresponde a la capa mas cercana, con radio=Radio(1), la
!segunda mitad tiene radio=Radio(2), se toman puntos saltados.
DO i=1,SIZE(CargasDisolvente,1)/2
  j=i+SIZE(CargasDisolvente,1)/2
  k=2*i
  CargasDisolvente(i,1)=Radio(1)*SIN(Angulos(k-1,1))*COS(Angulos(k-1,2))
  CargasDisolvente(i,2)=Radio(1)*SIN(Angulos(k-1,1))*SIN(Angulos(k-1,2))
  CargasDisolvente(i,3)=Radio(1)*COS(Angulos(k-1,1))
  CargasDisolvente(j,1)=Radio(2)*SIN(Angulos(k,1))*COS(Angulos(k,2))
  CargasDisolvente(j,2)=Radio(2)*SIN(Angulos(k,1))*SIN(Angulos(k,2))
  CargasDisolvente(j,3)=Radio(2)*COS(Angulos(k,1))
END DO

END SUBROUTINE PosicionesCargas

!-----
! Esta subrutina calcula las cargas que representan al disolvente
!-----
! 32:      Fichero en el que se almacenan las configuraciones del disolvente
! Intervalo: Intervalo entre cada dos configuraciones que se leen
! Dist:    Distancia entre dos puntos
! Error:   Variable para controlar los errores
! i, j:    Contadores
!-----

SUBROUTINE CalcularCargas
  USE Parametros
  USE Resultados, ONLY: Puntos,CargasDisolvente,CargasCapa,IntMM, &
    CampoMedio,ConfigNum,KBoltzmann
  IMPLICIT NONE
  INTERFACE

```



```

SUBROUTINE LeerConfig(Config)
  INTEGER, INTENT(IN) :: Config
END SUBROUTINE LeerConfig
SUBROUTINE AjusteLagrange(Tipo)
  CHARACTER, INTENT(IN) :: Tipo*3
END SUBROUTINE AjusteLagrange
END INTERFACE
DOUBLE PRECISION :: Dist
INTEGER :: Intervalo,Error,i,j

!Crea los puntos donde se calculara el potencial
CALL PuntosPotencial

!Se abre el fichero donde se escribiran las configuraciones del disolvente
OPEN(32,STATUS='REPLACE',ACTION='READWRITE',FILE='configs',IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_abrir_el_fichero_con_las_configuraciones_del_
    disolvente.'
  STOP
END IF

!Se van tomando las distintas configuraciones
Intervalo=1
IF (NumConfig > 1) Intervalo=(NumConfigTot-NumConfigIni)/(NumConfig-1)
IF (Intervalo < 1) THEN
  WRITE(21,666) 'No_hay_suficientes_configuraciones_en_la_dinamica/dump.'
  STOP
END IF
IF (.NOT. ALLOCATED(ConfigNum)) ALLOCATE(ConfigNum(0:NumConfig))
ConfigNum=0
DO i=1,NumConfig
  !Se lee la configuracion a estudiar
  CALL LeerConfig(NumConfigIni-1+(i-1)*Intervalo)
  !Se obtienen las posiciones de los atomos y se calcula el potencial
  CALL CalcularPotencial
  ConfigNum(i)=ConfigNum(NumConfig)
END DO
Puntos(:,4:5)=Puntos(:,4:5)/NumConfig
IntMM(Ciclo,2)=IntMM(Ciclo,2)
IntMM(Ciclo,:)=IntMM(Ciclo,)/NumConfig
IntMM(Ciclo,4)=KBoltzmann*Temperatura*LOG(IntMM(Ciclo,4))
CampoMedio(Ciclo,,:)=CampoMedio(Ciclo,,:)/NumConfig

IF (PrimeraCapa == 'SI' .OR. MaxIterOpt > 0) THEN
  IF (Ciclo <= MaxIter+1) CALL ReducirCapa
END IF

!Resta el potencial debido a las cargas de la capa de solvatacion
IF (PrimeraCapa == 'SI') THEN
  DO i=1,SIZE(Puntos,1)
    DO j=1,SIZE(CargasCapa,1)
      Dist=SQRT(SUM((CargasCapa(j,1:3)-Puntos(i,1:3))**2))
      Puntos(i,4:5)=Puntos(i,4:5)-CargasCapa(j,4)/Dist
    END DO
  END DO
END IF
CLOSE(32)

```

```

        !Se ajustan las cargas al potencial
        CALL AjusteLagrange('Dis')
        !Tambien se ajustan al potencial de las cargas permanentes si hace falta
        IF (Polarizacion == 'SI') THEN
            CALL AjusteLagrange('QDi')
        ELSE
            CargasDisolvente(:,5)=CargasDisolvente(:,4)
        END IF

        DEALLOCATE (Puntos)

666   FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE CalcularCargas

!-----
! Esta subrutina calcula las coordenadas de los puntos de una malla que quedan
! dentro de una cavidad de esferas proporcionales a los radios de Van der Waals,
! donde se calculara el potencial
!-----
! RadioVdW:  Radios de Van der Waals de los distintos elementos
! Inc:       Separacion entre los puntos de la malla en las tres dimensiones
! Maximo:    Maximias coordenadas x,y,z de la malla
! Minimo:    Minimas coordenadas x,y,z de la malla
! Punto:     Coordenadas del punto que se esta analizando
! Dentro:    Matriz que indica si cada punto esta dentro de la cavidad
! Error:     Variable para controlar los errores
! i,j,k,l:   Contadores
!-----
SUBROUTINE PuntosPotencial
    USE Parametros
    USE Sistema, ONLY: Solute
    USE Resultados, ONLY: Puntos,AngstromAtomica
    IMPLICIT NONE
    DOUBLE PRECISION, DIMENSION(54) :: RadioVdW
    DOUBLE PRECISION, DIMENSION(3)  :: Punto,Inc,Maximo,Minimo
    INTEGER, DIMENSION(0:DivMalla,0:DivMalla,0:DivMalla) :: Dentro
    INTEGER :: Error,i,j,k,l

    RadioVdW=0.0D0
    RadioVdW(1)=1.2D0
    RadioVdW(2)=1.4D0
    RadioVdW(6)=1.7D0
    RadioVdW(7)=1.55D0
    RadioVdW(8)=1.52D0
    RadioVdW(9)=1.47D0
    RadioVdW(10)=1.54D0
    RadioVdW(14)=2.1D0
    RadioVdW(15)=1.8D0
    RadioVdW(16)=1.8D0
    RadioVdW(17)=1.75D0
    RadioVdW(18)=1.88D0
    RadioVdW(33)=1.85D0
    RadioVdW(34)=1.9D0
    RadioVdW(35)=1.85D0
    RadioVdW(36)=2.02D0
    RadioVdW(52)=2.06D0
    RadioVdW(53)=1.98D0

```

```

RadioVdW(54)=2.16D0
RadioVdW=RadioVdW*FactorRadios*AngstromAtomica

!Se definen los extremos de la cavidad ocupada por el soluto
Maximo=0.0D0
Minimo=0.0D0
DO i=1,SIZE(Soluto)
  Maximo(1)=MAX(Maximo(1),Soluto(i) %x+RadioVdW(Soluto(i) %NumAt))
  Maximo(2)=MAX(Maximo(2),Soluto(i) %y+RadioVdW(Soluto(i) %NumAt))
  Maximo(3)=MAX(Maximo(3),Soluto(i) %z+RadioVdW(Soluto(i) %NumAt))
  Minimo(1)=MIN(Minimo(1),Soluto(i) %x-RadioVdW(Soluto(i) %NumAt))
  Minimo(2)=MIN(Minimo(2),Soluto(i) %y-RadioVdW(Soluto(i) %NumAt))
  Minimo(3)=MIN(Minimo(3),Soluto(i) %z-RadioVdW(Soluto(i) %NumAt))
END DO
Inc=(Maximo-Minimo)/DivMalla

!Para cada punto de la malla inicial se comprueba si queda dentro de
!la cavidad y si es asi se marca el punto
Dentro=0
DO k=0,DivMalla
  Punto(3)=Minimo(3)+Inc(3)*k
  DO j=0,DivMalla
    Punto(2)=Minimo(2)+Inc(2)*j
    DO i=0,DivMalla
      Punto(1)=Minimo(1)+Inc(1)*i
      DO l=1,SIZE(Soluto)
        IF ((Punto(1)-Soluto(l) %x)**2+(Punto(2)-Soluto(l) %y)**2+ &
          (Punto(3)-Soluto(l) %z)**2 > &
          RadioVdW(Soluto(l) %NumAt)**2) CYCLE
        Dentro(i,j,k)=1
        EXIT
      END DO
    END DO
  END DO
END DO

!Se calculan los puntos que hay que tener en cuenta
ALLOCATE(Puntos(COUNT(Dentro == 1),5),STAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_crear_la_matriz_con_los_puntos_del_potencial.'
  STOP
END IF
Puntos=0.0D0
l=0
DO i=0,DivMalla
  DO j=0,DivMalla
    DO k=0,DivMalla
      IF (Dentro(i,j,k) == 0) CYCLE
      l=l+1
      Puntos(l,1)=Minimo(1)+Inc(1)*i
      Puntos(l,2)=Minimo(2)+Inc(2)*j
      Puntos(l,3)=Minimo(3)+Inc(3)*k
    END DO
  END DO
END DO
666 FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

```

```

END SUBROUTINE PuntosPotencial

!-----
! Esta subrutina halla las posiciones de los atomos del disolvente y del
! soluto en una configuracion determinada y calcula el potencial que generan
! las cargas permanentes y los dipolos inducidos en los puntos de una malla
!-----
! 999:           Fichero temporal "999.tmp" de donde se lee la configuracion
! CentroMasas:  Posiciones de los centros de masas de todas las moleculas
! Cuat:         Cuaterniones de todas las moleculas (orientacion)
! Dis,Sol:     Geometria y cargas del disolvente y del soluto
! Sol2:        Geometria y cargas del soluto en el ciclo anterior
! CampoQ:      Campo electrico generado por las cargas permanentes en las
!              posiciones de los centros de masas
! Campo:       Campo electrico total en los centros de masas
! Dipolos:     Dipolos inducidos sobre cada molecula de disolvente
! Rot:         Matriz de rotacion de la molecula que se esta considerando
! Pol:         Polarizabilidades de las moleculas de disolvente
! AtDis:       Posiciones de los atomos de disolvente
! Celda:       Dimension de la celdilla de simulacion (se supone cubica)
! Dist:        Variable para almacenar distancias
! Dif:         Diferencias entre dipolos inducidos (para calculo consistente)
! Pot1,Pot2,Pot3,Pot4: Potenciales o energias calculados en cada momento
! S,F:         Variables para la correccion del campo a distancias cortas
! NTipoPot:    Numero que indica el tipo de potencial
! Eps,Sig6:    Parametros del potencial Lennard-Jones
! Dist4,Dist6,Dist8,Dist12: Parametros del potencial generico
! Preci:       Precision para el calculo consistente de dipolos inducidos
! Aux,Aux2:    Variables auxiliares (vector y matriz)
! Mols:        Numero de moleculas a considerar
! Error:       Variable para controlar los errores
! i,j,k,l,ii,kk: Contadores
!-----

SUBROUTINE CalcularPotencial
  USE Parametros
  USE Sistema, ONLY: Solute,Disolvente,SoluteAnt,TipoPotencial,Potencial, &
    MoleculasDisolvente,MoleculasSolute,PolDisolvente
  USE Resultados, ONLY: Puntos,IntMM,CampoMedio,ConfigNum,KBoltzmann, &
    AngstromAtomica
  IMPLICIT NONE
  DOUBLE PRECISION, DIMENSION(:,,:), ALLOCATABLE :: CentroMasas,Cuat,Dis, &
    Sol,Sol2,CampoQ,Campo,Dipolos
  DOUBLE PRECISION, DIMENSION(:,,:), ALLOCATABLE :: Pol,AtDis
  DOUBLE PRECISION, DIMENSION(3) :: Aux
  DOUBLE PRECISION, DIMENSION(3,3) :: Rot,Aux2
  DOUBLE PRECISION :: Celda,Dist,Dif,Pot1,Pot2,Pot3,Pot4,S,F
  DOUBLE PRECISION :: Eps,Sig6,Dist4,Dist6,Dist8,Dist12
  DOUBLE PRECISION, PARAMETER :: Preci=1.0D-3
  INTEGER :: Mols,NTipoPot,Error,i,j,k,l,ii,kk

  !Se definen los datos iniciales
  Mols=MoleculasDisolvente+MoleculasSolute
  ALLOCATE (CentroMasas (Mols,3),Cuat (Mols,4),Dis (SIZE (Disolvente),4), &
    Sol (SIZE (Solute),4),AtDis (MoleculasDisolvente,SIZE (Disolvente),3), &
    Sol2 (SIZE (SoluteAnt),4))
  IF (Polarizacion == 'SI') ALLOCATE (Pol (MoleculasDisolvente,3,3))
  Dis(:,1)=Disolvente(:) %x
  Dis(:,2)=Disolvente(:) %y

```

```

Dis(:,3)=Disolvente(:) %z
Dis(:,4)=Disolvente(:) %q
Sol(:,1)=Soluto(:) %x
Sol(:,2)=Soluto(:) %y
Sol(:,3)=Soluto(:) %z
Sol(:,4)=Soluto(:) %q

!Se lee la configuracion
OPEN(999,STATUS='OLD',ACTION='READ',RECL=10000,FILE='999.tmp', &
  IOSTAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_abrir_el_fichero_"999.tmp".'
  STOP
END IF
READ(999,*) Celda
Celda=Celda*AngstromAtomica
READ(999,*) ((CentroMasas(i,j),j=1,3),i=1,Mols)
CentroMasas=CentroMasas*AngstromAtomica
Cuat(:,1)=1.0D0
Cuat(:,2:4)=0.0D0
IF (SIZE(Soluto) > 1 .AND. SIZE(Disolvente) > 1) &
  READ(999,*) ((Cuat(i,j),j=1,4),i=1,Mols)
IF (SIZE(Soluto) == 1 .AND. SIZE(Disolvente) > 1) &
  READ(999,*) ((Cuat(i,j),j=1,4),i=1,MoleculasDisolvente)
IF (SIZE(Soluto) > 1 .AND. SIZE(Disolvente) == 1) &
  READ(999,*) ((Cuat(i,j),j=1,4),i=MoleculasDisolvente+1,Mols)
CLOSE(999)

!Se comprueba si el radio de corte es adecuado
IF (RadioCapa > Celda/2.0D0) THEN
  RadioCapa = Celda/2.0D0
  WRITE(21,102) 'Se_reduce_la_capa_de_solvatacion,_RadioCapa=',RadioCapa
END IF

!Se calculan las posiciones en la configuracion y se centra el sistema en
!el soluto. "Aux2" es la matriz de rotacion del soluto
Aux2(1,1)=Cuat(Mols,1)**2+Cuat(Mols,2)**2-Cuat(Mols,3)**2-Cuat(Mols,4)**2
Aux2(2,2)=Cuat(Mols,1)**2-Cuat(Mols,2)**2+Cuat(Mols,3)**2-Cuat(Mols,4)**2
Aux2(3,3)=Cuat(Mols,1)**2-Cuat(Mols,2)**2-Cuat(Mols,3)**2+Cuat(Mols,4)**2
Aux2(1,2)=2.0D0*(Cuat(Mols,2)*Cuat(Mols,3)-Cuat(Mols,1)*Cuat(Mols,4))
Aux2(2,1)=2.0D0*(Cuat(Mols,2)*Cuat(Mols,3)+Cuat(Mols,1)*Cuat(Mols,4))
Aux2(2,3)=2.0D0*(Cuat(Mols,3)*Cuat(Mols,4)-Cuat(Mols,1)*Cuat(Mols,2))
Aux2(3,2)=2.0D0*(Cuat(Mols,3)*Cuat(Mols,4)+Cuat(Mols,1)*Cuat(Mols,2))
Aux2(3,1)=2.0D0*(Cuat(Mols,2)*Cuat(Mols,4)-Cuat(Mols,1)*Cuat(Mols,3))
Aux2(1,3)=2.0D0*(Cuat(Mols,2)*Cuat(Mols,4)+Cuat(Mols,1)*Cuat(Mols,3))
k=0
DO i=1,MoleculasDisolvente
  !Se calcula la posicion del centro de masas
  Aux(:)=ANINT((CentroMasas(i,:)-CentroMasas(Mols,:))/Celda)
  CentroMasas(i,:)=CentroMasas(i,:)-CentroMasas(Mols,:)-Celda*Aux(:)
  Dist=SUM(CentroMasas(i,:)*CentroMasas(i,:))
  IF (Dist > Celda*Celda/4.0D0) CYCLE
  k=k+1
  CentroMasas(k,:)=MATMUL(TRANSPPOSE(Aux2(:,:)),CentroMasas(i,:))
  WRITE(32,*) CentroMasas(k,:)
  !Se calcula la matriz de rotacion de la molecula
  Rot(1,1)=Cuat(i,1)**2+Cuat(i,2)**2-Cuat(i,3)**2-Cuat(i,4)**2
  Rot(2,2)=Cuat(i,1)**2-Cuat(i,2)**2+Cuat(i,3)**2-Cuat(i,4)**2

```

```

Rot(3,3)=Cuat(i,1)**2-Cuat(i,2)**2-Cuat(i,3)**2+Cuat(i,4)**2
Rot(1,2)=2.0D0*(Cuat(i,2)*Cuat(i,3)-Cuat(i,1)*Cuat(i,4))
Rot(2,1)=2.0D0*(Cuat(i,2)*Cuat(i,3)+Cuat(i,1)*Cuat(i,4))
Rot(2,3)=2.0D0*(Cuat(i,3)*Cuat(i,4)-Cuat(i,1)*Cuat(i,2))
Rot(3,2)=2.0D0*(Cuat(i,3)*Cuat(i,4)+Cuat(i,1)*Cuat(i,2))
Rot(3,1)=2.0D0*(Cuat(i,2)*Cuat(i,4)-Cuat(i,1)*Cuat(i,3))
Rot(1,3)=2.0D0*(Cuat(i,2)*Cuat(i,4)+Cuat(i,1)*Cuat(i,3))
!Se calculan las posiciones de los atomos de la molecula
DO j=1,SIZE(Dis,1)
  AtDis(k,j,:)=MATMUL(TRANSPOSE(Aux2(:,)),MATMUL(Rot(:,),Dis(j,1:3)))
  AtDis(k,j,:)=AtDis(k,j,:)+CentroMasas(k,:)
  WRITE(32,*) AtDis(k,j,:)
END DO
!Se rota el tensor de polarizabilidad
IF (Polarizacion == 'SI') THEN
  Pol(k,:)=MATMUL(Rot(:,),MATMUL(PolDisolvente,TRANSPOSE(Rot(:,))))
  Pol(k,:)=MATMUL(TRANSPOSE(Aux2(:,)),MATMUL(Pol(k,:),Aux2(:,)))
END IF
END DO
WRITE(32,103)
DEALLOCATE(Cuat)

!Se quitan las moleculas fuera del radio de corte ("Celda"/2)
CentroMasas(k+1:MoleculasDisolvente,:)=0.0D0
AtDis(k+1:MoleculasDisolvente,:)=0.0D0
IF (Polarizacion == 'SI') Pol(k+1:MoleculasDisolvente,:)=0.0D0
Mols=k
ConfigNum(NumConfig)=ConfigNum(NumConfig)+k*SIZE(Disolvente)

!Se calculan los dipolos inducidos
IF (Polarizacion == 'SI') THEN
  ALLOCATE(Dipolos(Mols,3),CampoQ(Mols,3))
  Dipolos=0.0D0
  CampoQ=0.0D0
  DO i=1,Mols
    !Se calcula el campo electrico generado por las cargas permanentes
    DO k=1,SIZE(Dis,1)
      DO j=1,Mols
        IF (j == i) CYCLE
        Aux(:)=CentroMasas(i,:)-AtDis(j,k,:)
        Dist=SQRT(SUM(Aux*Aux))
        CampoQ(i,:)=CampoQ(i,:)+Dis(k,4)*Aux(:)/(Dist*Dist*Dist)
      END DO
    END DO
    DO k=1,SIZE(Sol,1)
      Aux(:)=CentroMasas(i,:)-Sol(k,1:3)
      Dist=SQRT(SUM(Aux*Aux))
      CampoQ(i,:)=CampoQ(i,:)+Sol(k,4)*Aux(:)/(Dist*Dist*Dist)
    END DO
    Dipolos(i,:)=MATMUL(Pol(i,:,:),CampoQ(i,:))
  END DO
  IF (DipInd == 'CON') THEN
    ALLOCATE(Campo(Mols,3))
    S=(PolDisolvente(1,1)+PolDisolvente(2,2)+PolDisolvente(3,3))/3
    S=3.1407D0*S**(1.0D0/3.0D0)
    Dif=Preci+1.0D0
    DO WHILE(Dif >= Preci)
      Dif=0.0D0
    END DO
  END IF
END IF

```

```

Campo=CampoQ
DO i=1,Mols
  !Se annade la contribucion debida a los dipolos inducidos
  DO j=i+1,Mols
    Aux(:)=(CentroMasas(i,:)-CentroMasas(j,:))
    Dist=SQRT(SUM(Aux*Aux))
    F=MIN(1.0D0,Dist/S)
    DO k=1,3
      Aux2(k,:)=3.0D0*F**3*Aux(k)*Aux(:)/(Dist*Dist*Dist*Dist*Dist)
      Aux2(k,k)=Aux2(k,k)-(4.0D0-3.0D0*F)*F*F*F/(Dist*Dist*Dist)
    END DO
    Campo(i,:)=Campo(i,)+MATMUL(Aux2,Dipolos(j,:))
    Campo(j,:)=Campo(j,)+MATMUL(Aux2,Dipolos(i,:))
  END DO
  Aux(:)=Dipolos(i,:)
  Dipolos(i,:)=MATMUL(Pol(i,,:),Campo(i,:))
  Dif=MAX(Dif,MAXVAL(ABS(Dipolos(i,)-Aux(:))))
END DO
END DO
ELSE
  ALLOCATE(Campo(1,3))
END IF
Campo(1,:)=0.0D0
DO i=1,Mols
  !Se elimina la contribucion del soluto y se calcula U(p,q)
  DO j=1,SIZE(Sol,1)
    Aux(:)=CentroMasas(i,:)-Sol(j,1:3)
    Dist=SQRT(SUM(Aux*Aux))
    CampoQ(i,:)=CampoQ(i,)-Sol(j,4)*Aux(:)/(Dist*Dist*Dist)
  END DO
  IntMM(Ciclo,2)=IntMM(Ciclo,2)-SUM(Dipolos(i,)*CampoQ(i,:))
  !Se calcula el campo generado por el disolvente en el origen
  DO j=1,SIZE(Dis,1)
    Aux(:)=AtDis(i,j,:)
    Dist=SQRT(SUM(Aux*Aux))
    Campo(1,:)=Campo(1,)+Dis(j,4)*Aux(:)/(Dist*Dist*Dist)
  END DO
  Aux(:)=CentroMasas(i,:)
  Dist=SQRT(SUM(Aux*Aux))
  DO k=1,3
    Aux2(k,:)=3.0D0*Aux(k)*Aux(:)/(Dist*Dist*Dist*Dist*Dist)
    Aux2(k,k)=Aux2(k,k)-1.0D0/(Dist*Dist*Dist)
  END DO
  Campo(1,:)=Campo(1,)+MATMUL(Aux2,Dipolos(i,:))
END DO
DEALLOCATE(CampoQ)
ELSE
  ALLOCATE(Campo(1,3))
  Campo=0.0D0
  !Se calcula el campo generado por el disolvente en el origen
  DO i=1,Mols
    DO j=1,SIZE(Dis,1)
      Aux(:)=AtDis(i,j,:)
      Dist=SQRT(SUM(Aux*Aux))
      Campo(1,:)=Campo(1,)+Dis(j,4)*Aux(:)/(Dist*Dist*Dist)
    END DO
  END DO
END IF

```

```

CampoMedio(Ciclo,1,:)=CampoMedio(Ciclo,1,:)+Campo(1,:)
CampoMedio(Ciclo,2,:)=CampoMedio(Ciclo,2,:)+Campo(1,1)*Campo(1,:)
CampoMedio(Ciclo,3,:)=CampoMedio(Ciclo,3,:)+Campo(1,2)*Campo(1,:)
CampoMedio(Ciclo,4,:)=CampoMedio(Ciclo,4,:)+Campo(1,3)*Campo(1,:)
DEALLOCATE(Campo)

!Calcula el potencial generado sobre la malla del soluto
DO i=1,SIZE(Puntos,1)
  Pot1=0.0D0
  Pot2=0.0D0
  DO j=1,Mols
    DO k=1,SIZE(Dis,1)
      Pot1=Pot1+Dis(k,4)/SQRT(SUM((Puntos(i,1:3)-AtDis(j,k,:))**2))
    END DO
    IF (Polarizacion == 'SI') THEN
      Aux(:)=Puntos(i,1:3)-CentroMasas(j,:)
      Dist=SQRT(SUM(Aux*Aux))
      Pot2=Pot2+SUM(Dipolos(j,:)*Aux(:))/(Dist*Dist*Dist)
    END IF
  END DO
  Puntos(i,4)=Puntos(i,4)+Pot1+Pot2
  Puntos(i,5)=Puntos(i,5)+Pot1
END DO

!Calcula las interacciones medias (MM) soluto-disolvente
SELECT CASE(TRIM(TipoPotencial))
  CASE('lennard-jones')
    NTipoPot=0
  CASE('generic')
    NTipoPot=1
END SELECT
Pot1=0.0D0
Pot2=0.0D0
Pot3=0.0D0
Sol2(:,1)=SolutoAnt(:) %x
Sol2(:,2)=SolutoAnt(:) %y
Sol2(:,3)=SolutoAnt(:) %z
Sol2(:,4)=SolutoAnt(:) %q
DO i=1,SIZE(Sol,1)
  ii=Soluto(i) %id
  DO j=1,Mols
    DO k=1,SIZE(Dis,1)
      !Interaccion q-q
      Dist=SQRT(SUM((AtDis(j,k,:)-Sol(i,1:3))**2))
      Pot1=Pot1+Sol(i,4)*Dis(k,4)/Dist
      Pot2=Pot2+Sol2(i,4)*Dis(k,4)/Dist
      !Interaccion de Van der Waals
      kk=Disolvente(k) %id
      IF (Potencial(kk,ii,1) < 0.1D0) CYCLE
      SELECT CASE(NTipoPot)
        CASE (0)
          Eps=Potencial(kk,ii,2)
          Sig6=(Potencial(kk,ii,3)/Dist)**6
          Pot3=Pot3+Eps*(Sig6*Sig6-Sig6)
        CASE (1)
          Dist4=Dist*Dist*Dist*Dist
          Dist6=Dist4*Dist*Dist
          Dist8=Dist4*Dist4
      END SELECT
    END DO
  END DO
END DO

```



```

        Dist12=Dist6*Dist6
        Pot3=Pot3+Potencial(kk,ii,1)*EXP(-Potencial(kk,ii,2)*Dist)+ &
            Potencial(kk,ii,3)/Dist12-Potencial(kk,ii,4)/Dist4- &
            Potencial(kk,ii,5)/Dist6-Potencial(kk,ii,6)/Dist8
    END SELECT
END DO
END DO
END DO
Pot4=Pot3
!Si hay cambio de geometria se calculan las interacciones con la
!geometria del ciclo anterior
IF (MaxIterOpt /= 0) THEN
    Pot2=0.0D0
    Pot4=0.0D0
    DO i=1,SIZE(Sol2,1)
        ii=SolutAnt(i)%id
        DO j=1,Mols
            DO k=1,SIZE(Dis,1)
                !Interaccion q-q
                Dist=SQRT(SUM((AtDis(j,k,:)-Sol2(i,1:3))**2))
                Pot2=Pot2+Sol2(i,4)*Dis(k,4)/Dist
                !Interaccion de Van der Waals
                kk=Disolvente(k)%id
                IF (Potencial(kk,ii,1) < 0.1D0) CYCLE
                SELECT CASE(NTipoPot)
                    CASE (0)
                        Eps=Potencial(kk,ii,2)
                        Sig6=(Potencial(kk,ii,3)/Dist)**6
                        Pot4=Pot4+Eps*(Sig6*Sig6-Sig6)
                    CASE (1)
                        Dist4=Dist*Dist*Dist*Dist
                        Dist6=Dist4*Dist*Dist
                        Dist8=Dist4*Dist4
                        Dist12=Dist6*Dist6
                        Pot4=Pot4+Potencial(kk,ii,1)*EXP(-Potencial(kk,ii,2)*Dist)+ &
                            Potencial(kk,ii,3)/Dist12-Potencial(kk,ii,4)/Dist4- &
                            Potencial(kk,ii,5)/Dist6-Potencial(kk,ii,6)/Dist8
                END SELECT
            END DO
        END DO
    END DO
END IF

IntMM(Ciclo,1)=IntMM(Ciclo,1)+Pot1      !Interaccion q-q
IntMM(Ciclo,3)=IntMM(Ciclo,3)+Pot3      !Interaccion de Van der Waals
IntMM(Ciclo,4)=IntMM(Ciclo,4)+ &      !Variacion de energia libre
    EXP((Pot1+Pot3-Pot2-Pot4)/(KBoltzmann*Temperatura))

IF (Polarizacion == 'SI') DEALLOCATE(Dipolos,Pol)
DEALLOCATE(CentroMasas,AtDis,Dis,Sol,Sol2)

102  FORMAT (A,F12.8)
103  FORMAT (80('-'))
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE CalcularPotencial
!-----

```

```

! Lee las posiciones de los atomos del fichero donde estan guardadas. Tambien
! une las cargas de la capa de solvatacion que estan suficientemente cercanas
! entre si.
!-----
! Cargas:      Cargas correspondientes a las configuraciones del disolvente
! Numero:      Matriz para saber cuantas cargas ocupan la misma posicion
! Dist:        Distancia entre dos puntos al cuadrado
! Pos:         Posicion del centro de masas de la molecula que se esta leyendo
! Prox:        "Proximidad" elevado al cuadrado
! Rad:         "RadioCapa" elevado al cuadrado
! Preci:       Precision para considerar las cargas nulas
! NumCargas:  Numero de cargas en la capa de solvatacion
! Error:       Variable para controlar los errores
! i,j,k:       Contadores
!-----
SUBROUTINE ReducirCapa
  USE Parametros
  USE Sistema, ONLY: Disolvente
  USE Resultados, ONLY: CargasCapa,Puntos,ConfigDisolv,ConfigNum
  IMPLICIT NONE
  INTEGER, DIMENSION(:), ALLOCATABLE :: Numero
  DOUBLE PRECISION, DIMENSION(:), ALLOCATABLE :: Cargas
  DOUBLE PRECISION, DIMENSION(3) :: Pos
  DOUBLE PRECISION :: Dist,Prox,Rad
  DOUBLE PRECISION, PARAMETER :: Preci=1.0E-10
  INTEGER :: NumCargas,Error,i,j,k

  Prox=Proximidad*Proximidad
  Rad=RadioCapa*RadioCapa
  NumCargas=ConfigNum(NumConfig)

  !Lee las posiciones de todos los atomos, si el centro de masas queda
  !fuera de la capa de solvatacion, elimina la molecula
  REWIND(32)
  IF (ALLOCATED(ConfigDisolv)) DEALLOCATE(ConfigDisolv)
  ALLOCATE(ConfigDisolv(NumCargas,3),Numero(NumCargas),Cargas(NumCargas), &
    STAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_crear_la_matriz_con_todos_los_atomos_del_
      disolvente.'
    STOP
  END IF
  Numero(:)=1
  DO i=1,NumConfig
    DO j=ConfigNum(i-1),ConfigNum(i)-1,SIZE(Disolvente)
      READ(32,*) Pos(:)
      DO k=1,SIZE(Disolvente)
        READ(32,*) ConfigDisolv(j+k,1:3)
        Cargas(j+k)=Disolvente(k) %q
      END DO
      IF (SUM(Pos*Pos) > Rad) Numero(j+1:j+SIZE(Disolvente))=0
    END DO
    READ(32,*)
  END DO
  Cargas(:)=Cargas(:)/NumConfig

  IF (PrimeraCapa == 'SI') THEN
    !Si dos atomos estan suficientemente cerca, suma las cargas

```

```

DO i=1,NumCargas-1
  IF (Numero(i) == 0) CYCLE
  DO j=i+1,NumCargas
    IF (Numero(j) == 0) CYCLE
    Dist=SUM((ConfigDisolv(i,1:3)-ConfigDisolv(j,1:3))**2)
    IF (Dist <= Prox) THEN
      Cargas(i)=Cargas(i)+Cargas(j)
      Numero(i)=Numero(i)+1
      Numero(j)=0
    END IF
  END DO
  IF (ABS(Cargas(i)) <= Preci) Numero(i)=0
END DO

!Se guarda la capa de solvatacion reducida
IF (ALLOCATED(CargasCapa)) DEALLOCATE(CargasCapa)
ALLOCATE(CargasCapa(COUNT(Numero /= 0),5),STAT=Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'Error_al_crear_la_matriz_con_las_cargas_de_la_capa_de_
  solvatacion.'
  STOP
END IF
j=0
DO i=1,SIZE(Cargas,1)
  IF (Numero(i) == 0) CYCLE
  j=j+1
  CargasCapa(j,1:3)=ConfigDisolv(i,:)
  CargasCapa(j,4)=Cargas(i)
END DO
END IF

DEALLOCATE(Numero,Cargas)
IF (MaxIterOpt == 0) DEALLOCATE(ConfigDisolv)

666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE ReducirCapa

!-----
! Calcula la fluctuacion del campo electrico en el centro de masas del soluto
!-----
! CentroMasas: Centro de masas del soluto
! Campo:      Campo electrico de cada configuracion
! Aux:       Vector auxiliar
! Dist:      Distancia de los atomos del disolvente al centro de masas
! i, j, k:   Contadores
!-----
SUBROUTINE CalcularCampo
  USE Parametros, ONLY: Ciclo,NumConfig
  USE Sistema, ONLY: Soluto,Disolvente
  USE Resultados, ONLY: ConfigDisolv,ConfigNum,CampoMedio
  IMPLICIT NONE
  DOUBLE PRECISION, DIMENSION(3) :: CentroMasas,Campo,Aux
  DOUBLE PRECISION :: Dist
  INTEGER :: i,j,k

  !Se calcula el centro de masas del soluto
  CentroMasas(1)=SUM(Soluto(:) %x*Soluto(:) %Masa)/SUM(Soluto(:) %Masa)

```

```

CentroMasas(2)=SUM(Soluto(:) %y*Soluto(:) %Masa)/SUM(Soluto(:) %Masa)
CentroMasas(3)=SUM(Soluto(:) %z*Soluto(:) %Masa)/SUM(Soluto(:) %Masa)

!Se calcula el campo electrico y el campo al cuadrado promedios de todas
!las configuraciones
DO i=1,NumConfig
  Campo=0.0D0
  DO j=ConfigNum(i-1)+1,ConfigNum(i)
    k=MOD(j-1,SIZE(Disolvente))+1
    Aux(:)=ConfigDisolv(j,:)-CentroMasas(:)
    Dist=SQRT(SUM(Aux*Aux))
    Campo(:)=Campo(:)+Disolvente(k) %q*Aux(:)/(Dist*Dist*Dist)
  END DO
  CampoMedio(Ciclo,1,:)=CampoMedio(Ciclo,1,:)+Campo(:)
  CampoMedio(Ciclo,2,:)=CampoMedio(Ciclo,2,:)+Campo(1)*Campo(:)
  CampoMedio(Ciclo,3,:)=CampoMedio(Ciclo,3,:)+Campo(2)*Campo(:)
  CampoMedio(Ciclo,4,:)=CampoMedio(Ciclo,4,:)+Campo(3)*Campo(:)
END DO
CampoMedio(Ciclo,(:, :))=CampoMedio(Ciclo,(:, :))/NumConfig

END SUBROUTINE CalcularCampo

!-----
! Ajusta los valores de una serie de cargas para reproducir un potencial dado
! por el metodo de los multiplicadores de Lagrange (restriccion de carga total)
!-----
! Tipo:           Que se va ajustar: 'Dis', 'QDi' o 'Sol'
! Cargas:        Matriz con las posiciones de las cargas a ajustar
! PuntosA:       Matriz el potencial al que hay que ajustar las cargas
! Ecuaciones:    Coeficientes del sistema de ecuaciones a resolver
! TerminosI:    Terminos independientes del sistema de ecuaciones
! NumCargas:     Numero de cargas a ajustar
! Dist1,Dist2:   Distancias entre dos puntos
! Err:           Error relativo medio del ajuste
! Rms:          Error cuadratico medio del ajuste
! RRms:         Error cuadratico medio relativo del ajuste
! Error:        Variable para controlar los errores
! i, j, k:      Contadores
!-----
SUBROUTINE AjusteLagrange(Tipo)
  USE Parametros
  USE Sistema, ONLY: Soluto,Carga
  USE Resultados, ONLY: CargasDisolvente,CargasCapa,Puntos, &
    AngstromAtomica
  IMPLICIT NONE
  INTERFACE
    SUBROUTINE ResolverLU(Coef,Term,Num,Error)
      DOUBLE PRECISION, DIMENSION(Num,Num), INTENT(INOUT) :: Coef
      DOUBLE PRECISION, DIMENSION(Num), INTENT(INOUT) :: Term
      INTEGER, INTENT(IN) :: Num
      INTEGER, INTENT(OUT) :: Error
    END SUBROUTINE ResolverLU
  END INTERFACE
  CHARACTER, INTENT(IN) :: Tipo*3
  DOUBLE PRECISION, DIMENSION(:, :), ALLOCATABLE :: Ecuaciones,Cargas,PuntosA
  DOUBLE PRECISION, DIMENSION(:), ALLOCATABLE :: TerminosI
  DOUBLE PRECISION :: Dist1,Dist2,Err,Rms,RRms
  INTEGER :: NumCargas,Error,i,j,k

```

```

!Se toman los datos iniciales segun cual sea el ajuste
SELECT CASE(Tipo)
!Si se ajusta el disolvente, el potencial esta calculado en una malla y
!las posiciones de las cargas son las de "CargasDisolvente". Si se
!ajustan las cargas permanentes es igual pero se toma el potencial gene-
!rado por las cargas permanentes
CASE ('Dis','QDi')
  ALLOCATE(PuntosA(SIZE(Puntos,1),5),STAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_crear_la_matriz_con_los_puntos_del_potencial.'
    STOP
  END IF
  IF (Tipo == 'Dis') THEN
    PuntosA(:,1:4)=Puntos(:,1:4)
  ELSE
    PuntosA(:,1:3)=Puntos(:,1:3)
    PuntosA(:,4)=Puntos(:,5)
  END IF
  PuntosA(:,5)=0.0D0
  NumCargas=SIZE(CargasDisolvente,1)
  ALLOCATE(Cargas(NumCargas,4),STAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_crear_la_matriz_con_las_cargas_para_ajustar.'
    STOP
  END IF
  Cargas(:,1:3)=CargasDisolvente(:,1:3)
!Si se ajusta el soluto, el potencial esta calculado en las cargas del
!del disolvente y las cargas estan en los atomos del soluto, ademas hay
!que tener en cuenta las equivalencias
CASE ('Sol')
  IF (Ciclo > 0 .AND. PrimeraCapa == 'SI') THEN
    ALLOCATE(PuntosA(SIZE(CargasDisolvente,1)+SIZE(CargasCapa,1),5), &
      STAT=Error)
  ELSE
    ALLOCATE(PuntosA(SIZE(CargasDisolvente,1),5),STAT=Error)
  END IF
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_crear_la_matriz_con_los_puntos_del_potencial.'
    STOP
  END IF
  PuntosA(1:SIZE(CargasDisolvente,1),1:3)=CargasDisolvente(:,1:3)
  PuntosA(1:SIZE(CargasDisolvente,1),4)=CargasDisolvente(:,6)
  IF (Ciclo > 0 .AND. PrimeraCapa == 'SI') THEN
    PuntosA(SIZE(CargasDisolvente,1)+1:,1:3)=CargasCapa(:,1:3)
    PuntosA(SIZE(CargasDisolvente,1)+1:,4)=CargasCapa(:,5)
  END IF
  PuntosA(:,5)=0.0D0
  ALLOCATE(Cargas(SIZE(Soluto),4),STAT=Error)
  IF (Error /= 0) THEN
    WRITE(21,666) 'Error_al_crear_la_matriz_con_las_cargas_para_ajustar.'
    STOP
  END IF
  Cargas(:,1)=Soluto(:) %x
  Cargas(:,2)=Soluto(:) %y
  Cargas(:,3)=Soluto(:) %z
END SELECT

```

```

!Se construyen las ecuaciones:
!a(i,j)=SUMA[1/(r(ik)*r(jk))] r(ik) -> distancia entre carga i y punto k
!b(i)=SUMA[V(k)/r(ik)] V(k) -> potencial en el punto k
!Ademas se annade la condicion de la carga total (0 para el disolvente)
ALLOCATE(Ecuaciones(SIZE(Cargas,1)+1,SIZE(Cargas,1)+1))
ALLOCATE(TerminosI(SIZE(Cargas,1)+1))
Ecuaciones=0.0D0
TerminosI=0.0D0
DO k=1,SIZE(PuntosA,1)
  DO i=1,SIZE(Cargas,1)
    Dist1=SQRT(SUM((Cargas(i,1:3)-PuntosA(k,1:3))**2))
    TerminosI(i)=TerminosI(i)+PuntosA(k,4)/Dist1
    DO j=1,SIZE(Cargas,1)
      Dist2=SQRT(SUM((Cargas(j,1:3)-PuntosA(k,1:3))**2))
      Ecuaciones(i,j)=Ecuaciones(i,j)+1.0D0/(Dist1*Dist2)
    END DO
    Ecuaciones(i,SIZE(Cargas,1)+1)=1.0D0
    Ecuaciones(SIZE(Cargas,1)+1,i)=1.0D0
  END DO
END DO
IF (Tipo == 'Sol') TerminosI(SIZE(Cargas,1)+1)=DBLE(Carga)

!Se resuelve el sistema de ecuaciones
CALL ResolverLU(Ecuaciones,TerminosI,SIZE(Cargas,1)+1,Error)
IF (Error /= 0) THEN
  WRITE(21,666) 'La_matriz_para_el_ajuste_de_las_cargas_es_singular.'
  STOP
END IF
Cargas(:,4)=TerminosI(1:SIZE(Cargas,1))

!Se calculan los errores (la columna 5 de PuntosA contiene el potencial
!calculado con las cargas ajustadas)
!Err=SUMA[|difpot/pot|]/n difpot -> dif entre pot real y calculado
!Rms=SQRT[SUMA[(difpot)**2]/n] n -> numero de puntos en el potencial
!RRms=SQRT[SUMA[(difpot)**2]/SUMA[pot]**2] pot -> potencial real
Err=0.0D0
Rms=0.0D0
RRms=0.0D0
DO k=1,SIZE(PuntosA,1)
  DO i=1,SIZE(Cargas,1)
    Dist1=SQRT(SUM((Cargas(i,1:3)-PuntosA(k,1:3))**2))
    PuntosA(k,5)=PuntosA(k,5)+Cargas(i,4)/Dist1
  END DO
END DO
Err=SUM(ABS((PuntosA(:,5)-PuntosA(:,4))/PuntosA(:,4)))/SIZE(PuntosA,1))
Rms=SQRT(SUM((PuntosA(:,5)-PuntosA(:,4))**2)/SIZE(PuntosA,1))
RRms=SUM(((PuntosA(:,5)-PuntosA(:,4))/PuntosA(:,4))**2)/SIZE(PuntosA,1))
RRms=SQRT(RRms)

!Se da la salida segun el tipo de ajuste
SELECT CASE(Tipo)
!Si se ajusta el disolvente, las cargas van a "CargasDisolvente"
CASE ('Dis')
  WRITE(21,301) 'Ajuste_al_potencial_total_del_disolvente:',Err,Rms,RRms,&
    SUM(Cargas(:,4))
  CargasDisolvente(:,4)=Cargas(:,4)
!Si se ajusta a las cargas permanentes del disolvente, tambien va a
!"CargasDisolvente", pero a la columna 5

```

```

CASE ('QDi')
  WRITE(21,301) 'Ajuste_al_potencial_de_las_cargas_permanentes:',Err, &
    Rms,RRms,SUM(Cargas(:,4))
  CargasDisolvente(:,5)=Cargas(:,4)
  !Si se ajusta el soluto, las cargas van a "Soluto"
CASE ('Sol')
  WRITE(21,301) 'Ajuste_al_potencial_total_del_soluto:',Err,Rms,RRms, &
    SUM(Cargas(:,4))
  Soluto(:) %q=Cargas(:,4)
END SELECT

DEALLOCATE(Cargas,PuntosA,Ecuaciones,TerminosI)

301  FORMAT (80('-')/A/'Error:',F14.10,'_Rms:',F14.10,'_RRms:',F14.10/ &
  'Carga_total:',F14.10/80('-'))
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE AjusteLagrange

!-----
! Resuelve un sistema de ecuaciones lineales por el metodo de descomposicion LU
!-----
! Coef:  Matriz con los coeficientes de las ecuaciones, a la salida contiene
!        la descomposicion LU
! Term:  Vector con los terminos independientes, a la salida contiene las sol.
! Num:   Numero de ecuaciones e incognitas en el sistema
! Error: 0 si no pasa nada, 1 si la matriz es singular
! Factor: Vector con los factores para la normalizacion de las ecuaciones
! Preci: Precision para compararar con cero
! Aux:   Variable axiliar para los intercambios
! p:     Numero de la ecuacion que hara de pivote
! i,j:   Contadores
!-----
SUBROUTINE ResolverLU(Coef,Term,Num,Error)
  IMPLICIT NONE
  INTEGER, INTENT(IN) :: Num
  DOUBLE PRECISION, DIMENSION(Num,Num), INTENT(INOUT) :: Coef
  DOUBLE PRECISION, DIMENSION(Num), INTENT(INOUT) :: Term
  INTEGER, INTENT(OUT) :: Error
  DOUBLE PRECISION, DIMENSION(Num) :: Factor
  DOUBLE PRECISION :: Aux
  INTEGER :: i,j,p
  DOUBLE PRECISION, PARAMETER :: Preci=1.0D-20

  Error=0
  !Se calculan los factores de normalizacion, si hay alguna fila de ceros,
  !la matriz es singular
  Factor=MAXVAL(ABS(Coef),DIM=2)
  IF (ANY(Factor < Preci)) THEN
    Error=1
    RETURN
  END IF
  Factor=1.0D0/Factor

  !Se localiza la mejor ecuacion para hacer de pivote y se pone en el lugar
  !adecuado, tambien se intercambian los terminos independientes
  !Despues se modifican los coeficientes
  DO i=1,Num

```

```

p=i-1+SUM(MAXLOC (Factor (i:Num) *ABS (Coef (i:Num, i) )))
IF (i /= p) THEN
  DO j=1, Num
    Aux=Coef (i, j)
    Coef (i, j)=Coef (p, j)
    Coef (p, j)=Aux
  END DO
  Aux=Term (i)
  Term (i)=Term (p)
  Term (p)=Aux
  Factor (p)=Factor (i)
END IF
IF (ABS (Coef (i, i)) < Preci) THEN
  Error=1
  RETURN
END IF
Coef (i+1:Num, i)=Coef (i+1:Num, i)/Coef (i, i)
DO j=i+1, Num
  Coef (i+1:Num, j)=Coef (i+1:Num, j)-Coef (i+1:Num, i)*Coef (i, j)
END DO
END DO

!Se resuelve el sistema por doble sustitucion (primero directa, y luego
!inversa)
DO i=2, Num
  Term (i)=Term (i)-SUM (Term (1:i-1)*Coef (i, 1:i-1))
END DO
Term (Num)=Term (Num)/Coef (Num, Num)
DO i=Num-1, 1, -1
  Term (i)=(Term (i)-SUM (Term (i+1:Num)*Coef (i, i+1:Num)))/Coef (i, i)
END DO

END SUBROUTINE ResolverLU

!-----
!Desplaza y gira una molecula para colocarla centrada en su centro de masas
!y orientada segun sus ejes principales de inercia
!-----
!Tipo: Que molecula se va a mover ('Sol' o 'Dis')
!Ordenar: Dice si se ordenan o no los ejes segun el momento de inercia
!Molecula: Matriz con las posiciones y masas de los atomos de la molecula
!CentroMasas: Posicion del centro de masas antes de modificar la molecula
!Inercia: Tensor de inercia de la molecula original
!Ejes: Matriz de vectores propios del tensor de inercia (ejes princip.)
!Aux, Aux2: Variable y vector auxiliares
!Error: Variable para controlar los errores
!i, j, k: Contadores
!-----

SUBROUTINE GirarMolecula (Tipo, Ordenar)
  USE Sistema, ONLY: Soluto, Disolvente, PolDisolvente
  IMPLICIT NONE
  INTERFACE
    SUBROUTINE Diagonalizar (Mat, Vec, Val, Ord, Error)
      DOUBLE PRECISION, DIMENSION (:, :), INTENT (INOUT) :: Mat
      DOUBLE PRECISION, DIMENSION (:, :), INTENT (OUT) :: Vec
      DOUBLE PRECISION, DIMENSION (:), INTENT (OUT) :: Val
      CHARACTER*2, INTENT (IN) :: Ord
      INTEGER, INTENT (OUT) :: Error
    END SUBROUTINE
  END INTERFACE

```



```

      END SUBROUTINE Diagonalizar
END INTERFACE
CHARACTER, INTENT(IN) :: Tipo*3,Ordenar*2
DOUBLE PRECISION, DIMENSION(:,,:), ALLOCATABLE :: Molecula
DOUBLE PRECISION, DIMENSION(3) :: CentroMasas,Aux2
DOUBLE PRECISION, DIMENSION(3,3) :: Inercia,Ejes
DOUBLE PRECISION :: Aux
INTEGER :: i,j,k,Error

!Se copia la molecula que se va a modificar
SELECT CASE(Tipo)
CASE ('Dis')
  ALLOCATE(Molecula(SIZE(Disolvente),4))
  Molecula(:,1)=Disolvente(:) %x
  Molecula(:,2)=Disolvente(:) %y
  Molecula(:,3)=Disolvente(:) %z
  Molecula(:,4)=Disolvente(:) %Masa
CASE ('Sol')
  ALLOCATE(Molecula(SIZE(Soluto),4))
  Molecula(:,1)=Soluto(:) %x
  Molecula(:,2)=Soluto(:) %y
  Molecula(:,3)=Soluto(:) %z
  Molecula(:,4)=Soluto(:) %Masa
END SELECT

!Se desplaza la molecula para dejar el centro de masas en el 0,0,0
DO i=1,3
  CentroMasas(i)=SUM(Molecula(:,i)*Molecula(:,4))/SUM(Molecula(:,4))
  Molecula(:,i)=Molecula(:,i)-CentroMasas(i)
END DO

!Se calcula el tensor de inercia de la molecula centrada
Aux=SUM((Molecula(:,1)*Molecula(:,1)+Molecula(:,2)*Molecula(:,2)+ &
  Molecula(:,3)*Molecula(:,3))*Molecula(:,4))
DO i=1,3
  Inercia(i,i)=Aux-SUM(Molecula(:,i)*Molecula(:,i)*Molecula(:,4))
  DO j=i+1,3
    Inercia(i,j)=-SUM(Molecula(:,i)*Molecula(:,j)*Molecula(:,4))
    Inercia(j,i)=Inercia(i,j)
  END DO
END DO

!Se diagonaliza el tensor de inercia
CALL Diagonalizar(Inercia,Ejes,Aux2,Ordenar,Error)

IF (Ordenar == 'Si') THEN
  !Mantiene la quiralidad, comprobando que el x e2 = e3
  Aux2(1)=(Ejes(2,1)*Ejes(3,2)-Ejes(2,2)*Ejes(3,1))/Ejes(1,3)
  Aux2(2)=(Ejes(3,1)*Ejes(1,2)-Ejes(3,2)*Ejes(1,1))/Ejes(2,3)
  Aux2(3)=(Ejes(1,1)*Ejes(2,2)-Ejes(1,2)*Ejes(2,1))/Ejes(3,3)
  IF (Aux2(1) < 0 .AND. Aux2(2) < 0 .AND. Aux2(3) < 0) &
    Ejes(:,3)=-Ejes(:,3)
ELSE
  !Se asegura de que los valores de la diagonal son positivos
  !Solo funciona si la matriz de rotacion puede ser aproximadamente 1
  DO i=1,3
    k=SUM(MAXLOC(ABS(Ejes(i,:))))
    IF (k /= i) THEN

```

```

        DO j=1,3
            Aux=Ejes(j,i)
            Ejes(j,i)=Ejes(j,k)
            Ejes(j,k)=Aux
        END DO
    END IF
    IF (Ejes(i,i) < 0.0D0) Ejes(:,i)=-Ejes(:,i)
END DO
END IF

!Gira la molecula
Molecula(:,1:3)=MATMUL(Molecula(:,1:3),Ejes)

!Se copia la molecula modificada en el lugar adecuado
SELECT CASE(Tipo)
CASE ('Dis')
    Disolvente(:) %x=Molecula(:,1)
    Disolvente(:) %y=Molecula(:,2)
    Disolvente(:) %z=Molecula(:,3)
    PolDisolvente=MATMUL(PolDisolvente,TRANSPPOSE(Ejes))
    PolDisolvente=MATMUL(Ejes,PolDisolvente)
CASE ('Sol')
    Soluto(:) %x=Molecula(:,1)
    Soluto(:) %y=Molecula(:,2)
    Soluto(:) %z=Molecula(:,3)
END SELECT
DEALLOCATE(Molecula)

END SUBROUTINE GirarMolecula

!-----
! Esta subrutina diagonaliza una matriz por el metodo de Jacobi y devuelve
! sus vectores y valores propios, ordenados o no
!-----
! Mat:      Matriz a diagonalizar
! Val:      Vector que contiene los valores propios
! Vec:      Matriz que contiene los vectores propios
! Ord:      'SI' si se ordenan los vectores propios, 'NO' si no se hace
! Num:      Dimension de las matrices y vectores
! Max:      Numero maximo de vueltas para diagonalizar
! Suma:     Suma de los elementos extradiagonales
! Tol:      Tolerancia para hacer comparaciones
! Zeta,Tau: Variables para la diagonalizacion
! t,c,s,h:  Variables auxiliares para la diagonalizacion
! e:        Vector auxiliar para la diagonalizacion
! Error:    Variable para controlar los errores
! i,j,k:    Contadores
!-----
SUBROUTINE Diagonalizar(Mat,Vec,Val,Ord,Error)
    IMPLICIT NONE
    DOUBLE PRECISION, DIMENSION(:, :), INTENT(INOUT) :: Mat
    DOUBLE PRECISION, DIMENSION(:, :), INTENT(OUT) :: Vec
    DOUBLE PRECISION, DIMENSION(:), INTENT(OUT) :: Val
    CHARACTER*2, INTENT(IN) :: Ord
    INTEGER, INTENT(OUT) :: Error
    DOUBLE PRECISION, DIMENSION(:), ALLOCATABLE :: e
    DOUBLE PRECISION :: t,c,s,h,Suma,Tol,Zeta,Tau
    INTEGER :: i,j,k,Num,Max

```

```

!Se inician los datos
Max=5000
Error=1
Num=SIZE(Mat,1)
ALLOCATE(e(Num))
Vec(:, :)=0.0D0
DO i=1,Num
  Vec(i,i)=1.0D0
  Val(i)=Mat(i,i)
END DO

!Se diagonaliza la matriz
DO i=1,Max
  Suma=0.0D0
  DO j=1,Num-1
    Suma=Suma+SUM(ABS(Mat(j,j+1:Num)))
  END DO
  IF (Suma == 0.0D0) EXIT
  DO j=1,Num-1
    DO k=j+1,Num
      h=Mat(j,k)
      Tol=100.0D0*ABS(h)
      IF (i < 4 .AND. Tol < 20.0D0*Suma/(Num*Num)) CYCLE
      IF (Val(k)+Tol == Val(k) .AND. Val(j)+Tol == Val(j)) THEN
        Mat(j,k) = 0.0D0
        CYCLE
      END IF
      Zeta=0.5D0*(Val(k)-Val(j))/h
      IF (Val(k)-Val(j)+Tol == Val(k)-Val(j)) THEN
        t=h/(Val(k)-Val(j))
      ELSE
        t=1.0D0/(ABS(Zeta)+SQRT(Zeta*Zeta+1.0D0))
        IF (Zeta < 0.0D0) t=-t
      END IF
      c=1.0D0/SQRT(t*t+1.0D0)
      s=t*c
      Tau=s/(1.0D0+c)
      Val(j)=Val(j)-t*h
      Val(k)=Val(k)+t*h
      e(1:j-1)=Mat(1:j-1,j)
      e(j+1:Num)=Mat(j,j+1:Num)
      Mat(1:j-1,j)=e(1:j-1)-s*(Mat(1:j-1,k)+Tau*e(1:j-1))
      Mat(1:j-1,k)=Mat(1:j-1,k)+s*(e(1:j-1)-Tau*Mat(1:j-1,k))
      Mat(j,j+1:k-1)=e(j+1:k-1)-s*(Mat(j+1:k-1,k)+Tau*e(j+1:k-1))
      Mat(j+1:k-1,k)=Mat(j+1:k-1,k)+s*(e(j+1:k-1)-Tau*Mat(j+1:k-1,k))
      Mat(j,k+1:Num)=e(k+1:Num)-s*(Mat(k,k+1:Num)+Tau*e(k+1:Num))
      Mat(k,k+1:Num)=Mat(k,k+1:Num)+s*(e(k+1:Num)-Tau*Mat(k,k+1:Num))
      Mat(j,k)=0.0D0
      e(:)=Vec(:,j)
      Vec(:,j)=e(:)-s*(Vec(:,k)+Tau*e(:))
      Vec(:,k)=Vec(:,k)+s*(e(:)-Tau*Vec(:,k))
    END DO
  END DO
END DO

IF (Suma == 0.0D0) Error=0

```

```
!Se ordenan los vectores y valores propios
IF (Ord == 'SI') THEN
  DO i=1, Num-1
    j=SUM(MINLOC(Val(i:Num)))+i-1
    IF (j /= i) THEN
      e(1)=Val(i)
      Val(i)=Val(j)
      Val(j)=e(1)
      e(:)=Vec(:,i)
      Vec(:,i)=Vec(:,j)
      Vec(:,j)=e(:)
    END IF
  END DO
END IF

END SUBROUTINE Diagonalizar
```

## F.7. Optim.f90

```

!CONTIENE:
!OptimizarGeometria
!CalcularGradHessPot (Tipo)
!BuscarMinimo
!EnergPunto (Factor)
!ActualizarHessiana
!CalcularIncremento
!EscribeOptim(Iter)

!-----
! Esta subrutina optimiza la geometria del soluto teniendo en cuenta las fuerzas
! sobre los atomos ejercidas por el disolvente (cargas y potencial intermolecu-
! lar).
!-----
! 22:   Fichero donde se escribe cada paso de la optimizacion
! 32:   Fichero que contiene las configuraciones del disolvente
! Num:  Numero de coordenadas del soluto (3*Atomos)
! Iter:  Numero de iteracion dentro de la optimizacion
! Fin:   Variable para controlar cuando termina la optimizacion
! Error: Variable para controlar los errores
! i:    Contador
!-----
SUBROUTINE OptimizarGeometria
  USE Parametros
  USE Sistema, ONLY: Soluto
  USE Optimizacion
  USE Resultados, ONLY: Energia, EPol, CargasDisolvente, CargasCapa
  IMPLICIT NONE
  INTERFACE
    SUBROUTINE EscribeOptim(Iter)
      INTEGER, INTENT(IN) :: Iter
    END SUBROUTINE EscribeOptim
    SUBROUTINE EjecutarAbInicio(Tipo)
      CHARACTER*4, INTENT(IN) :: Tipo
    END SUBROUTINE EjecutarAbInicio
    SUBROUTINE LeerAbInicio(Tipo)
      CHARACTER*4, INTENT(IN) :: Tipo
    END SUBROUTINE LeerAbInicio
    SUBROUTINE CalcularGradHessPot(Tipo)
      CHARACTER*4, INTENT(IN) :: Tipo
    END SUBROUTINE CalcularGradHessPot
  END INTERFACE
  INTEGER :: i, j, Num, Iter, Fin, Error

  Num=3*SIZE(Soluto)
  ALLOCATE(Gradiente(Num), GradPot(Num), Incremento(Num), Coordenadas(Num), &
    Hessiana(Num,Num), GradAnt(Num))
  IF (CalcHessiana /= 0) ALLOCATE(HessPot(Num,Num))

  !Se abre el fichero donde se escribiran los resultados de la optimizacion
  OPEN(22, STATUS='REPLACE', ACTION='WRITE', FILE='optim'//Extension, &
    IOSTAT=Error)
  IF (Error /= 0) THEN
    WRITE(21, 666) 'Error_al_abrir_el_fichero_de_salida_de_la_optimizacion.'
    STOP

```

```

END IF
WRITE(21,101) 'Optimizacion_de_la_geometria...'

!Se toman las coordenadas iniciales para optimizar
DO i=1,Num,3
  Coordenadas(i)=Solut((i+2)/3) %x
  Coordenadas(i+1)=Solut((i+2)/3) %y
  Coordenadas(i+2)=Solut((i+2)/3) %z
END DO

!Se realiza el calculo inicial de energia, gradientes y hessiana (si es
!necesaria). Se suman las contribuciones del potencial intermolecular
IF (CalcHessiana == 0) THEN
  CALL EjecutarAbInicio('Grad')
  CALL LeerAbInicio('Grad')
  CALL CalcularGradHessPot('Grad')
  Gradiente=Gradiente+GradPot
  Hessiana(:, :)=0.0D0
  DO i=1,Num
    Hessiana(i,i)=1.0D0
  END DO
ELSE
  CALL EjecutarAbInicio('Hess')
  CALL LeerAbInicio('Hess')
  CALL CalcularGradHessPot('Hess')
  Gradiente=Gradiente+GradPot
  Hessiana=Hessiana+HessPot
END IF
EPol(Ciclo)=Energia(Ciclo)
EPol(Ciclo)=EPol(Ciclo)-SUM(CargasDisolvente(:,4)*CargasDisolvente(:,6))
IF (PrimeraCapa == 'SI') EPol(Ciclo)=EPol(Ciclo)- &
  SUM(CargasCapa(:,4)*CargasCapa(:,5))
CALL EscribeOptim(0)

Fin=0
IF (SQRT(SUM(Gradiente(:)*Gradiente(:))) < ConvGradiente) Fin=1
Iter=0
DO WHILE (Fin == 0 .AND. Iter < MaxIterOpt)
  Iter=Iter+1
  CALL CalcularIncremento
  Paso=MaxPaso*MaxPaso/SUM(Incremento*Incremento)

  !Se busca la minima energia en la direccion dada por "Incremento" o se
  !normaliza a "MaxPaso" si es mayor
  IF (BusquedaLineal == 'SI') THEN
    CALL BuscarMinimo
  ELSE IF (Paso < 1.0D0) THEN
    Incremento=Incremento*SQRT(Paso)
  END IF

  Coordenadas=Coordenadas+Incremento
  DO i=1,Num,3
    Solut((i+2)/3) %x=Coordenadas(i)
    Solut((i+2)/3) %y=Coordenadas(i+1)
    Solut((i+2)/3) %z=Coordenadas(i+2)
  END DO
  !Se calculan los gradientes y la hessiana si es necesario
  IF (CalcHessiana == 0) THEN

```

```

    GradAnt=Gradiente
    CALL EjecutarAbInitio('Grad')
    CALL LeerAbInitio('Grad')
    CALL CalcularGradHessPot('Grad')
    Gradiente=Gradiente+GradPot
    ELSE IF (MOD(Iter,CalcHessiana) == 0) THEN
    CALL EjecutarAbInitio('Hess')
    CALL LeerAbInitio('Hess')
    CALL CalcularGradHessPot('Hess')
    Gradiente=Gradiente+GradPot
    Hessiana=Hessiana+HessPot
    ELSE
    GradAnt=Gradiente
    CALL EjecutarAbInitio('Grad')
    CALL LeerAbInitio('Grad')
    CALL CalcularGradHessPot('Grad')
    Gradiente=Gradiente+GradPot
    END IF
    CALL EscribirOptim(Iter)
    !Se comprueba la convergencia
    IF (SQRT(SUM(Gradiente(:)*Gradiente(:))) < ConvGradiente) THEN
    Fin=1
    ELSE IF (SQRT(SUM(Incremento(:)*Incremento(:))) < MinPaso) THEN
    Fin=2
    !Si no hay convergencia, se actualiza la hessiana
    ELSE IF (CalcHessiana == 0 .OR. MOD(Iter,CalcHessiana) /= 0) THEN
    CALL ActualizarHessiana
    END IF
    END DO

    SELECT CASE (Fin)
    CASE (0)
    WRITE(21,101) 'Se_alcanzo_el_limite_maximo_de_iteraciones.'
    WRITE(22,101) 'Se_alcanzo_el_limite_maximo_de_iteraciones.'
    CASE (1)
    WRITE(21,101) 'Se_alcanzo_la_convergencia_en_el_gradiente.'
    WRITE(22,101) 'Se_alcanzo_la_convergencia_en_el_gradiente.'
    CASE (2)
    WRITE(21,101) 'La_variacion_de_la_geometria_no_alcanza_el_minimo.'
    WRITE(22,101) 'La_variacion_de_la_geometria_no_alcanza_el_minimo.'
    END SELECT

    WRITE(22,101)
    WRITE(22,101) 'Hessiana_(u.a.):'
    WRITE(22,*) ((Hessiana(i,j),j=1,i),i=1,Num)

    CLOSE(22)

    DEALLOCATE(Gradiente,GradPot,Incremento,Coordenadas,Hessiana,GradAnt)
    IF (CalcHessiana /= 0) DEALLOCATE(HessPot)

101  FORMAT (A)
666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE OptimizarGeometria

!-----
! Esta subrutina calcula la contribucion a la energia, gradiente y hessiana

```

```

! debida al potencial intermolecular especificado en la dinamica molecular
!-----
! Tipo:           Tipo de calculo a realizar ('Nada','Grad','Hess')
! NTipo:         Numero que indica el tipo de calculo
! Dist:         Distancia entre dos atomos
! Eps,Sig*,Dist*: Parametros para el potencial Lennard-Jones
! Par*:         Parametros para el potencial generico
! Aux1,Aux2:     Variables auxiliares para simplificar
! i,j,k,m,ii,jj: Contadores
!-----
SUBROUTINE CalcularGradHessPot(Tipo)
  USE Parametros
  USE Sistema, ONLY: Solute,Disolvente,TipoPotencial,Potencial
  USE Resultados, ONLY: ConfigDisolv
  USE Optimizacion
  IMPLICIT NONE
  CHARACTER*4, INTENT(IN) :: Tipo
  DOUBLE PRECISION, DIMENSION(3) :: Dist
  DOUBLE PRECISION :: Eps,Sig6,Sig12,Dist1,Dist2,Dist4,Par1,Par2,Par3,Par4,&
    Par5,Par6,Aux1,Aux2
  INTEGER :: NTipo,i,j,k,m,ii,jj

  SELECT CASE(Tipo)
    CASE('Nada')
      NTipo=0
    CASE('Grad')
      NTipo=1
    CASE('Hess')
      NTipo=2
  END SELECT

  !Se da valor inicial a los resultados
  EnergPot=0.0D0
  IF (Tipo == 'Grad' .OR. Tipo == 'Hess') GradPot=0.0D0
  IF (Tipo == 'Hess') HessPot=0.0D0

  !Se realiza el calculo considerando la media de todas las configuraciones
  SELECT CASE (TRIM(TipoPotencial))
    CASE('lennard-jones')
      DO j=1,SIZE(ConfigDisolv,1)
        m=MOD(j-1,SIZE(Disolvente))+1
        jj=Disolvente(m)%id
        DO i=1,SIZE(Solute)
          ii=Solute(i)%id
          IF (Potencial(ii,jj,1) < 0.1D0) CYCLE
          Dist(1)=Solute(i)%x-ConfigDisolv(j,1)
          Dist(2)=Solute(i)%y-ConfigDisolv(j,2)
          Dist(3)=Solute(i)%z-ConfigDisolv(j,3)
          Dist2=SUM(Dist*Dist)
          Eps=Potencial(ii,jj,2)
          Sig6=(Potencial(ii,jj,3)*Potencial(ii,jj,3)/Dist2)**3
          Sig12=Sig6*Sig6
          EnergPot=EnergPot+Eps*(Sig12-Sig6)
          IF (NTipo == 0) CYCLE
          k=3*i-3
          Aux1=6.0D0*Eps*(Sig6/Dist2-2.0D0*Sig12/Dist2)
          GradPot(k+1)=GradPot(k+1)+Aux1*Dist(1)
          GradPot(k+2)=GradPot(k+2)+Aux1*Dist(2)
        END DO
      END DO
    END CASE
  END SELECT

```



```

GradPot (k+3)=GradPot (k+3)+Aux1*Dist (3)
IF (NTipo == 1) CYCLE
Dist4=Dist2*Dist2
Aux2=24.0D0*Eps*(7.0D0*Sig12/Dist4-2.0D0*Sig6/Dist4)
HessPot (k+1,k+2)=HessPot (k+1,k+2)+Aux2*Dist (1)*Dist (2)
HessPot (k+1,k+3)=HessPot (k+1,k+3)+Aux2*Dist (1)*Dist (3)
HessPot (k+2,k+3)=HessPot (k+2,k+3)+Aux2*Dist (2)*Dist (3)
HessPot (k+1,k+1)=HessPot (k+1,k+1)+Aux2*Dist (1)*Dist (1)+Aux1
HessPot (k+2,k+2)=HessPot (k+2,k+2)+Aux2*Dist (2)*Dist (2)+Aux1
HessPot (k+3,k+3)=HessPot (k+3,k+3)+Aux2*Dist (3)*Dist (3)+Aux1
END DO
END DO
CASE('generic')
DO j=1,SIZE(ConfigDisolv,1)
m=MOD(j-1,SIZE(Disolvente))+1
jj=Disolvente(m)%id
DO i=1,SIZE(Soluto)
ii=Soluto(i)%id
IF (Potencial(ii,jj,1) < 0.1) CYCLE
Dist (1)=Soluto(i)%x-ConfigDisolv(j,1)
Dist (2)=Soluto(i)%y-ConfigDisolv(j,2)
Dist (3)=Soluto(i)%z-ConfigDisolv(j,3)
Dist2=SUM(Dist*Dist)
Dist4=Dist2*Dist2
Dist1=SQRT(Dist2)
Par1=Potencial(ii,jj,2)
Par2=EXP(-Potencial(ii,jj,3)*Dist1)
Par3=Potencial(ii,jj,4)/(Dist4*Dist4*Dist4)
Par4=-Potencial(ii,jj,5)/Dist4
Par5=-Potencial(ii,jj,6)/(Dist4*Dist2)
Par6=-Potencial(ii,jj,7)/(Dist4*Dist4)
EnergPot=EnergPot+Par1*Par2+Par3+Par4+Par5+Par6
IF (NTipo == 0) CYCLE
k=3*i-3
Par1=-Par1*Potencial(ii,jj,3)/Dist1
Par3=-12.0D0*Par3/Dist2
Par4=-4.0D0*Par4/Dist2
Par5=-6.0D0*Par5/Dist2
Par6=-8.0D0*Par6/Dist2
Aux1=Par1*Par2+Par3+Par4+Par5+Par6
GradPot (k+1)=GradPot (k+1)+Aux1*Dist (1)
GradPot (k+2)=GradPot (k+2)+Aux1*Dist (2)
GradPot (k+3)=GradPot (k+3)+Aux1*Dist (3)
IF (NTipo == 1) CYCLE
Par1=Par1*Potencial(ii,jj,3)/Dist2
Par3=-14.0D0*Par3/Dist2
Par4=-6.0D0*Par4/Dist2
Par5=-8.0D0*Par5/Dist2
Par6=-10.0D0*Par6/Dist2
Aux2=Par1*Par2+Par3+Par4+Par5+Par6
HessPot (k+1,k+2)=HessPot (k+1,k+2)+Aux2*Dist (1)*Dist (2)
HessPot (k+1,k+3)=HessPot (k+1,k+3)+Aux2*Dist (1)*Dist (3)
HessPot (k+2,k+3)=HessPot (k+2,k+3)+Aux2*Dist (2)*Dist (3)
HessPot (k+1,k+1)=HessPot (k+1,k+1)+Aux2*Dist (1)*Dist (1)+Aux1
HessPot (k+2,k+2)=HessPot (k+2,k+2)+Aux2*Dist (2)*Dist (2)+Aux1
HessPot (k+3,k+3)=HessPot (k+3,k+3)+Aux2*Dist (3)*Dist (3)+Aux1
END DO
END DO

```

```

END SELECT

EnergPot=EnergPot/NumConfig
IF (Tipo == 'Grad' .OR. Tipo == 'Hess') GradPot(:)=GradPot(:)/NumConfig
!La matriz HessPot ha de ser simetrica
IF (Tipo == 'Hess') THEN
  DO i=1,SIZE(HessPot,1)
    HessPot(i,i)=HessPot(i,i)/NumConfig
    DO j=i+1,SIZE(HessPot,1)
      HessPot(i,j)=HessPot(i,j)/NumConfig
      HessPot(j,i)=HessPot(i,j)
    END DO
  END DO
END IF

END SUBROUTINE CalcularGradHessPot

!-----
! Busca (y encuentra) el minimo de energia con una sola variable por el metodo
! de Brent.
!-----
! Aureo: Numero aureo, (1+SQRT(5))/2
! Preci: Precision relativa a alcanzar en el minimo
! IterMax: Numero maximo de iteraciones para encontrar el minimo
! Valor: Seis puntos con sus valores de energia correspondientes
! Aux: Variable auxiliar
! Med: Posicion media entre dos valores
! D,E,F,P,Q,R: Variables utilizadas por los algoritmos
! Iter: Numero de iteracion
! Fin: Variable para controlar cuando ha de finalizar la busqueda
!-----
SUBROUTINE BuscarMinimo
  USE Parametros, ONLY: Ciclo
  USE Resultados, ONLY: Energia
  USE Optimizacion
  IMPLICIT NONE
  INTERFACE
    DOUBLE PRECISION FUNCTION EnergPunto(Factor)
      DOUBLE PRECISION, INTENT(IN) :: Factor
    END FUNCTION EnergPunto
  END INTERFACE
  DOUBLE PRECISION, PARAMETER :: Aureo=1.618034D0,Preci=1.0D-6
  DOUBLE PRECISION, DIMENSION(6,2) :: Valor
  DOUBLE PRECISION, DIMENSION(2) :: Aux
  DOUBLE PRECISION :: Med,D,E,F,P,Q,R
  INTEGER, PARAMETER :: IterMax=100
  INTEGER :: Iter,Fin

  !Se toman dos valores iniciales para comenzar la busqueda
  Valor(1,1)=0.0D0
  Valor(2,1)=MIN(1.0D0,0.5D0*Paso)
  Valor(1,2)=Energia(Ciclo)+EnergPot
  Valor(2,2)=EnergPunto(Valor(2,1))

  !Se buscan dos puntos (1 y 3) entre los cuales esta el minimo.
  IF (Valor(2,2) > Valor(1,2)) THEN
    Aux(:)=Valor(1,:)
    Valor(1,:)=Valor(2,:)
  
```

```

    Valor(2,:)=Aux(:)
END IF
Valor(3,1)=Valor(2,1)+Aureo*(Valor(2,1)-Valor(1,1))
Valor(3,2)=EnergPunto(Valor(3,1))
DO WHILE (Valor(2,2) >= Valor(3,2))
    !El punto 4 es el minimo de la parabola que pasa por 1, 2 y 3
    R=(Valor(2,1)-Valor(1,1))*(Valor(2,2)-Valor(3,2))
    Q=(Valor(2,1)-Valor(3,1))*(Valor(2,2)-Valor(1,2))
    Valor(4,1)=Valor(2,1) - 0.5D0* &
        ((Valor(2,1)-Valor(1,1))*R-(Valor(2,1)-Valor(3,1))*Q)/(R-Q)
    Valor(4,2)=EnergPunto(Valor(4,1))
    Fin=0
    !Si 4 esta entre 2 y 3
    IF ((Valor(4,1)-Valor(2,1))*(Valor(3,1)-Valor(4,1)) > 0.0D0) THEN
        IF (Valor(4,2) < Valor(3,2)) THEN
            Valor(1,:)=Valor(2,:)
            Valor(2,:)=Valor(4,:)
            Fin=1
        ELSE IF (Valor(4,2) > Valor(2,2)) THEN
            Valor(3,:)=Valor(4,:)
            Fin=1
        ELSE
            Valor(4,1)=Valor(3,1)+Aureo*(Valor(3,1)-Valor(2,1))
            Valor(4,2)=EnergPunto(Valor(4,1))
        END IF
        !Si 4 no esta mas alla de 3
        ELSE IF ((Valor(4,1)-Valor(3,1))*(Valor(3,1)-Valor(1,1)) < 0.0D0) THEN
            Valor(4,1)=Valor(3,1)+Aureo*(Valor(3,1)-Valor(2,1))
            Valor(4,2)=EnergPunto(Valor(4,1))
        END IF
        !Se toman los puntos 2, 3 y 4
        IF (Fin == 0) THEN
            Valor(1,:)=Valor(2,:)
            Valor(2,:)=Valor(3,:)
            Valor(3,:)=Valor(4,:)
        END IF
    END DO

    !Se localiza con mayor precision la posicion del minimo
    IF (Valor(1,1) > Valor(3,1)) THEN
        Aux(:)=Valor(1,:)
        Valor(1,:)=Valor(3,:)
        Valor(3,:)=Aux(:)
    END IF
    IF (Valor(1,2) < Valor(3,2)) THEN
        Valor(5,:)=Valor(3,:)
        Valor(6,:)=Valor(1,:)
    ELSE
        Valor(5,:)=Valor(3,:)
        Valor(6,:)=Valor(1,:)
    END IF
    Iter=0
    Fin=0
    DO WHILE (Fin == 0 .AND. Iter < IterMax)
        Iter=Iter+1
        F=E
        E=D
        Med=0.5D0*(Valor(1,1)+Valor(3,1))

```

```

IF (ABS((Valor(3,1)-Valor(1,1))/Med) < 2.0D0*Preci) THEN
  Fin=1
  EXIT
END IF
!Se calcula la aproximacion parabolica
R=(Valor(2,1)-Valor(6,1))*(Valor(2,2)-Valor(5,2))
Q=(Valor(2,1)-Valor(5,1))*(Valor(2,2)-Valor(6,2))
P=(Valor(2,1)-Valor(6,1))*R-(Valor(2,1)-Valor(5,1))*Q
!Si no es adecuada, se da un paso "aureo"
IF (ABS(D) > 0.5D0*F .OR. &
  D/(Valor(3,1)-Valor(2,1)) >= 1.0D0 .OR. &
  D/(Valor(2,1)-Valor(1,1)) >= 1.0D0) THEN
  IF (Valor(2,1) < Med) THEN
    D=(2.0D0-Aureo)*(Valor(3,1)-Valor(2,1))
  ELSE
    D=(2.0D0-Aureo)*(Valor(1,1)-Valor(2,1))
  END IF
END IF
!Se comprueba la energia y se actua en consecuencia
Valor(4,1)=Valor(2,1)+D
Valor(4,2)=EnergPunto(Valor(4,1))
IF (Valor(4,2) <= Valor(2,2)) THEN
  IF (D < 0.0D0) THEN
    Valor(3,:)=Valor(2,:)
  ELSE
    Valor(1,:)=Valor(2,:)
  END IF
  Valor(5,:)=Valor(6,:)
  Valor(6,:)=Valor(2,:)
  Valor(2,:)=Valor(4,:)
ELSE
  IF (D < 0.0D0) THEN
    Valor(1,:)=Valor(4,:)
  ELSE
    Valor(3,:)=Valor(4,:)
  END IF
  IF (Valor(4,2) <= Valor(6,2)) THEN
    Valor(5,:)=Valor(6,:)
    Valor(6,:)=Valor(4,:)
  ELSE IF (Valor(4,2) <= Valor(5,2)) THEN
    Valor(5,:)=Valor(4,:)
  END IF
END IF
END DO
IF (Fin == 0) WRITE(6,101) 'Se_alcanzo_el_maximo_de_iteraciones_(Brent)'

!Se actualiza el incremento para que coincida con la posicion del minimo
Incremento=Incremento*Valor(2,1)

101  FORMAT(A)

END SUBROUTINE BuscarMinimo

```

```

!-----
! Funcion que devuelve el valor de la energia para cierta geometria
!-----
! Factor: Factor que multiplica al incremento para modificar la geometria
! i:      Contador

```

```

!-----
DOUBLE PRECISION FUNCTION EnergPunto(Factor)
  USE Parametros
  USE Sistema, ONLY: Soluto
  USE Resultados, ONLY: Energia
  USE Optimizacion
  IMPLICIT NONE
  INTERFACE
    SUBROUTINE EjecutarAbInicio(Tipo)
      CHARACTER*4, INTENT(IN) :: Tipo
    END SUBROUTINE EjecutarAbInicio
    SUBROUTINE LeerAbInicio(Tipo)
      CHARACTER*4, INTENT(IN) :: Tipo
    END SUBROUTINE LeerAbInicio
    SUBROUTINE CalcularGradHessPot(Tipo)
      CHARACTER*4, INTENT(IN) :: Tipo
    END SUBROUTINE CalcularGradHessPot
  END INTERFACE
  DOUBLE PRECISION, INTENT(IN) :: Factor
  INTEGER :: i

  !Se calcula la nueva geometria del soluto
  DO i=1,SIZE(Coordenadas),3
    Soluto((i+2)/3) %x=Coordenadas(i)+Factor*Incremento(i)
    Soluto((i+2)/3) %y=Coordenadas(i+1)+Factor*Incremento(i+1)
    Soluto((i+2)/3) %z=Coordenadas(i+2)+Factor*Incremento(i+2)
  END DO

  !Se calcula la energia para esa geometria
  CALL EjecutarAbInicio('Nada')
  CALL LeerAbInicio('Nada')
  CALL CalcularGradHessPot('Nada')
  EnergPunto=Energia(Ciclo)+EnergPot

END FUNCTION

!-----
! Actualiza la matriz Hessiana por el metodo m-BFGS o m-Powell
!-----
! IncGrad: Diferencia entre los gradientes de los ultimos pasos
! Aux:      Vector: Hessiana x Incremento
! vJ,vU:    Vectores auxiliares
! a,b:      Escalares
! i,j:      Contadores
!-----
SUBROUTINE ActualizarHessiana
  USE Parametros, ONLY: EstTransicion
  USE Optimizacion
  IMPLICIT NONE
  DOUBLE PRECISION, DIMENSION(:), ALLOCATABLE :: IncGrad,Aux,vJ,vU
  DOUBLE PRECISION :: a,b
  INTEGER :: i,j

  i=SIZE(Gradiente)
  ALLOCATE(IncGrad(i),Aux(i),vJ(i),vU(i))

  !Metodo m-BFGS, metodo m-Powell para estados de transicion
  IncGrad(:)=Gradiente(:)-GradAnt(:)

```

```

Aux=MATMUL(Hessiana(:, :), Incremento(:))
vJ(:)=IncGrad(:)-Aux(:)
IF (EstTransicion == 'SI') THEN
  Aux(:)=Incremento(:)
ELSE
  a=SUM(Aux(:)*IncGrad(:))*SUM(Aux(:)*IncGrad(:))/ &
    (SUM(Aux(:)*Aux(:))*SUM(IncGrad(:)*IncGrad(:)))
  b=1.0D0-a
  Aux(:)=a*IncGrad(:)+b*Aux(:)
END IF
vU(:)=Aux(:)/SUM(Incremento(:)*Aux(:))
a=SUM(vJ(:)*Incremento(:))
DO i=1, SIZE(Hessiana,1)
  Hessiana(i,i)=Hessiana(i,i)+2.0D0*vJ(i)*vU(i)-a*vU(i)*vU(i)
  DO j=i+1, SIZE(Hessiana,1)
    Hessiana(i,j)=Hessiana(i,j)+vJ(i)*vU(j)+vU(i)*vJ(j)-a*vU(i)*vU(j)
    Hessiana(j,i)=Hessiana(i,j)
  END DO
END DO

DEALLOCATE (IncGrad,Aux,vJ,vU)

END SUBROUTINE ActualizarHessiana

!-----
! Calcula el desplazamiento en las coordenadas para el proximo paso por el
! metodo RFO (Jon Baker, J. Comp. Chem. 7 (1968) 385-395)
!-----
! Num:      Dimension de las matrices y vectores: 3*N+1
! Matriz:   Matriz a diagonalizar
! Valores:  Vector que contiene los valores propios
! Vectores: Matriz que contiene los vectores propios
! F:        Proyeccion del gradiente sobre los vectores propios
! Masa:     Vector con las masas de los atomos
! Aux,Aux2: Vectores auxiliares
! Lambda:   Desplazamiento de los valores propios
! Tau:      Desplazamiento del valor propio para estado de transicion
! Suma:     Variable para el calculo de lambda
! Modo:     El modo en el que se maximiza la energia
! Preci:    Precision para no dividir por cero
! Error:    Variable para controlar los errores
! i,j,k:    Contadores
!-----
SUBROUTINE CalcularIncremento
  USE Parametros
  USE Optimizacion, ONLY: Gradiente, Hessiana, Incremento
  USE Sistema, ONLY: Soluto
  IMPLICIT NONE
  INTERFACE
    SUBROUTINE Diagonalizar(Mat,Vec,Val,Ord,Error)
      DOUBLE PRECISION, DIMENSION(:, :), INTENT(INOUT) :: Mat
      DOUBLE PRECISION, DIMENSION(:, :), INTENT(OUT) :: Vec
      DOUBLE PRECISION, DIMENSION(:), INTENT(OUT) :: Val
      CHARACTER*2, INTENT(IN) :: Ord
      INTEGER, INTENT(OUT) :: Error
    END SUBROUTINE Diagonalizar
  END INTERFACE
  DOUBLE PRECISION, DIMENSION(:), ALLOCATABLE :: Valores,F,Masa

```

```

DOUBLE PRECISION, DIMENSION(:, :), ALLOCATABLE :: Matriz, Vectores
DOUBLE PRECISION, DIMENSION(3) :: Aux, Aux2
DOUBLE PRECISION :: Lambda, Tau, Suma
DOUBLE PRECISION, PARAMETER :: Preci=1.0D-8
INTEGER :: Num, Modo, i, j, k, Error

Num=SIZE(Gradiente)
ALLOCATE (Valores (Num), Matriz (Num, Num), Vectores (Num, Num), F (Num), Masa (Num))

!Convierte hessiana y gradiente a coordenadas ponderadas
DO i=1, Num
  j=(i-1)/3+1
  Masa(i)=SQRT(Soluto(j) %Masa)
  IF (Masa(i) > Preci) Masa(i)=1.0D0/Masa(i)
END DO
F(:)=Gradiente(:)*Masa(:)
DO i=1, Num
  Matriz(i, :)=Hessiana(i, :)*Masa(i)*Masa(:)
END DO

CALL Diagonalizar(Matriz, Vectores, Valores, 'SI', Error)
IF (Error /= 0) THEN
  WRITE(21, 666) 'No se ha podido diagonalizar la matriz hessiana.'
  STOP
END IF

!Se calcula la proyeccion del gradiente sobre los vectores propios
F(:)=MATMUL(TRANSPOSE(Vectores(:, :)), F(:))

!Se maximiza el modo "Modo", se minimizan los demas, j es el menor
Modo=0
j=1
IF (EstTransicion == 'SI') Modo=1
IF (Modo == 1) j=2

!Se calcula Lambda: Sum(F_i^2/(lambda-b_i))=lambda)
Lambda=MIN(0.0D0, Valores(j))
DO k=1, 1000
  Tau=Lambda
  Suma=0.0D0
  DO i=j+1, Num
    IF ((i /= Modo) .AND. (ABS(Lambda-Valores(i)) > Preci)) &
      Suma=Suma+F(i)*F(i)/(Lambda-Valores(i))
  END DO
  Lambda=0.5D0*(Valores(j)+Suma)
  Lambda=Lambda-SQRT(Lambda*Lambda+F(j)*F(j)-Valores(j)*Suma)
  Lambda=(Tau+Lambda+Lambda)/3.0D0
  IF (ABS(Tau-Lambda) < Preci) EXIT
END DO

!Se calcula Tau si se va a optimizar un estado de transicion
IF (Modo /= 0) THEN
  Tau=0.5D0*Valores(Modo)
  Tau=Tau+SQRT(Tau*Tau+F(Modo)*F(Modo))
END IF

!Se calcula el paso por el metodo P-RFO
Incremento(:)=0.0D0
DO i=1, Num

```

```

        IF (i == Modo) THEN
            IF (ABS(Tau-Valores(i)) > Preci) &
                Incremento(:)=Incremento(:)+Vectores(:,i)*F(i)/(Tau-Valores(i))
            ELSE
                IF (ABS(Lambda-Valores(i)) > Preci) &
                    Incremento(:)=Incremento(:)+Vectores(:,i)*F(i)/(Lambda-Valores(i))
                END IF
            END DO

        !Si el paso es demasiado grande, se toma el Newton-Raphson
        Tau=0.0D0
        DO i=1,Num
            IF (Valores(i) > Preci) Tau=Tau+F(i)*F(i)/(Valores(i)*Valores(i))
        END DO
        IF (Tau/SUM(Incremento*Incremento) < Preci) THEN
            Incremento(:)=0.0D0
            DO i=1,Num
                IF (ABS(Valores(i)) > Preci) THEN
                    IF (i == Modo) THEN
                        Incremento(:)=Incremento(:)+Vectores(:,i)*F(i)/ABS(Valores(i))
                    ELSE
                        Incremento(:)=Incremento(:)+Vectores(:,i)*F(i)/-ABS(Valores(i))
                    END IF
                END IF
            END DO
        END IF

        !Convierte a coordenadas cartesianas (no ponderadas)
        Incremento(:)=Incremento(:)*Masa(:)

        DEALLOCATE(Matriz,Vectores,Valores,F,Masa)

666  FORMAT (80('-')/A/'El_programa_se_detiene'/80('-'))

END SUBROUTINE CalcularIncremento

!-----
! Escribe en el fichero de salida la geometria y los gradientes actuales
!-----
! Iter: Numero de iteracion actual en la optimizacion
! Num:  Numero de coordenadas
! i:    Contador
!-----
SUBROUTINE EscribeOptim(Iter)
    USE Parametros, ONLY: Ciclo
    USE Sistema, ONLY: Soluto
    USE Optimizacion, ONLY: Coordenadas,Gradiente,EnergPot
    USE Resultados, ONLY: Energia,EPol
    IMPLICIT NONE
    INTEGER, INTENT(IN) :: Iter
    INTEGER :: Num,i

    Num=SIZE(Coordenadas)
    IF (Iter == 0) THEN
        WRITE(22,667) ' INICIO'
        WRITE(22,303) EPol(Ciclo)-Energia(0)
    ELSE
        WRITE(22,668) ' ITERACION_',Iter
    END IF

```



```
END IF
IF (EnergPot < 0.0D0) THEN
  WRITE(22,302) Energia(Ciclo),ABS(EnergPot),Energia(Ciclo)+EnergPot
ELSE
  WRITE(22,301) Energia(Ciclo),ABS(EnergPot),Energia(Ciclo)+EnergPot
END IF
WRITE(22,101) '      Geometria_(u.a.):'
DO i=1,Num,3
  WRITE(22,103) (i+2)/3,Soluto((i+2)/3) %Nombre,Coordenadas(i:i+2)
END DO
WRITE(22,101) '      Gradientes_(u.a.):'
DO i=1,Num,3
  WRITE(22,103) (i+2)/3,Soluto((i+2)/3) %Nombre,Gradiente(i:i+2)
END DO
WRITE(22,304) SQRT(SUM(Gradiente(:)*Gradiente(:)))
WRITE(22,102)

101  FORMAT (A)
102  FORMAT (80('='))
103  FORMAT (' Atomo_',I2,2X,A,2X,3(F20.12))
301  FORMAT (' Energia_(u.a.):_',F20.12,'+_+',F15.12,'_=_',F20.12)
302  FORMAT (' Energia_(u.a.):_',F20.12,'_-'',F15.12,'_=_',F20.12)
303  FORMAT (' Energia_de_polarizacion_inicial_(u.a.):_',F20.12)
304  FORMAT (' Modulo_del_gradiente_(u.a.):_',F20.12)
667  FORMAT (20X,A/80('-'))
668  FORMAT (20X,A,I2/80('-'))

END SUBROUTINE EscribeOptim
```



---

## BIBLIOGRAFÍA

- [1] K. Morokuma, L. Pedersen, «Molecular-orbital studies of hydrogen bonds. An ab initio calculation for dimeric H<sub>2</sub>O». *J. Chem. Phys.* **48** (1968) 3275-3282  
[Citado en pág. 25]
- [2] G. Alagona, A. Pullman, E. Scrocco, J. Tomasi, «Quantum mechanical studies of environmental effects in biomolecules. I. Hydration of formamide». *Int. J. Pept. Protein Res.* **5** (1973) 251-259  
[Citado en pág. 25]
- [3] A. Pullman, B. Pullman, «New paths in the molecular orbital approach to solvation of biological molecules». *Q. Rev. Biophys.* **7** (1975) 505-566  
[Citado en pág. 25]
- [4] R. Car, M. Parrinello, «Unified approach for molecular dynamics and density functional theory». *Phys. Rev. Lett.* **55** (1985) 2471-2474 [Citado en pág. 26]
- [5] J. A. Barker, D. Henderson, «What is "liquid"? Understanding the states of matter». *Rev. Mod. Phys.* **48** (1976) 587-671 [Citado en pág. 26]
- [6] J.-P. Hansen, I. R. McDonald, *Theory of Simple Liquids*, 2.<sup>a</sup> ed., Academic Press (1990) [Citado en págs. 26 y 56]
- [7] J. K. Percus, G. J. Yevick, «Analysis of classical statistical mechanics by means of collective coordinates». *Phys. Rev.* **110** (1958) 1-13 [Citado en pág. 26]
- [8] D. Chandler, H. C. Andersen, «Optimized cluster expansions for classical fluids. II. Theory of molecular liquids». *J. Chem. Phys.* **57** (1972) 1930-1937  
[Citado en págs. 26 y 55]
- [9] D. Chandler, «Derivation of an integral equation for pair correlation functions in molecular fluids». *J. Chem. Phys.* **59** (1973) 2742-2746  
[Citado en pág. 26]

- [10] J. L. Lebowitz, J. K. Percus, «Mean spherical model for lattice gases with extended hard cores and continuum fluids». *Phys. Rev.* **144** (1966) 251-258  
[Citado en pág. 26]
- [11] T. Morita, «Theory of classical fluids – hyper-netted chain approximation. 1. Formulation for a one-component system». *Prog. Theor. Phys.* **20** (1958) 920-938  
[Citado en pág. 26]
- [12] J. M. J. van Leeuwen, J. Groeneveld, J. de Boer, «New method for the calculation of the pair correlation function. I». *Physica* **25** (1959) 792-808  
[Citado en pág. 26]
- [13] G. S. Rushbrooke, «On the hyper-chain approximation in the theory of classical fluids». *Physica* **26** (1960) 259-265  
[Citado en pág. 26]
- [14] L. Verlet, «On the theory of classical fluids». *Nuovo Cimento Ser. X* **18** (1960) 77-101  
[Citado en pág. 26]
- [15] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press (1987)  
[Citado en págs. 26, 60 y 114]
- [16] N. C. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, «Equation of state calculations by fast computing machines». *J. Chem. Phys.* **21** (1953) 1087-1092  
[Citado en págs. 27 y 62]
- [17] B. J. Alder, T. E. Wainwright, «Phase transition for a hard sphere system». *J. Chem. Phys.* **27** (1957) 1208-1209  
[Citado en págs. 27 y 63]
- [18] A. Rahman, F. H. Stillinger, «Molecular dynamics study of liquid water». *J. Chem. Phys.* **55** (1971) 3336-3359  
[Citado en págs. 27 y 63]
- [19] H. S. Frank, W. Y. Wen, «Structural aspects of ion-solvent interaction in aqueous solutions: A suggested picture of water structure». *Discuss. Faraday Soc.* **24** (1957) 133-140  
[Citado en pág. 27]
- [20] E. Clementi, W. Kołos, G. C. Lie, G. Ranghino, «Non-additivity of interaction in water trimers». *Int. J. Quantum Chem.* **17** (1980) 377-398 [Citado en pág. 27]
- [21] J. E. H. Köhler, W. Sängler, B. Lesyng, «Cooperative effects in extended hydrogen bonded systems involving O–H groups. Ab initio studies of the cyclic S<sub>4</sub> water tetramer». *J. Comput. Chem.* **8** (1987) 1090-1098  
[Citado en pág. 27]
- [22] G. Chałasiński, S. M. Cybulski, M. M. Szcześniak, S. Sheiner, «Nonadditive effects in HF and HCl trimers». *J. Chem. Phys.* **91** (1989) 7048-7056  
[Citado en pág. 27]

- [23] G. Chałasiński, M. M. Szczeniak, P. Cieplak, S. Sheiner, «Ab initio study of intermolecular potential of H<sub>2</sub>O trimer». *J. Chem. Phys.* **94** (1991) 2873-2883 [Citado en pág. 27]
- [24] M. M. Szczeniak, R. A. Kendall, G. Chałasiński, «Ab initio study of the nonadditive effects in the trimer of ammonia». *J. Chem. Phys.* **95** (1991) 5169-5178 [Citado en pág. 27]
- [25] A. Wallqvist, B. J. Berne, «Effective potentials for liquid water using polarizable and nonpolarizable models». *J. Phys. Chem.* **97** (1993) 13841-13851 [Citado en pág. 27]
- [26] S. W. Rick, S. J. Stuart, «Dynamical fluctuating charge force fields: Application to liquid water». *J. Chem. Phys.* **101** (1994) 6141-6156 [Citado en pág. 27]
- [27] R. Kelterbaum, N. Turki, A. Rahmouni, E. Kochanski, «About the nature of intermolecular three-body forces in ionic systems: The case of protonated hydrates». *J. Chem. Phys.* **100** (1994) 1589-1598 [Citado en pág. 27]
- [28] P. A. Bash, M. J. Field, M. Karplus, «Free energy perturbation method for chemical reactions in the condensed phase: A dynamic approach based on a combined quantum and molecular mechanics potential». *J. Am. Chem. Soc.* **109** (1987) 8092-8094 [Citado en pág. 27]
- [29] J. Gao, X. Xia, «A priori evaluation of aqueous polarisation effects through Monte Carlo QM/MM simulations». *Science* **258** (1992) 631-635 [Citado en pág. 27]
- [30] J. Gao, «Absolute free energy of solvation from Monte Carlo simulations using combined quantum and molecular mechanical potentials». *J. Phys. Chem.* **96** (1992) 537-540 [Citado en pág. 27]
- [31] I. Tuñón, E. Silla, J. Bertrán, «Proton solvation in liquid water: An ab initio study using the continuum model». *J. Phys. Chem.* **97** (1993) 5547-5552 [Citado en pág. 27]
- [32] J. T. Blair, K. Krogh-Jespersen, R. M. Levy, «Solvent effects on optical absorption spectra: The  $^1A_1 \rightarrow ^1A_2$  transition of formaldehyde in water». *J. Am. Chem. Soc.* **111** (1989) 6948-6956 [Citado en pág. 27]
- [33] R. M. Levy, D. B. Kitchen, J. T. Blair, K. Krogh-Jespersen, «Molecular dynamics simulation of time-resolved fluorescence and nonequilibrium solvation of formaldehyde in water». *J. Phys. Chem.* **94** (1990) 4470-4476 [Citado en pág. 27]

- [34] D. Rinaldi, J.-L. Rivail, «Molecular polarizabilities and dielectric effect of the medium in the liquid state. Theoretical study of the water molecule and its dimers». *Theor. Chim. Acta* **32** (1973) 57-70 [Citado en págs. 28 y 54]
- [35] O. Tapia, O. Goscinski, «Self-consistent reaction field theory of solvent effects». *Mol. Phys.* **29** (1975) 1653-1661 [Citado en págs. 28 y 54]
- [36] S. Miertuš, E. Scrocco, J. Tomasi, «Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects». *Chem. Phys.* **55** (1981) 117-129 [Citado en págs. 28 y 54]
- [37] P. Hobza, R. Zahradník, *Intermolecular Complexes. The Role of Van der Waals Systems in Physical Chemistry and in the Biodisciplines*, núm. 52 en *Studies in Physical and Theoretical Chemistry*, Elsevier (1988) [Citado en pág. 28]
- [38] S. Ten-no, F. Hirata, S. Kato, «A hybrid approach for the solvent effect on the electronic structure of a solute based on the RISM and Hartree-Fock equations». *Chem. Phys. Lett.* **214** (1993) 391-396 [Citado en págs. 28 y 56]
- [39] S. Ten-no, F. Hirata, S. Kato, «Reference interaction site model self-consistent field study for solvation effect on carbonyl compounds in aqueous solution». *J. Chem. Phys.* **100** (1994) 7443-4753 [Citado en págs. 28 y 56]
- [40] M. L. Sánchez, M. Á. Aguilar, F. J. Olivares del Valle, «Study of solvent effects by means of averaged solvent electrostatic potentials obtained from molecular dynamics data». *J. Comput. Chem.* **18** (1997) 313-322 [Citado en págs. 28, 73 y 77]
- [41] M. L. Sánchez, M. Á. Aguilar, F. J. Olivares del Valle, «A mean field approach that combines quantum mechanics and molecular dynamics simulation: The water molecule in liquid water». *J. Mol. Struct. (Theochem)* **426** (1998) 181-190 [Citado en págs. 28, 73 y 118]
- [42] M. L. Sánchez, M. E. Martín, M. Á. Aguilar, F. J. Olivares del Valle, «Solvent effects by means of averaged solvent electrostatic potentials: Coupled method». *J. Comput. Chem.* **21** (2000) 705-715 [Citado en págs. 28, 73 y 80]
- [43] M. Born, J. R. Oppenheimer, «Zur Quantentheorie der Molekeln». *Ann. Phys. (Leipzig)* **84** (1927) 457-484 [Citado en pág. 34]
- [44] D. R. Hartree, «The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods». *Proc. Cambridge Phil. Soc.* **24** (1928) 89-110 [Citado en págs. 36 y 113]

- [45] D. R. Hartree, «The wave mechanics of an atom with a non-Coulomb central field. Part II. Some results and discussion». *Proc. Cambridge Phil. Soc.* **24** (1928) 111-132 [Citado en págs. 36 y 113]
- [46] V. A. Fock, «Näherungsmethode zur Lösung der quantenmechanischen Mehrkörperproblems». *Z. Physik* **61** (1930) 126-148 [Citado en págs. 36 y 113]
- [47] J. C. Slater, «The theory of complex spectra». *Phys. Rev.* **34** (1929) 1293-1322 [Citado en pág. 36]
- [48] J. Almlöf, R. Ahlrichs, «Notes on Hartree-Fock theory and related topics», en *European Summerschool in Quantum Chemistry 2003* (B. O. Roos, P.-O. Widmark, eds.), vol. I, 3.<sup>a</sup> ed., Lund University, Lund (2003) págs. 169-240 [Citado en pág. 37]
- [49] I. N. Levine, *Química Cuántica*, 1.<sup>a</sup> ed., Editorial AC, Madrid (1977) [Citado en pág. 37]
- [50] C. C. J. Roothaan, «New developments in molecular orbital theory». *Rev. Mod. Phys.* **23** (1951) 69-89 [Citado en pág. 39]
- [51] G. G. Hall, «The molecular orbital theory of chemical valency. VIII - A method of calculating ionization potentials». *Proc. R. Soc. Lond. Ser. A* **205** (1951) 541-552 [Citado en pág. 39]
- [52] C. Møller, M. S. Plesset, «Note on an approximation treatment for many-electron systems». *Phys. Rev.* **46** (1934) 618-622 [Citado en págs. 41 y 113]
- [53] S. F. Boys, «Electronic wavefunctions. I. A general method of calculation for stationary states of any molecular system». *Proc. R. Soc. Lond. Ser. A* **200** (1950) 542-554 [Citado en pág. 42]
- [54] J. A. Pople, J. S. Binkley, R. Seeger, «Theoretical models incorporating electron correlation». *Int. J. Quantum Chem. Symp.* **10** (1976) 1-19 [Citado en pág. 42]
- [55] L. G. Yaffe, W. A. Goddard III, «Orbital optimization in electronic wave functions; equations for quadratic and cubic convergence of general multi-configuration wave functions». *Phys. Rev. A* **13** (1976) 1682-1691 [Citado en pág. 43]
- [56] B. O. Roos, P. R. Taylor, P. E. M. Siegbahn, «A complete active space SCF (CASSCF) using a density matrix formulated super-CI approach». *Chem. Phys.* **48** (1980) 157-173 [Citado en págs. 43 y 113]
- [57] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski, «Second-order perturbation theory with a CASSCF reference function». *J. Phys. Chem.* **94** (1990) 5483-5488 [Citado en pág. 44]

- [58] K. Andersson, P.-Å. Malmqvist, B. O. Roos, «Second-order perturbation theory with a complete active space self-consistent field reference function». *J. Chem. Phys.* **96** (1992) 1218-1226 [Citado en pág. 44]
- [59] J. Čížek, «On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods». *J. Chem. Phys.* **45** (1966) 4256-4266 [Citado en pág. 44]
- [60] J. Čížek, «On the use of the cluster expansion and the technique of diagrams in calculations of correlation effects in atoms and molecules», en *Correlation Effects in Atoms and Molecules* (R. Lefebvre, C. Moser, eds.), núm. 14 en *Advances in Chemical Physics*, Interscience Publishers (1969) págs. 35-89 [Citado en pág. 44]
- [61] J. E. Campbell, «On a law of combination of operators bearing on the theory of continuous transformation groups». *Proc. London Math. Soc.* **28** (1897) 381-390 [Citado en pág. 45]
- [62] H. F. Baker, «Alternants and continuous groups». *Proc. London Math. Soc.* **3** (1905) 24-47 [Citado en pág. 45]
- [63] F. Hausdorff, «Die symbolische Exponentialformel in der Gruppentheorie». *Ber. Verh. Sächs. Akad. Wiss. Leipzig, Math.-Naturwiss. Klasse* **58** (1906) 19-48 [Citado en pág. 45]
- [64] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, núm. 16 en *International Series of Monographs on Chemistry*, Oxford University Press (1989) [Citado en págs. 45 y 113]
- [65] P. Hohenberg, W. Kohn, «Inhomogeneous electron gas». *Phys. Rev.* **136** (1964) B864-B871 [Citado en pág. 46]
- [66] W. Kohn, L. J. Sham, «Self-consistent equations including exchange and correlation effects». *Phys. Rev.* **140** (1965) A1133-A1138 [Citado en pág. 46]
- [67] J. C. Slater, «A simplification of the Hartree-Fock method». *Phys. Rev.* **81** (1951) 385-390 [Citado en pág. 48]
- [68] P. A. M. Dirac, «Note on exchange phenomena in the Thomas atom». *Proc. Cambridge Phil. Soc.* **26** (1930) 376-385 [Citado en pág. 48]
- [69] D. M. Ceperley, B. J. Alder, «Ground state of the electron gas by a stochastic method». *Phys. Rev. Lett.* **45** (1980) 566-569 [Citado en pág. 48]



- [70] S. H. Vosko, L. Wilk, M. Nusair, «Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis». *Can. J. Phys.* **58** (1980) 1200-1211 [Citado en pág. 48]
- [71] A. D. Becke, «Density-functional exchange-energy approximation with correct asymptotic behavior». *Phys. Rev. A* **38** (1988) 3098-3100 [Citado en págs. 49 y 113]
- [72] C. Lee, W. Yang, R. G. Parr, «Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density». *Phys. Rev. B* **37** (1988) 785-789 [Citado en págs. 49 y 113]
- [73] A. D. Becke, «Density-functional thermochemistry. III. The role of exact exchange». *J. Chem. Phys.* **98** (1993) 5648-5652 [Citado en págs. 49 y 113]
- [74] J. Tomasi, M. Persico, «Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent». *Chem. Rev.* **94** (1994) 2027-2094 [Citado en pág. 51]
- [75] M. Born, «Volumen und Hydratationswärme der Ionen». *Z. Physik* **1** (1920) 45-48 [Citado en pág. 52]
- [76] A. A. Rashin, B. Honig, «Reevaluation of the Born model of ion hydration». *J. Phys. Chem.* **89** (1985) 5588-5593 [Citado en pág. 52]
- [77] R. P. Bell, «The electrostatic energy of dipole molecules in different media». *Trans. Faraday Soc.* **27** (1931) 797-802 [Citado en pág. 52]
- [78] L. Onsager, «Electric moments of molecules in liquids». *J. Am. Chem. Soc.* **58** (1936) 1486-1493 [Citado en pág. 52]
- [79] C. J. F. Böttcher. *Physica* **5** (1938) 635 [Citado en pág. 52]
- [80] J. G. Kirkwood, «Theory of solutions of molecules containing widely separated charges with special application to zwitterions». *J. Chem. Phys.* **2** (1934) 351-361 [Citado en pág. 53]
- [81] J. G. Kirkwood, F. H. Westheimer, «The electrostatic influence of substituents on the dissociation constants of organic acids. I». *J. Chem. Phys.* **6** (1938) 506-512 [Citado en pág. 53]
- [82] J. G. Kirkwood, F. H. Westheimer, «The electrostatic influence of substituents on the dissociation constants of organic acids. II». *J. Chem. Phys.* **6** (1938) 513-517 [Citado en pág. 53]

- [83] J.-L. Rivail, D. Rinaldi, «A quantum chemical approach to dielectric solvent effects in molecular liquids». *Chem. Phys.* **18** (1976) 233-242  
[Citado en págs. 54 y 164]
- [84] D. Rinaldi, M. F. Ruiz-López, J.-L. Rivail, «Ab initio SCF calculations on electrostatically solvated molecules using a deformable three axes ellipsoidal cavity». *J. Chem. Phys.* **78** (1983) 834-838  
[Citado en págs. 54 y 164]
- [85] V. Dillet, D. Rinaldi, J.-L. Rivail, «Liquid-state quantum chemistry: An improved cavity model». *J. Phys. Chem.* **98** (1994) 5034-5039  
[Citado en págs. 54 y 164]
- [86] C. J. Cramer, D. G. Truhlar, «An SCF solvation model for the hydrophobic effect and absolute free energies of aqueous solvation». *Science* **256** (1992) 213-217  
[Citado en pág. 54]
- [87] S. Miertuš, J. Tomasi, «Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes». *Chem. Phys.* **65** (1982) 239-245  
[Citado en pág. 54]
- [88] F. J. Olivares del Valle, J. Tomasi, «Electron correlation and solvation effects. I. Basic formulation and preliminary attempt to include the electron correlation in the quantum mechanical polarizable continuum model so as to study solvation phenomena». *Chem. Phys.* **150** (1991) 139-150  
[Citado en pág. 54]
- [89] B. Mennucci, J. Tomasi, «Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries». *J. Chem. Phys.* **106** (1996) 5151-5158  
[Citado en pág. 54]
- [90] A. Klamt, G. Schüürmann, «COSMO: A new approach to dielectric screening in solvent with explicit expressions for the screening energy and its gradient». *J. Chem. Soc. Perkin Trans. 2* (1993) 799-805  
[Citado en pág. 55]
- [91] T. N. Truong, E. V. Stefanovich, «A new method for incorporating solvent effect into the classical, ab initio molecular orbital and density functional theory frameworks for arbitrary shape cavity». *Chem. Phys. Lett.* **240** (1995) 253-260  
[Citado en pág. 55]
- [92] E. Cancès, B. Mennucci, «New applications of integral equation methods for solvation continuum models: Ionic solutions and liquid crystals». *J. Math. Chem.* **23** (1993) 309-326  
[Citado en pág. 55]
- [93] L. S. Ornstein, F. Zernike, «Accidental deviations of density and opalescence at the critical point of a single substance». *Proc. Acad. Sci. (Amsterdam)* **17** (1914) 793-806  
[Citado en pág. 56]

- [94] F. Hirata (ed.), *Molecular Theory of Solvation*, núm. 24 en *Understanding Chemical Reactivity*, Kluwer Academic Publishers (2004) [Citado en pág. 56]
- [95] A. Warshel, M. Levitt, «Theoretical studies of enzymatic reactions: Dielectric electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme». *J. Mol. Biol.* **103** (1976) 227-249 [Citado en págs. 56 y 67]
- [96] A. Warshel, «Calculations of chemical processes in solutions». *J. Phys. Chem.* **83** (1979) 1640-1652 [Citado en pág. 56]
- [97] A. Warshel, S. T. Russell, «Calculations of electrostatic interactions in biological systems and in solutions». *Q. Rev. Biophys.* **17** (1984) 283-422 [Citado en pág. 56]
- [98] A. Warshel, R. M. Weiss, «An empirical valence bond approach for comparing reactions in solutions and in enzymes». *J. Am. Chem. Soc.* **102** (1980) 6218-6226 [Citado en págs. 57 y 69]
- [99] A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, John Wiley & Sons (1991) [Citado en págs. 57 y 67]
- [100] L. C. Pauling, «The nature of the chemical bond. IV. The energy of single bonds and the relative electronegativity of atoms». *J. Am. Chem. Soc.* **54** (1932) 3570-3582 [Citado en pág. 57]
- [101] J. Åqvist, M. Fothergill, A. Warshel, «Computer simulation of the carbon dioxide/bicarbonate interconversion step in human carbonic anhydrase I». *J. Am. Chem. Soc.* **115** (1993) 631-635 [Citado en pág. 57]
- [102] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, «A second generation force field for the simulation of proteins, nucleic acids, and organic molecules». *J. Am. Chem. Soc.* **117** (1995) 5179-5197 [Citado en pág. 59]
- [103] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, «CHARMM: A program for macromolecular energy, minimization, and dynamics calculations». *J. Comput. Chem.* **4** (1983) 187-217 [Citado en pág. 59]
- [104] W. L. Jorgensen, J. D. Madura, C. J. Swenson, «Optimized intermolecular potential functions for liquid hydrocarbons». *J. Am. Chem. Soc.* **106** (1984) 6638-6646 [Citado en págs. 59 y 116]
- [105] W. L. Jorgensen, «Optimized intermolecular potential functions for liquid alcohols». *J. Phys. Chem.* **90** (1986) 1276-1284 [Citado en págs. 59, 115 y 116]

- [106] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, «Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids». *J. Am. Chem. Soc.* **118** (1996) 11225-11236  
[Citado en págs. 59 y 116]
- [107] N. L. Allinger, «Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms». *J. Am. Chem. Soc.* **99** (1977) 8127-8134  
[Citado en pág. 59]
- [108] T. A. Halgren, «The representation of Van der Waals (vdW) interactions in molecular mechanics force fields: Potential form, combination rules, and vdW parameters». *J. Am. Chem. Soc.* **114** (1992) 7827-7843 [Citado en pág. 59]
- [109] P. P. Ewald, «Die Berechnung optischer und elektrostatischer Gitterpotentiale.» *Ann. Phys. (Leipzig)* **64** (1921) 253-287 [Citado en págs. 60 y 114]
- [110] D. Frenkel, B. Smit, *Understanding Molecular Simulations. From Algorithms to Applications*, 2.<sup>a</sup> ed., núm. 1 en Computational Science Series, Academic Press (2002) [Citado en págs. 60, 92 y 114]
- [111] N. C. Metropolis, S. Ulam, «The Monte Carlo method». *J. Am. Stat. Assoc.* **44** (1949) 335-341 [Citado en pág. 61]
- [112] K. A. Fichtorn, W. H. Weinberg, «Theoretical foundations of dynamical Monte Carlo simulations». *J. Chem. Phys.* **95** (1991) 1090-1096  
[Citado en pág. 63]
- [113] B. J. Alder, T. E. Wainwright, «Studies in molecular dynamics. I. General method». *J. Chem. Phys.* **31** (1959) 459-466 [Citado en pág. 63]
- [114] A. Rahman, «Correlations in the motion of atoms in liquid argon». *Phys. Rev.* **136** (1964) A405-A411 [Citado en pág. 63]
- [115] L. Verlet, «Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules». *Phys. Rev.* **159** (1967) 98-103  
[Citado en pág. 64]
- [116] R. W. Hockney, J. W. Eastwood, *Computer Simulation Using Particles*, McGraw-Hill (1981) [Citado en pág. 64]
- [117] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, «A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters». *J. Chem. Phys.* **76** (1982) 637-649 [Citado en pág. 64]

- [118] S. Nosé, «A molecular dynamics method for simulations in the canonical ensemble». *Mol. Phys.* **52** (1984) 255-268 [Citado en págs. 65 y 114]
- [119] W. G. Hoover, «Canonical dynamics: Equilibrium phase-space distributions». *Phys. Rev. A* **31** (1985) 1695-1697 [Citado en págs. 65 y 114]
- [120] D. J. Evans, S. Murad, «Singularity free algorithm for molecular dynamics simulation of rigid polyatomics». *Mol. Phys.* **34** (1977) 327-331 [Citado en pág. 66]
- [121] G. S. Pawley, M. T. Dove, «Quaternion-based reorientation conditions for molecular dynamics analyses». *Mol. Phys.* **55** (1985) 1147-1157 [Citado en pág. 66]
- [122] K. Refson, G. S. Pawley, «Molecular dynamics studies of the condensed phases of n-butane and their transitions. I. Techniques and model results». *Mol. Phys.* **61** (1987) 669-692 [Citado en pág. 66]
- [123] V. Luzhkov, A. Warshel, «Microscopic models for quantum mechanical calculations of chemical processes in solutions: LD/AMPAC and SCAAS/AMPAC calculations of solvation energies». *J. Comput. Chem.* **13** (1992) 199-213 [Citado en pág. 67]
- [124] J. Åqvist, A. Warshel, «Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches». *Chem. Rev.* **93** (1993) 2523-2544 [Citado en pág. 67]
- [125] M. J. Field, P. A. Bash, M. Karplus, «A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations». *J. Comput. Chem.* **11** (1990) 700-733 [Citado en pág. 67]
- [126] G. Monard, X. Prat-Resina, A. González-Lafont, J. M. Lluch, «Determination of enzymatic reaction pathways using QM/MM methods». *Int. J. Quantum Chem.* **93** (2003) 229-244 [Citado en pág. 68]
- [127] R. P. Muller, A. Warshel, «Ab initio calculations of free energy barriers for chemical reactions in solution». *J. Phys. Chem.* **99** (1995) 17516-17524 [Citado en pág. 69]
- [128] D. Wei, D. R. Salahub, «A combined density functional and molecular dynamics simulation of a quantum water molecule in aqueous solution». *Chem. Phys. Lett.* **224** (1994) 291-296 [Citado en pág. 69]
- [129] T. A. Wesolowski, A. Warshel, «Ab initio free energy perturbation calculations of solvation free energy using the frozen density functional approach». *J. Phys. Chem.* **98** (1994) 5183-5187 [Citado en pág. 69]

- [130] R. V. Stanton, D. S. Hartsough, K. M. Merz Jr, «An examination of a density functional/molecular mechanical coupled potential». *J. Comput. Chem.* **16** (1995) 113-128 [Citado en pág. 69]
- [131] I. Tuñón, M. T. C. Martins-Costa, C. Millot, M. F. Ruiz-López, J.-L. Rivail, «A coupled density functional–molecular mechanics Monte Carlo simulation method: The water molecule in liquid water». *J. Comput. Chem.* **17** (1996) 19-29 [Citado en págs. 69 y 163]
- [132] J. Gao, «An automated procedure for simulating chemical reactions in solution. Application to the decarboxylation of 3-carboxybenzoxazole in water». *J. Am. Chem. Soc.* **117** (1995) 8600-8607 [Citado en pág. 69]
- [133] T. N. Truong, E. V. Stefanovich, «Development of a perturbative approach for Monte Carlo simulations using a hybrid ab initio QM/MM method». *Chem. Phys. Lett* **256** (1996) 348-352 [Citado en pág. 69]
- [134] I. Tuñón, M. T. C. Martins-Costa, C. Millot, M. F. Ruiz-López, «A hybrid density functional-classical molecular dynamics simulation of a water molecule in liquid water». *J. Mol. Model.* **1** (1995) 196-201 [Citado en págs. 69, 163 y 164]
- [135] I. Tuñón, M. T. C. Martins-Costa, C. Millot, M. F. Ruiz-López, «Molecular dynamics simulations of elementary chemical processes in liquid water using combined density functional and molecular mechanics potentials. I. Proton transfer in strongly H-bonded complexes». *J. Chem. Phys.* **106** (1997) 3633-3642 [Citado en págs. 69 y 163]
- [136] S. Chalmet, M. F. Ruiz-López, «Molecular dynamics simulation of formamide in water using density functional theory and classical potentials». *J. Chem. Phys.* **111** (1999) 1117-1125 [Citado en págs. 69, 133, 134 y 163]
- [137] M. E. Martín, M. L. Sánchez, F. J. Olivares del Valle, M. Á. Aguilar, «A theoretical study of liquid alcohols using averaged solvent electrostatic potentials obtained from molecular dynamics simulations: Methanol, ethanol and propanol». *J. Chem. Phys.* **116** (2002) 1613-1620 [Citado en pág. 73]
- [138] A. Muñoz Losa, I. Fdez. Galván, M. E. Martín, M. Á. Aguilar, «Theoretical study of liquid hydrogen fluoride. Application of the averaged solvent electrostatic potential/molecular dynamics method». *J. Phys. Chem. B* **107** (2003) 5043-5047 [Citado en págs. 75 y 171]
- [139] L. E. Chirlian, M. M. Francl, «Atomic charges derived from electrostatic potentials: A detailed study». *J. Comput. Chem.* **8** (1987) 894-905 [Citado en págs. 80 y 114]

- [140] C. M. Breneman, K. B. Wiberg, «Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis». *J. Comput. Chem.* **11** (1990) 361-373  
[Citado en págs. 80 y 114]
- [141] M. L. Sánchez, M. E. Martín, M. Á. Aguilar, F. J. Olivares del Valle, «Solvent effects on the  $^1(n,\pi^*)$  transition of formaldehyde in liquid water. A QM/MM study using the mean field approximation». *Chem. Phys. Lett.* **310** (1999) 195-200  
[Citado en pág. 83]
- [142] M. E. Martín, M. L. Sánchez, F. J. Olivares del Valle, M. Á. Aguilar, «A multiconfigurational self-consistent field/molecular dynamics study of the  $(n \rightarrow \pi)^1$  transition of carbonyl compounds in liquid water». *J. Chem. Phys.* **113** (2000) 6308-6315  
[Citado en pág. 83]
- [143] M. E. Martín, M. L. Sánchez, M. Á. Aguilar, F. J. Olivares del Valle, «Multiconfigurational self-consistent and molecular mechanics simulation of solvent effects on the  $n \rightarrow \pi^*$  blue shift of pyrimidine». *J. Mol. Struct. (Theochem)* **537** (2001) 213-222  
[Citado en pág. 83]
- [144] B. Linder, «Reaction-field techniques and their applications to intermolecular forces», en *Intermolecular Forces* (J. O. Hirschfelder, ed.), núm. 12 en *Advances in Chemical Physics*, Interscience Publishers (1967) págs. 225-281  
[Citado en pág. 86]
- [145] G. Karlström, B. Halle, «A fluctuation approach to solvation in polar fluids». *J. Chem. Phys.* **99** (1993) 8056-8062  
[Citado en pág. 86]
- [146] J.-L. Rivail, D. Rinaldi, «Liquid-state quantum chemistry: Computational applications of the polarizable continuum models», en *Computational Chemistry: Reviews of Current Trends* (J. Leszczynski, ed.), vol. 1, World Scientific (1996) págs. 139-174  
[Citado en pág. 87]
- [147] C. J. Cramer, D. G. Truhlar, «Continuum solvation models: Classical and quantum mechanical implementations», en *Reviews in Computational Chemistry* (K. B. Lipkowitz, D. B. Boyd, eds.), vol. 6, VCH Publishers (1995) págs. 1-72  
[Citado en pág. 87]
- [148] S. Miertuš, J. Bartoš, M. Trebatická, «Dependance of atomic radii and volumes on the electron distribution in solute molecule and on solute-solvent interaction». *J. Mol. Liq.* **33** (1987) 139-156  
[Citado en pág. 87]
- [149] M. Á. Aguilar, F. J. Olivares del Valle, «Solute-solvent interactions. A simple procedure for constructing the solvent cavity for retaining a molecular solute». *Chem. Phys.* **129** (1989) 439-450  
[Citado en pág. 87]



- [150] P. A. Kollman, «Free energy calculations: Applications to chemical and biochemical phenomena». *Chem. Rev.* **93** (1993) 2395-2417 [Citado en pág. 92]
- [151] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons (1999) [Citado en pág. 92]
- [152] A. W. Leach, *Molecular Modelling. Principles and Applications*, 2.<sup>a</sup> ed., Prentice Hall (2001) [Citado en págs. 92 y 102]
- [153] C. Chipot, D. A. Pearlman, «Free energy calculations. The long and winding gilded road». *Mol. Sim.* **28** (2002) 1-12 [Citado en pág. 92]
- [154] R. W. Zwanzig, «High-temperature equation of state by a perturbation method. I. Nonpolar gases». *J. Chem. Phys.* **22** (1954) 1420-1426 [Citado en pág. 93]
- [155] Y. Zhang, H. Liu, W. Yang, «Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined ab initio QM/MM potential energy surface». *J. Chem. Phys.* **112** (2000) 3483-3492 [Citado en págs. 96 y 197]
- [156] N. Okuyama-Yoshida, M. Nagaoka, T. Yamabe, «Transition-state optimization on free energy surface: Toward solution chemical reaction ergodography». *Int. J. Quantum Chem.* **70** (1998) 95-103 [Citado en pág. 99]
- [157] N. Okuyama-Yoshida, M. Nagaoka, T. Yamabe, «Structure optimization via free energy gradient method: Application to glycine zwitterion in aqueous solution». *J. Chem. Phys.* **113** (2000) 3519-3524 [Citado en págs. 99 y 130]
- [158] H. Hirao, Y. Nagae, M. Nagaoka, «Transition-state optimization by the free energy gradient method: Application to aqueous-phase Menshutkin reaction between ammonia and methyl chloride». *Chem. Phys. Lett.* **348** (2001) 350-356 [Citado en págs. 99, 152, 155 y 157]
- [159] C. Peng, P. Y. Ayala, H. B. Schlegel, M. J. Frisch, «Using redundant internal coordinates to optimize equilibrium geometries and transition states». *J. Comp. Chem.* **17** (1996) 49-56 [Citado en pág. 100]
- [160] R. Fletcher, *Practical Methods of Optimization*, 2.<sup>a</sup> ed., John Wiley & Sons (1987) [Citado en pág. 102]
- [161] T. Schlick, «Optimization methods in computational chemistry», en *Reviews in Computational Chemistry* (K. B. Lipkowitz, D. B. Boyd, eds.), vol. 3, VCH Publishers (1992) págs. 1-71 [Citado en pág. 102]



- 
- [162] H. B. Schlegel, «Exploring potential energy surfaces for chemical reactions: An overview of some practical methods». *J. Comput. Chem.* **24** (2003) 1514-1527 [Citado en pág. 102]
- [163] R. Fletcher, C. M. Reeves, «Function minimization by conjugate gradients». *Comput. J.* **7** (1964) 149-154 [Citado en pág. 103]
- [164] E. Polak, *Computational Methods in Optimization: A Unified Approach*, Academic Press (1971) [Citado en pág. 103]
- [165] W. C. Davidon, «Variance algorithm for minimization». *Comput. J.* **10** (1968) 406-410 [Citado en págs. 105 y 201]
- [166] R. Fletcher, «Function minimization without evaluating derivatives – A review». *Comput. J.* **8** (1965) 33-41 [Citado en págs. 105 y 201]
- [167] R. Fletcher, M. J. D. Powell, «A rapidly convergent descent method for minimization». *Comput. J.* **6** (1963) 163-168 [Citado en págs. 105 y 201]
- [168] C. G. Broyden, «The convergence of a class of double rank minimization algorithms. Part I». *J. Inst. Math. Appl.* **6** (1970) 76-90 [Citado en págs. 105 y 201]
- [169] C. G. Broyden, «The convergence of a class of double rank minimization algorithms. Part II». *J. Inst. Math. Appl.* **6** (1970) 222-231 [Citado en págs. 105 y 201]
- [170] R. Fletcher, «A new approach to variable metric algorithms». *Comput. J.* **13** (1970) 317-322 [Citado en págs. 105 y 201]
- [171] D. Goldfarb, «A family of variable metric methods derived by variational means». *Math. Comp.* **24** (1970) 23-26 [Citado en págs. 105 y 201]
- [172] D. F. Shanno, «Conditioning of quasi-Newton methods for function minimization». *Math. Comp.* **24** (1970) 647-656 [Citado en págs. 105 y 201]
- [173] J. Simons, P. Jørgensen, H. Taylor, J. Ozment, «Walking on potential energy surfaces». *J. Phys. Chem.* **87** (1983) 2745-2753 [Citado en pág. 105]
- [174] A. Banerjee, N. Adams, J. Simons, R. Shepard, «Search for stationary points on surfaces». *J. Phys. Chem.* **89** (1985) 52-57 [Citado en págs. 105 y 107]
- [175] J. Baker, «An algorithm for the location of transition states». *J. Comp. Chem.* **7** (1986) 385-395 [Citado en pág. 105]

- [176] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery Jr, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, Ö. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Men-  
nucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. A. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andrés, C. González, M. Head-Gordon, E. S. Replogle, J. A. Pople, *Gaussian 98 (Revision A.11.3)*, Gaussian, Inc., Pittsburgh PA (2001)  
[Citado en pág. 113]
- [177] J. P. Perdew, «Density-functional approximation for the correlation energy of the inhomogeneous electron gas». *Phys. Rev. B* **33** (1986) 8822-8824  
[Citado en pág. 113]
- [178] A. D. Becke, «A new mixing of Hartree-Fock and local density-functional theories». *J. Chem. Phys.* **98** (1993) 1372-1377  
[Citado en pág. 113]
- [179] K. Raghavachari, J. S. Binkley, R. Seeger, J. A. Pople, «Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions». *J. Chem. Phys.* **72** (1980) 650-654  
[Citado en pág. 113]
- [180] A. D. McLean, G. S. Chandler, «Contracted Gaussian basis sets for molecular calculations. I. Second row atoms,  $Z=11-18$ ». *J. Chem. Phys.* **72** (1980) 5639-5648  
[Citado en pág. 113]
- [181] T. H. Dunning Jr, «Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen». *J. Chem. Phys.* **90** (1989) 1007-1023  
[Citado en pág. 113]
- [182] R. A. Kendall, T. H. Dunning Jr, R. J. Harrison, «Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions». *J. Chem. Phys.* **96** (1992) 6796-6806  
[Citado en pág. 113]
- [183] D. E. Woon, T. H. Dunning Jr, «Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminium through argon». *J. Chem. Phys.* **98** (1993) 1358-1371  
[Citado en pág. 113]
- [184] P.-O. Widmark, P.-Å. Malmqvist, B. O. Roos, «Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. I. First row atoms». *Theor. Chim. Acta* **77** (1990) 291-306  
[Citado en pág. 114]

- [185] A. St-Amant, D. R. Salahub, «New algorithm for the optimization of geometries in local density functional theory». *Chem. Phys. Lett.* **169** (1990) 387-392  
[Citado en págs. 114 y 177]
- [186] N. Godbout, D. R. Salahub, J. Andzelm, E. Wimmer, «Optimization of Gaussian-type basis sets for local spin density functional calculations. Part I. Boron through neon, optimization technique and validation». *Can. J. Chem.* **70** (1992) 560-571  
[Citado en págs. 114 y 177]
- [187] K. Refson, «Moldy: A portable molecular dynamics simulation program for serial and parallel computers». *Comput. Phys. Commun.* **126** (2000) 310-329  
[Citado en pág. 114]
- [188] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, «Comparison of simple potential functions for simulating liquid water». *J. Chem. Phys.* **79** (1983) 926-935  
[Citado en págs. 115 y 116]
- [189] R. B. Lawrance, M. W. P. Strandberg, «Centrifugal distortion in asymmetric top molecules. I. Ordinary formaldehyde,  $\text{H}_2\text{C}^{12}\text{O}$ ». *Phys. Rev.* **83** (1951) 363-369  
[Citado en pág. 115]
- [190] R. W. Kilb, C. C. Lin, E. B. Wilson Jr, «Calculation of energy levels for internal torsion and over-all rotation. II.  $\text{CH}_3\text{CHO}$  type molecules; acetaldehyde spectra». *J. Chem. Phys.* **26** (1957) 1695-1703  
[Citado en pág. 115]
- [191] J. D. Swalen, C. C. Costain, «Internal rotation in molecules with two internal rotors: Microwave spectrum of acetone». *J. Chem. Phys.* **31** (1959) 1562-1574  
[Citado en pág. 115]
- [192] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta Jr, P. Weiner, «A new force field for molecular mechanical simulation of nucleic acids and proteins». *J. Am. Chem. Soc.* **106** (1984) 765-784  
[Citado en pág. 116]
- [193] S. J. Weiner, P. A. Kollman, D. T. Nguyen, D. A. Case, «An all atom force field for simulations of proteins and nucleic acids». *J. Comput. Chem.* **7** (1986) 230-252  
[Citado en pág. 116]
- [194] W. L. Jorgensen, C. J. Swenson, «Optimized intermolecular potential functions for amides and peptides. Structure and properties of liquid amides». *J. Am. Chem. Soc.* **107** (1985) 569-578  
[Citado en pág. 116]
- [195] S. N. Ha, A. Giammona, M. Field, J. W. Brady, «A revised potential-energy surface for molecular studies of carbohydrates». *Carbohydr. Res.* **180** (1988) 207-221  
[Citado en pág. 116]

- [196] M. Ferrario, M. Haughney, I. R. McDonald, M. L. Klein, «Molecular-dynamics simulation of aqueous mixtures: Methanol, acetone, and ammonia». *J. Chem. Phys.* **93** (1990) 5156-5166 [Citado en pág. 116]
- [197] H. A. Carlson, T. B. Nguyen, M. Orozco, W. L. Jorgensen, «Accuracy of free energies of hydration for organic molecules from 6-31G\*-derived partial charges». *J. Comput. Chem.* **14** (1993) 1240-1249 [Citado en pág. 116]
- [198] J. Gao, X. Xia, «A two-dimensional energy surface for a type II S<sub>N</sub>2 reaction in aqueous solution». *J. Am. Chem. Soc.* **115** (1993) 9667-9675 [Citado en págs. 116, 152, 155 y 156]
- [199] A. Muñoz Losa, I. Fdez. Galván, M. E. Martín, M. Á. Aguilar, «An averaged solvent electrostatic potential/molecular dynamics study of the influence of the electron correlation on the properties of liquid hydrogen fluoride». *J. Mol. Struct. (Theochem)* **632** (2003) 227-234 [Citado en pág. 118]
- [200] J. Applequist, J. R. Carl, K.-K. Fung, «Atom dipole interaction model for molecular polarizability. application to polyatomic molecules and determination of atom polarizabilities». *J. Am. Chem. Soc.* **94** (1972) 2952-2960 [Citado en pág. 119]
- [201] L. H. Johnston, R. P. Srivastava, R. M. Lees, «Laser-Stark spectrum of methyl alcohol at 890.761 GHz with the HCN laser». *J. Mol. Spectrosc.* **84** (1980) 1-40 [Citado en pág. 119]
- [202] R. K. Kakar, C. R. Quade, «Microwave rotational spectrum and internal rotation in gauche ethyl alcohol». *J. Chem. Phys.* **72** (1980) 4300-4307 [Citado en pág. 119]
- [203] A. A. Abdurakhmanov, E. I. Veliyulin, R. A. Ragimova, L. M. Imanov, «The microwave-spectrum of normal-propanol – the *gauche-gauche* conformer». *J. Struct. Chem.* **22** (1981) 28-33 [Citado en pág. 119]
- [204] A. A. Maryott, F. Buckley, «Table of dielectric constants and electric dipole movements of substances in the gaseous state», Circular 537, U.S. National Bureau of Standards (1953) [Citado en pág. 119]
- [205] J. W. C. Johns, A. R. W. McKellar, «Stark spectroscopy with the CO laser: The dipole moment of H<sub>2</sub>CO in the *v*<sub>2</sub>=2 state». *J. Chem. Phys.* **63** (1975) 1682-1685 [Citado en pág. 119]
- [206] P. H. Turner, A. P. Cox, «Microwave spectrum, structure, dipole moment and centrifugal distortion of nitrosomethane. Dipole moment of acetaldehyde». *J. Chem. Soc. Faraday Trans. II* **74** (1978) 533-559 [Citado en pág. 119]

- [207] R. Peter, H. Dreizler, «Das Mikrowellenspektrum von Aceton im Torsionsgrundzustand». *Z. Naturforsch.* **20A** (1965) 301-312 [Citado en pág. 119]
- [208] K. Coutinho, M. J. de Oliveira, S. Canuto, «Sampling configurations in Monte Carlo simulations for quantum mechanical studies of solvent effects». *Int. J. Quantum Chem.* **66** (1998) 249-253 [Citado en pág. 119]
- [209] K. Coutinho, S. Canuto, M. C. Zerner, «A Monte Carlo-quantum mechanics study of the solvatochromic shifts of the lowest transition of benzene». *J. Chem. Phys.* **112** (2000) 9874-9880 [Citado en pág. 119]
- [210] F. J. Olivares del Valle, M. Á. Aguilar, S. Tolosa, «Polarizable continuum model calculations including electron correlation in the ab initio wavefunction». *J. Mol. Struct. (Theochem)* **279** (1993) 223-231 [Citado en pág. 121]
- [211] S. E. Barrows, J. W. Storer, C. J. Cramer, A. D. French, D. G. Truhlar, «Factors controlling relative stability of anomers and hydroxymethyl conformers of glucopyranose». *J. Comput. Chem.* **19** (1998) 1111-1129 [Citado en págs. 125, 140 y 141]
- [212] T. A. Wesolowski, A. Warshel, «Frozen density functional approach for ab initio calculations of solvated molecules». *J. Phys. Chem.* **97** (1993) 8050-8053 [Citado en pág. 135]
- [213] P. Ahlström, A. Wallqvist, S. Engström, B. Jönsson, «A molecular dynamics study of polarizable water». *Mol. Phys.* **68** (1989) 563-581 [Citado en pág. 135]
- [214] K. Laasonen, M. Sprik, M. Parrinello, R. Car, «“Ab initio” liquid water». *J. Chem. Phys.* **99** (1993) 9080-9089 [Citado en pág. 135]
- [215] G. Jansen, F. Colonna, J. G. Ángyán, «Mixed quantum-classical calculations on the water molecule in liquid phase: Influence of a polarizable environment on electronic properties». *Int. J. Quantum Chem.* **58** (1996) 251-265 [Citado en pág. 135]
- [216] P. L. Silvestrelli, M. Parrinello, «Structural, electronic, and bonding properties of liquid water from first principles». *J. Chem. Phys.* **111** (1999) 3572-3580 [Citado en pág. 135]
- [217] D. Eisenberg, W. Kauzmann, *The Structure and Properties of Water*, Oxford University Press (1969) [Citado en pág. 135]
- [218] C. A. Coulson, D. Eisenberg, «Interactions of H<sub>2</sub>O molecules in ice». *Proc. R. Soc. Lond. Ser. A* **291** (1966) 445-453 [Citado en pág. 135]

- [219] E. Whalley, «A relation between the strengths of the orientation polarization and the infrared absorption of the O–H stretching vibrations of ice». *Chem. Phys. Lett.* **53** (1978) 449-451 [Citado en pág. 135]
- [220] Y. S. Badyal, M.-L. Saboungi, D. L. Price, S. D. Shastri, D. R. Haefner, A. K. Soper, «Electron distribution in water». *J. Chem. Phys.* **112** (2000) 9206-9208 [Citado en pág. 135]
- [221] C. J. Cramer, D. G. Truhlar, «Quantum chemical conformational analysis of glucose in aqueous solution». *J. Am. Chem. Soc.* **115** (1993) 5745-5753 [Citado en pág. 140]
- [222] B. D. Wladkowski, S. A. Chenoweth, K. E. Jones, J. W. Brown, «Exocyclic hydroxymethyl rotational conformers of  $\beta$ - and  $\alpha$ -D-glucopyranose in the gas phase and aqueous solution». *J. Phys. Chem. A* **102** (1998) 5086-5092 [Citado en págs. 140 y 141]
- [223] F. A. Momany, M. Appell, G. L. Strati, J. L. Willett, «B3LYP/6-311++G\*\* study of monohydrates of  $\alpha$ - and  $\beta$ -D-glucopyranose: Hydrogen bonding, stress energies, and effect of hydration on internal coordinates». *Carbohydr. Res.* **339** (2004) 553-567 [Citado en pág. 140]
- [224] J.-H. Lii, B. Ma, N. L. Allinger, «Importance of selecting proper basis set in quantum mechanical studies of potential energy surfaces of carbohydrates». *J. Comput. Chem.* **20** (1999) 1593-1603 [Citado en pág. 140]
- [225] A. J. Kirby, *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*, Springer-Verlag, Berlin (1983) [Citado en pág. 140]
- [226] S. J. Angyal, «The composition and conformation of sugars in solution». *Angew. Chem. Int. Ed.* **8** (1969) 157-166 [Citado en pág. 140]
- [227] P. L. Polavarapu, C. S. Ewig, «Ab initio computed molecular structures and energies of the conformers of glucose». *J. Comput. Chem.* **13** (1992) 1255-1261 [Citado en pág. 140]
- [228] S. E. Barrows, F. J. Dulles, C. J. Cramer, A. D. French, D. G. Truhlar, «Relative stability of alternative chair forms and hydroxymethyl conformations of  $\beta$ -D-glucopyranose». *Carbohydr. Res.* **276** (1995) 219-251 [Citado en págs. 140 y 141]
- [229] M. Appell, G. L. Strati, J. L. Willet, F. A. Momany, «B3LYP/6-311++G\*\* study of  $\alpha$ - and  $\beta$ -D-glucopyranose and 1,5-anhydro-D-glucitol:  ${}^4C_1$  and  ${}^1C_4$  chairs,  ${}^3O_B$  and  $B_{3,O}$  boats, and skew-boat conformations». *Carbohydr. Res.* **339** (2004) 537-551 [Citado en pág. 140]



- [230] C. J. Cramer, D. G. Truhlar, A. D. French, «Exo-anomeric effects on energies and geometries of different conformations of glucose and related systems in the gas phase and aqueous solution». *Carbohydr. Res.* **298** (1997) 1-14  
[Citado en pág. 141]
- [231] X. Prat-Resina, M. Garcia-Viloca, G. Monard, A. González-Lafont, J. M. Lluch, J. M. Bofill, J. M. Anglada, «The search for stationary points on a quantum mechanical/molecular mechanical potential-energy surface». *Theor. Chem. Acc.* **107** (2002) 147-153  
[Citado en págs. 142 y 202]
- [232] D. A. Pearlman, P. A. Kollman, «The overlooked bond-stretching contribution in free energy perturbation calculations». *J. Chem. Phys.* **94** (1991) 4532-4545  
[Citado en pág. 142]
- [233] D. A. Pearlman, «A comparison of alternative approaches to free energy calculations». *J. Phys. Chem.* **98** (1994) 1487-1493  
[Citado en pág. 142]
- [234] C. Höög, G. Widmalm, «Free energy simulations of D-xylose in water and methyl D-xylopyranoside in methanol». *J. Phys. Chem. B* **105** (2001) 6375-6379  
[Citado en pág. 145]
- [235] C. Molteni, M. Parrinello, «Glucose in aqueous solution by first principles molecular dynamics». *J. Am. Chem. Soc.* **120** (1998) 2168-2171  
[Citado en pág. 147]
- [236] Q. Liu, J. W. Brady, «Anisotropic solvent structuring in aqueous sugar solutions». *J. Am. Chem. Soc.* **118** (1996) 12276-12286  
[Citado en pág. 147]
- [237] Q. Liu, J. W. Brady, «Model dependence of the anisotropic structuring of solvent water around sugars in molecular dynamics simulations». *J. Phys. Chem. B* **101** (1997) 1317-1321  
[Citado en pág. 147]
- [238] H. Eyring, «The activated complex in chemical reactions». *J. Chem. Phys.* **3** (1935) 107-115  
[Citado en pág. 152]
- [239] N. A. Menshutkin. *Z. Physik. Chem.* **5** (1890) 589  
[Citado en pág. 152]
- [240] N. A. Menshutkin. *Z. Physik. Chem.* **6** (1890) 41  
[Citado en pág. 152]
- [241] T. N. Truong, T.-T. T. Truong, E. V. Stefanovich, «A general methodology for quantum modeling of free-energy profile of reactions in solution: An application to the Menshutkin  $\text{NH}_3 + \text{CH}_3\text{Cl}$  reaction in water». *J. Chem. Phys.* **107** (1997) 1881-1889  
[Citado en págs. 152, 154 y 155]
- [242] C. Amovilli, B. Mennucci, F. M. Floris, «MCSCF study of the  $\text{S}_{\text{N}}2$  Menshutkin reaction in aqueous solution within the polarizable continuum model». *J. Phys. Chem. B* **102** (1998) 3023-3028  
[Citado en págs. 152 y 155]

- [243] X. Fradera, L. Amat, M. Torrent, J. Mestres, P. Constans, E. Besalú, J. Martí, S. Simon, M. Lobato, J. M. Oliva, J. M. Luis, J. L. Andrés, M. Solà, R. Carbó-Dorca, M. Duran, «Analysis of the changes on the potential energy surface of Menshutkin reactions induced by external perturbations». *J. Mol. Struct. (Theochem)* **371** (1996) 171-183 [Citado en pág. 152]
- [244] K. Naka, H. Sato, A. Morita, F. Hirata, S. Kato, «RISM-SCF study of the free-energy profile of the Menshutkin-type reaction  $\text{NH}_3 + \text{CH}_3\text{Cl} \rightarrow \text{NH}_3\text{CH}_3^+ + \text{Cl}^-$  in aqueous solution». *Theor. Chem. Acc.* **102** (1999) 165-169 [Citado en págs. 152 y 155]
- [245] D. G. Truhlar, A. D. Isaacson, B. C. Garrett, «Generalized transition state theory», en *Theory of Chemical Reaction Dynamics* (M. Baer, ed.), vol. IV, CRC Press, Boca Raton, FL (1985) págs. 65-130 [Citado en pág. 153]
- [246] J. P. de Jongh, H. A. Dijkerman, «Microwave spectrum of methyl chloride in excited vibrational states». *J. Mol. Spectrosc.* **25** (1968) 129-130 [Citado en pág. 154]
- [247] M. D. Marshall, K. C. Izgi, J. S. Muentner, «IR-microwave double resonance studies of dipole moments in  $\nu_1$  and  $\nu_3$  states of ammonia». *J. Chem. Phys.* **107** (1997) 1037-1044 [Citado en pág. 154]
- [248] A. Halkier, P. R. Taylor, «A theoretical investigation of the equilibrium electric dipole moment of ammonia». *Chem. Phys. Lett.* **285** (1998) 133 [Citado en pág. 154]
- [249] G. S. Hammond, «A correlation of reaction rates». *J. Am. Chem. Soc.* **77** (1955) 334-338 [Citado en pág. 155]
- [250] J. Gao, «A priori computation of a solvent-enhanced  $\text{S}_{\text{N}}2$  reaction profile in water: the Menshutkin reaction». *J. Am. Chem. Soc.* **113** (1991) 7796-7797 [Citado en pág. 155]
- [251] K. Okamoto, S. Fukui, H. Shingu, «Kinetic studies of bimolecular nucleophilic substitution. 6. Rates of Menshutkin reaction of methyl iodide with methylamines and ammonia in aqueous solutions». *Bull. Chem. Soc. Jpn.* **40** (1967) 1920 [Citado en pág. 156]
- [252] D. B. Kimball, M. M. Haley, «Triazenes: A versatile tool in organic synthesis». *Angew. Chem. Int. Ed.* **41** (2002) 3338-3351 [Citado en pág. 161]
- [253] J. C. Scaiano, C. Chen, P. F. McGarry, «A flash photolysis and optoacoustic calorimetry study of the cis-trans isomerization of 1,3-diphenyltriazene». *J. Photochem. Photobiol. A: Chem.* **62** (1991) 75-81 [Citado en pág. 161]



- [254] M. Barra, N. Chen, «Kinetic studies of the thermal cis–trans isomerization of 1,3-diphenyltriazene in aqueous solution. Effects of acids and bases». *J. Org. Chem.* **65** (2000) 5739-5744 [Citado en pág. 161]
- [255] N. Chen, M. Barra, I. Lee, N. Chahal, «Substituent effects on the thermal cis-to-trans isomerization of 1,3-diphenyltriazenes in aqueous solution». *J. Org. Chem.* **67** (2002) 2271-2277 [Citado en pág. 161]
- [256] J. W. Sutherland, «Pulse radiolysis of aqueous hydrazine solutions. The triazene species». *J. Phys. Chem.* **83** (1979) 789-795 [Citado en pág. 161]
- [257] R. H. Smith Jr, C. L. Denlinger, R. Kupper, S. R. Koepke, C. J. Michejda, «Specific acid catalysis in the decomposition of trialkyltriazenes». *J. Am. Chem. Soc.* **106** (1984) 1056-1059 [Citado en pág. 161]
- [258] R. H. Smith Jr, C. L. Denlinger, R. Kupper, A. F. Mehl, C. J. Michejda, «Decomposition of 1,3-dialkyltriazenes in aqueous buffers: Kinetic and mechanistic studies». *J. Am. Chem. Soc.* **108** (1986) 3726-3730 [Citado en pág. 161]
- [259] R. H. Smith Jr, B. D. Wladkowski, A. F. Mehl, M. J. Cleveland, E. A. Rudrow, G. N. Chmurny, C. J. Michejda, «1,3-Dialkyltriazenes: Tautomeric equilibria and rates and products of decomposition». *J. Org. Chem.* **54** (1989) 1036-1042 [Citado en págs. 161 y 178]
- [260] M. N. Ramos, S. R. Pereira, «Conformation, electronic structure, and biological activity of antitumour triazenes». *J. Chem. Soc. Perkin Trans. 2* (1986) 131-133 [Citado en pág. 161]
- [261] A. Schmiedekamp, R. H. Smith Jr, C. J. Michejda, «Ab initio studies of triazenes in relation to experimental findings». *J. Org. Chem.* **53** (1988) 3433-3436 [Citado en pág. 161]
- [262] A. Schmiedekamp, I. A. Topol, S. K. Burt, H. Razafinjanahary, H. Chermette, T. Pfaltzgraff, C. J. Michejda, «Triazene proton affinities: A comparison between density functional, hartree–fock, and post-hartree–fock methods». *J. Comput. Chem.* **15** (1994) 875-892 [Citado en pág. 161]
- [263] P. Charbonneau, B. Jean-Claude, M. A. Whitehead, «Synthesis of a prodrug: A semi-empirical PM3 study». *J. Mol. Struct. (Theochem)* **574** (2001) 85-91 [Citado en pág. 161]
- [264] C. C. Pye, K. Vaughan, J. F. Glister, «An ab initio study of conformations and sigmatropic shifts in triazene and its mono-, di-, and trimethyl derivatives». *Can. J. Chem.* **80** (2002) 447-454 [Citado en págs. 161, 177 y 178]

- [265] M. Strnad, M. T. C. Martins-Costa, C. Millot, I. Tuñón, M. F. Ruiz-López, J.-L. Rivail, «Molecular dynamics simulations of elementary chemical processes in liquid water using combined density functional and molecular mechanics potentials. II. Charge separation processes». *J. Chem. Phys.* **106** (1997) 3643-3657 [Citado en pág. 163]
- [266] S. Chalmet, W. Harb, M. F. Ruiz-López, «Computer simulation of amide bond formation in aqueous solution». *J. Phys. Chem. A* **105** (2001) 11574-11581 [Citado en pág. 163]
- [267] J. Keck, «Statistical investigation of dissociation cross-sections for diatoms». *Discuss. Faraday Soc.* **33** (1962) 173-182 [Citado en pág. 163]
- [268] J. B. Anderson, «Statistical theories of chemical reactions. Distributions in the transition region». *J. Chem. Phys.* **58** (1973) 4684-4692 [Citado en pág. 163]
- [269] D. Chandler, «Barrier crossings: Classical theory of rare but important events», en *Classical and Quantum Dynamics in Condensed Phase Simulations* (B. J. Berne, G. Ciccotti, D. F. Coker, eds.), World Scientific (1998) págs. 3-23 [Citado en pág. 163]
- [270] M. E. Martín, M. Á. Aguilar, S. Chalmet, M. F. Ruiz-López, «A comparative study of two QM/MM methods testing the validity of the mean field approximation». *Chem. Phys. Lett.* **344** (2001) 107-112 [Citado en pág. 163]
- [271] C. Curutchet, C. J. Cramer, D. G. Truhlar, M. F. Ruiz-López, D. Rinaldi, M. Orozco, F. J. Luque, «Electrostatic component of solvation: Comparison of SCRf continuum models». *J. Comput. Chem.* **24** (2003) 284-297 [Citado en pág. 164]
- [272] A. Bondi, «Van der Waals volumes and radii». *J. Phys. Chem.* **68** (1964) 441-451 [Citado en pág. 165]
- [273] J. P. Bergsma, J. R. Reimers, K. R. Wilson, J. T. Hynes, «Molecular dynamics of the A + BC reaction in rare gas solution». *J. Chem. Phys.* **85** (1986) 5625-5643 [Citado en pág. 179]
- [274] B. J. Gertner, K. R. Wilson, J. T. Hynes, «Nonequilibrium solvation effects on reaction rates for model S<sub>N</sub>2 reactions in water». *J. Chem. Phys.* **90** (1989) 3537-3558 [Citado en pág. 179]
- [275] B. C. Garret, D. G. Truhlar, «Criterion of minimum state density in the transition state theory of bimolecular reactions». *J. Chem. Phys.* **70** (1979) 1593-1598 [Citado en pág. 187]

- [276] D. G. Truhlar, B. C. Garret, «Variational transition-state theory». *Acc. Chem. Res.* **13** (1980) 440-448 [Citado en pág. 187]
- [277] B. T. Thole, «Molecular polarizabilities calculated with a modified dipole interaction». *Chem. Phys.* **59** (1981) 341-350 [Citado en pág. 194]



---

---

## ARTÍCULOS PUBLICADOS

---

---



## Theoretical Calculation of the Stark Component of the Solute–Solvent Interaction Energy. Validity of the Mean Field Approximation in the Study of Liquids and Solutions

M. L. Sánchez, M. E. Martín, I. Fdez. Galván, F. J. Olivares del Valle, and M. A. Aguilar\*

Dpto. Química-Física, Universidad de Extremadura. Avda de Elvas s/n, 06071 Badajoz, Spain

Received: January 11, 2002; In Final Form: March 13, 2002

We examine the Stark component of the solute–solvent interaction energy and check the validity of the mean field approximation (MFA) in the theoretical study of liquids and solutions. We considered two types of systems: methanol, ethanol, and propanol liquids and formaldehyde, acetaldehyde, and acetone in aqueous solution. We found that, independent of the level of calculation (HF, MP2, or MCSCF), the errors introduced by MFA are less than 5% in the interaction energy and less than 1% in the dipoles. We propose an approximate expression for the Stark component that reduces the errors in the interaction energy to below 1.6%.

### I. Introduction

Many of the most frequently used methods for the study of solvent effects make use of the mean field approximation (MFA). The different quantum versions of continuum models (SCRF,<sup>1</sup> PCM,<sup>2</sup> multipole expansions,<sup>3</sup> etc.) or the methods based on Langevin dipoles<sup>4</sup> are representative examples. In all these methods, the solute is represented quantum mechanically, whereas the solvent response is described classically through very simplified models that completely neglect the microscopic structure of the solvent and the possible presence of specific interactions. Because of this, several methods have been proposed in the past decade that combine MFA with a detailed description of the solvent structure obtained either from integral theories, the reference interaction site method in the case of RISM/HF,<sup>5</sup> or from simulations, when the averaged solvent electrostatic potential/molecular dynamics method, ASEP/MD,<sup>6</sup> is used.

In MFA-based methods, the energy and wave function of the solute molecule are obtained in the presence of the averaged perturbation generated by the solvent.<sup>2b</sup> The main advantage of this approximation is that it notably reduces the number of quantum calculations. For instance, whereas in traditional quantum mechanics/molecular mechanics<sup>7</sup> (QM/MM) methods (where MFA is not used), it is necessary to perform one quantum calculation for each chosen solvent configuration, ASEP/MD<sup>6</sup> or RISM/HF<sup>5</sup> require only 3–10 quantum calculations depending on the number of cycles needed to reach convergence.

As with any other approximation, MFA introduces errors into the evaluation of the solute–solvent interaction energy and the solute properties. These errors are associated with the neglect of the correlation between the motion of the solvent nuclei and the response of the solute electron polarizability; that is, MFA does not allow the solute to polarize in response to a change in the solvent nuclear configuration. This correlation energy is usually known as the Stark component.<sup>8,9</sup>

Despite the extended use of MFA in solvent effect theories, few studies have focused on the evaluation of the Stark component. Linder,<sup>8</sup> and later Karlstrom and Halle,<sup>9</sup> proposed expressions for this component based on continuum models. In both cases, the magnitude of the Stark component was related to the solute polarizability and the difference between the dipole

reaction factors evaluated at zero and optical frequencies:

$$\Delta W_{\text{Stark}} = -\frac{1}{2} K_{\text{B}} T \sum_l \sum_k \alpha_{lk} [g_{lk}(\epsilon_0) - g_{lk}(\epsilon_\infty)] \quad (1)$$

where  $K_{\text{B}}$  is the Boltzmann constant,  $T$  is the absolute temperature,  $\alpha_{lk}$  is the  $lk$  component of the solute polarizability tensor, and  $g_{lk}$  is the  $lk$  component of the Onsager factor tensor evaluated at zero and optical frequencies. Like the rest of the expressions based on continuum theories, eq 1 neglects the effects of the microscopic structure of the solvent and of the specific interactions. Furthermore, a value has to have been assigned to the cavity radius that holds the solute molecule, and this is not always easy.<sup>3b,10</sup>

As was indicated above, the ASEP/MD theory<sup>6</sup> combines MFA with MD calculations and, hence, does not suffer from the limitations of continuum models. We here use this theory to evaluate the Stark component calculated as the difference between the solute–solvent interaction energy when the MFA is used and when it is not. Our aim is 2-fold. First, we shall evaluate the Stark component for a few representative systems. This is an important preliminary step because the validity of the MFA-based theories depends on the supposition that this component remains small. Second, we shall show how an estimate of this component can easily be obtained from knowledge of the solvent electric field fluctuations at the position occupied by the solute.

The rest of the paper is organized as follows. In sections II and III, we detail the method that we follow in the determination of the magnitude of the Stark component and present an approximate expression that permits a rapid evaluation of this component. Section IV presents the results for various hydrogen-bonded systems using several levels of calculations. Section V presents our conclusions.

### II. Methods

**(a) Direct Determination of the Stark Component.** For a direct determination of the magnitude of the errors introduced by the MFA, we have to compare a full calculation of the energy and properties of a solute molecule in solution with the result when the MFA is introduced.

4814 *J. Phys. Chem. B*, Vol. 106, No. 18, 2002

We use a QM/MM method to obtain the solute wave function in solution. The solvent effect theory involving coupling between a quantum and a classical system has been widely discussed.<sup>7</sup> For clarity, we shall summarize the main points of the implementation of the method.

The Hamiltonian model is partitioned in the usual way:

$$H = H_{QM} + H_{MM} + H_{QM/MM} \quad (2)$$

with terms that correspond to the quantum part,  $H_{QM}$ , the classical part,  $H_{MM}$ , and the interaction between them,  $H_{QM/MM}$ . In our case, the quantum part is formed by the solute molecule, whereas the classical part includes all of the solvent molecules.

The energy and wave function of the solvated solute molecule are obtained by solving the effective Schrödinger equation for each configuration:

$$(H_{QM} + H_{QM/MM})|\Psi\rangle = E|\Psi\rangle \quad (3)$$

The interaction term,  $H_{QM/MM}$ , takes the following form:

$$\hat{H}_{QM/MM} = H_{QM/MM}^{\text{elect}} + H_{QM/MM}^{\text{vdw}} \quad (4)$$

$$\hat{H}_{QM/MM}^{\text{elect}} = \int dr \hat{\rho} \cdot \hat{V}_s(r; X_i) \quad (5)$$

where  $\hat{\rho}$  is the solute charge density and  $\hat{V}_s(r; X)$  is the electrostatic potential generated at the position  $r$  by the solvent configuration  $X_i$ . If the solvent is regarded as nonpolarizable, then this potential is nothing more than the electrostatic potential generated by the charges that represent the solvent molecules during the simulation. If the solvent is polarizable, we have also to include the contribution from the induced dipoles. The term  $H_{QM/MM}^{\text{vdw}}$  is the Hamiltonian for the van der Waals interaction, which is generally represented by a Lennard-Jones potential. Schrödinger's equation has to be solved for each point of configuration space, usually by sampling through molecular dynamics or Monte Carlo simulations.<sup>11</sup>

The desired properties are then obtained by averaging over the  $N$  solvent configurations

$$\langle A \rangle = \frac{1}{N} \sum_{i=1}^N A(X_i) \quad (6)$$

where the different  $A(X_i)$  values are obtained from the solute wave function,  $\Psi(X_i)$ , calculated in the presence of the solvent configuration  $X_i$ .

On the contrary, in MFA, one averages the solvent perturbation,  $\hat{H}_{QM/MM}^{\text{elect}}$ :

$$\hat{H}_{QM/MM}^{\text{elect}} = \int dr \hat{\rho} \cdot \langle \hat{V}_s(r) \rangle \quad (7)$$

where the brackets denote a statistical average. The term  $\langle \hat{V}_s(r) \rangle$  is usually known as the averaged solvent electrostatic potential (ASEP). Substituting (7) into (4) and (3), we obtain the energy,  $E_{MFA}$ , and wave function,  $\Psi_{MFA}$ , of the solute in the presence of the averaged solvent configuration. The obvious advantage of this approximation is that only one quantum calculation has to be performed. The Stark component is then calculated as the difference between the energies obtained with eqs 6 and 7, that is,  $W_{\text{Stark}} = \langle E \rangle - E_{MFA}$ .

Different mean field models can be obtained if we vary the description of the solute-solvent interaction term (point charges, point charges + polarizabilities, multipole expansion, pseudopotentials, etc.) or the way in which the different solvent configurations are obtained (MD, MC, integral theories, etc.)

In our case, we obtain the solvent configuration from MD, and the ASEP is represented by a set of point charges chosen in the following way.

1. We include the charges belonging to solvent molecules that, in any of the MD configurations, lie within a sphere of radius  $R$  chosen in such a way that it includes the first solvation shell. The value of any charge is then divided by the number of solvent configurations included in the calculation of the ASEP.

2. Next, a second set of charges is obtained by a least-squares fit to the values of the electrostatic potential originated by the solvent molecules lying beyond the first solvation shell.<sup>16</sup> The details have been given elsewhere.<sup>6a,6d</sup>

The total number of charges introduced into the perturbation Hamiltonian is about 5000.

#### (b) Approximate Evaluation of the Stark Component.

A very simple formula to evaluate the magnitude of the Stark component can be obtained if we represent the solute charge distribution by a dipole moment. In this case, the electrostatic solute-solvent interaction energy is

$$W(X_i) = \mu_0 E(X_i) + \frac{1}{2} (\alpha E(X_i)) E(X_i) = \mu_0 E(X_i) + \frac{1}{2} \alpha : [E(X_i) \otimes E(X_i)] \quad (8)$$

where  $E$  is the electric field generated by the solvent in the position occupied by the solute when the solvent is at the configuration  $X_i$ ,  $\alpha$  is the solute electronic polarizability tensor, “:” is the double scalar product of two second-order tensors defined as  $A:B = \sum_i \sum_j A_{ij} B_{ij}$ , and “ $\otimes$ ” denotes the tensor product.

The final energy is obtained by averaging over all of the configurations:

$$\langle W \rangle = \mu_0 \langle E \rangle + \frac{1}{2} \alpha : \langle E \otimes E \rangle \quad (9)$$

where the brackets denote averaged values.

In MFA, we calculate the energy of the averaged solvent configuration as

$$W = (\mu_0 + \frac{1}{2} \alpha \langle E \rangle) \langle E \rangle = \mu_0 \langle E \rangle + \frac{1}{2} \alpha : \langle E \rangle \otimes \langle E \rangle \quad (10)$$

The difference between (9) and (10) yields an approximate formula for the Stark component that is valid when the solute charge distribution is well described by a dipole moment:

$$\Delta W_{\text{Stark}} = \frac{1}{2} \alpha : (\langle E \otimes E \rangle - \langle E \rangle \otimes \langle E \rangle) \quad (11)$$

$$= \frac{1}{2} \sum_i \sum_j \alpha_{ij} (\langle E_i E_j \rangle - \langle E_i \rangle \langle E_j \rangle) \quad (12)$$

Equation 12 is obtained from eq 11 by resolving the double scalar and the tensor product. One sees that in this expression  $\Delta W_{\text{Stark}}$  is proportional to the solute polarizability and to the fluctuations of the electric field generated by the solvent at the position occupied by the solute.

### III. Details of the Calculations

We studied two types of systems, pure liquids and solutions. As examples of liquids, we considered methanol, ethanol, and propanol, and as examples of solutions, we considered formaldehyde, acetaldehyde, and acetone in aqueous solution. In all these cases, there are strong hydrogen bond interactions.

In the three pure liquid systems, the geometries were taken from Jorgensen et al.<sup>12</sup> and the basis set used was the aug-cc-pVDZ from Dunning et al.<sup>13</sup> This basis set yields values for



Theoretical Calculation of the Stark Component

**TABLE 1: In Vacuo Dipole Moments and Polarizabilities of the Methanol, Ethanol, and Propanol Molecules<sup>a</sup>**

	$\mu^0$	$\alpha^0$	$\mu^0(\text{exp})^{26,27}$	$\alpha^0(\text{exp})^{28,29}$
methanol	1.81	19.52	1.70	21.79
ethanol	1.72	31.00	1.68	34.48
propanol	1.65	42.12	1.55	45.49

<sup>a</sup> Dipole moments are in debyes. Polarizabilities are in au.

**TABLE 2: In Vacuo Dipole Moments and Polarizabilities of the Formaldehyde, Acetaldehyde, and Acetone Molecules<sup>a</sup>**

	$\mu^0$	$\alpha^0$	$\mu^0(\text{exp})$	$\alpha^0(\text{exp})$
formaldehyde	2.35	14.40	2.36 <sup>b</sup>	16.53 <sup>c</sup>
acetaldehyde	2.76	25.38	2.75 <sup>d</sup>	31.04 <sup>e</sup>
acetone	2.90	35.58	2.93 <sup>e</sup>	42.52 <sup>f</sup>

<sup>a</sup> Dipole moments are in debyes. Polarizabilities are in au. <sup>b</sup> Reference 30. <sup>c</sup> Reference 29. <sup>d</sup> Reference 31. <sup>e</sup> Reference 32. <sup>f</sup> Reference 33.

the in vacuo dipole moment and polarizability close to the experimental values (see Table 1). The dipole moments are overestimated by about 0.1 D, and the polarizabilities are underestimated by about 3 au. The experimental trends of decreasing dipole moments and increasing polarizabilities with the size of the alkyl groups are well reproduced. Wave functions and energies were calculated at Hartree–Fock (HF) level, and second-order Møller–Plesset (MP2) perturbation theory corrections to the energy (PTE) were added when indicated.

For formaldehyde,<sup>14</sup> acetaldehyde,<sup>15</sup> and acetone,<sup>16</sup> the experimental geometries and generally contracted basis sets of the atomic natural orbital (ANO) type<sup>17</sup> were used. The latter were obtained from 14s9p4d/8s primitive sets and contracted to 4s3p1d/2s. For each molecule, a Hartree–Fock (HF) calculation was first performed. The corresponding orbitals were used as starting orbitals in the subsequent CASSCF (complete active space self-consistent field) calculations. The active space is spanned by all of the configurations arising from six valence electrons in four orbitals. Table 2 gives the in vacuo dipole moments and polarizabilities for these molecules. The dipole moments are very well reproduced, but the polarizabilities are underestimated.

All of the quantum calculations were performed using the Gaussian 94 package.<sup>18</sup>

The MD calculations were performed using the program MOLLY.<sup>19</sup> In each case 128 (alcohols) or 215 (carbonyls) molecules were simulated at fixed intramolecular geometry by combining Lennard-Jones interatomic interactions<sup>12,20</sup> with electrostatic interactions. In the carbonyl–water system, the 214 solvent molecules were simulated by the TIP3P<sup>21</sup> model at fixed intramolecular geometry. Periodic boundary conditions were applied, and spherical cutoffs were used to truncate the molecular interactions at 9.0, 11.5, and 12.5 Å for methanol, ethanol, and propanol, respectively, and 9 Å for formaldehyde, acetaldehyde, and acetone. A time step of 0.5 fs was used. The electrostatic interaction was calculated with the Ewald method.<sup>22</sup> The temperature was fixed at 298 K by using a Nosé–Hoover<sup>23</sup> thermostat. Each simulation of the self-consistent process was run for 150 000 time steps (50 000 equilibration, 100 000 production). The solute and solvent coordinates were dumped at every 100 or 1000 steps for further analysis.

Once the different solvent configurations have been obtained, we calculate the solute–solvent interaction energy in the two ways indicated above. In the first, we perform a quantum calculation for each of the  $N$  solvent configurations and obtain  $N$  values of the interaction energy and the dipole moment. Next, the mean value is obtained (values in brackets). In the second

J. Phys. Chem. B, Vol. 106, No. 18, 2002 4815

**TABLE 3: Interaction Energy (in kcal/mol) and Dipole Moment (in debyes) of a Methanol Molecule in the Liquid State as a Function of the Number of Configurations Included in the Calculation of the Average<sup>a</sup>**

$N$ configurations	$\langle E \rangle$	$\langle \mu \rangle$	$N$ configurations	$\langle E \rangle$	$\langle \mu \rangle$
50	-19.28	2.47	333	-19.02	2.46
100	-19.07	2.46	500	-19.01	2.46
200	-19.02	2.46	1000	-19.01	2.46
250	-19.02	2.46			

<sup>a</sup> Results obtained at the Hartree–Fock level.

**TABLE 4: Interaction Energy (in kcal/mol) and Dipole Moment (in debyes) of a Methanol Molecule in the Liquid State as a Function of the Time Window Chosen to Calculate the Average<sup>a</sup>**

interval (ps)	$\langle E \rangle$	$\langle \mu \rangle$	interval (ps)	$\langle E \rangle$	$\langle \mu \rangle$	interval (ps)	$\langle E \rangle$	$\langle \mu \rangle$
0:5	18.95	2.41	20:25	18.73	2.48	35:40	20.13	2.49
5:10	18.78	2.44	25:30	19.26	2.50	40:45	20.64	2.52
10:15	17.44	2.39	30:35	20.79	2.56	45:50	17.59	2.38
15:20	17.86	2.46						

<sup>a</sup> Results obtained at the Hartree–Fock level.

type of calculation, from the  $N$  configurations, we obtain the averaged potential, ASEP, which is then introduced into the solute molecular Hamiltonian. By solving the Schrödinger equation, we get the values of the interaction energy and dipole moment in the mean field approximation.

#### IV. Results

Let us begin by analyzing the effect of the number of configurations used to calculate the averages on the values of the desired properties. Table 3 gives the results for methanol. It is necessary to clarify that, independent of their number, the configurations were chosen so as to span the complete time of the simulation, 50 ps. With 50 configurations (and hence 50 quantum calculations) the error (when compared with the 1000 configuration value) is only 0.3 kcal/mol in the energy and 0.01 D in the dipole moment. With 100 configurations, the errors decrease to 0.06 kcal/mol and 0.00 D, respectively. With 200 configurations the results have almost converged. It is evident that if the chosen configurations are uncorrelated,<sup>24</sup> as is the case, the results should converge very quickly. A very small number of configurations is enough to obtain an adequate description of the average properties of the system. The results do not depend on the particular set of configurations used as long as they span the same simulation time (different sets yield interaction energies that differ by less than 0.15 kcal/mol). The same conclusions were obtained by Coutinho et al.<sup>24</sup> in a study of the solvatochromic shifts of the lowest transition of benzene in several solvents. Because of the small magnitude of the errors obtained with 100 configurations, and in order to save computational time, in the rest of the paper the mean values are calculated by averaging over 100 configurations.

The results showed larger variations when different time windows were used to calculate the averages. For instance, in methanol, variations of about 3 kcal/mol in the electrostatic interaction energy can be found if the averages are calculated taken different windows of 5.0 ps (see Table 4). For the dipole moments, the differences can be 0.2 D. As expected, these fluctuations in the mean values decrease when we increase the size of the time window. From the results given in Tables 3 and 4, one can conclude that, to obtain good averaged values, more important than including many configurations is to choose the configurations so as to span the largest possible time interval.

**TABLE 5:** Interaction Energy, Stark Component (in kcal/mol), and Dipole Moment (in debyes) of Alcohols in the Liquid State Calculated as a Mean Value ( $\langle E \rangle$  and  $\langle \mu \rangle$ ) or with the Mean Field Approximation ( $E_{\text{MFA}}$  and  $\mu_{\text{MFA}}$ )

	$\langle E \rangle$	$E_{\text{MFA}}$	$W_{\text{Stark}}$	$\langle \mu \rangle$	$\mu_{\text{MFA}}$	$\langle \mu \rangle - \mu_{\text{MFA}}$
HF						
methanol 1000	-19.0	-18.6	0.4 (2.1%)	2.46	2.45	0.01 (0.4%)
methanol 100	-19.1	-18.7	0.4 (2.1%)	2.46	2.45	0.01 (0.4%)
ethanol 100	-16.5	-16.0	0.5 (3.0%)	2.27	2.25	0.02 (0.9%)
propanol 100	-14.3	-14.0	0.3 (2.1%)	2.15	2.13	0.02 (0.9%)
MP2						
methanol 1000	-18.3	-17.9	0.4 (2.2%)			
methanol 100	-18.3	-17.9	0.4 (2.2%)			
ethanol 100	-15.8	-15.4	0.4 (2.5%)			
propanol 100	-13.7	-13.5	0.2 (1.5%)			

**TABLE 6:** Interaction Energy, Stark Component (in kcal/mol), and Dipole Moment (in debyes) of Carbonyl Compounds in Aqueous Solution Calculated as a Mean Value ( $\langle E \rangle$  and  $\langle \mu \rangle$ ) or with the Mean Field Approximation ( $E_{\text{MFA}}$  and  $\mu_{\text{MFA}}$ )

	$\langle E \rangle$	$E_{\text{MFA}}$	$W_{\text{Stark}}$	$\langle \mu \rangle$	$\mu_{\text{MFA}}$	$\langle \mu \rangle - \mu_{\text{MFA}}$
CASSCF/ANO						
formaldehyde	-9.2	-8.8	0.4 (4.3%)	2.99	2.99	0.00 (0.0%)
acetaldehyde	-8.9	-8.5	0.4 (4.5%)	3.46	3.46	0.00 (0.0%)
acetone	-21.9	-21.1	0.8 (3.6%)	4.48	4.47	0.01 (0.2%)

We shall next analyze the validity of the MFA. Table 5 gives the values of the Stark energy for the liquid alcohols. For methanol, the results are for 1000 and 100 configurations. For ethanol and propanol, only 100 configurations were used. The Stark energy ranges between 0.3 and 0.8 kcal/mol, representing less than 5% in all of the cases. The differences in dipole moments are even lower: 0.01–0.02 D representing 0.4–1%. The inclusion of the electron correlation decreases the interaction energy, a fact already found in studies performed with continuum models,<sup>25</sup> but has very little effect on the value of the Stark component, which is unchanged or slightly lower.

Table 6 gives the values calculated for the Stark component in solutions of carbonyl compounds in liquid water. The properties are calculated now at the MCSCF level and with a basis set of ANO quality. The trend is very similar to that found for the alcohols: differences less than 4% in the energies and then 0.02 D in the dipole moments.

In sum, for all the systems considered and independent of the level of calculation (HF, MP2, or MCSCF), the Stark component ranges between 0.2 and 0.8 kcal/mol, with percentages representing 1.5 and 4.5%, respectively. It is clear that the Stark energy represents a very small part of the solute–solvent interaction energy. The behavior of the MFA is even better in the prediction of the molecular properties where the errors are less than 0.02 D (1%). As a consequence, one can conclude that, for the present systems, MFA is a very good approximation that does not introduce significant errors either in the energy or in molecular properties, and that permits a notable saving of computation time.

Although the values of the solute polarizabilities given in Tables 1 and 2 range between 14 and 36 au<sup>3</sup>, the magnitude of the Stark component, Tables 3 and 4, is almost constant and independent of the system or method of calculation. From eq 11, it is clear that the Stark component is proportional to the polarizability of the solute molecule, so that in principle, an increase in polarizability should be followed by an increase in the Stark energy. However, for a given series of compounds, alcohols for instance, increasing polarizability is associated with the increase in the size of the alkyl group. Because of volume effects of these groups, the solvent reaction field decreases as

**TABLE 7:** Polarizability (in au), Electric Field Fluctuation (in au), and Comparison between the Values of the Stark Component Calculated Exactly,  $\Delta W_{\text{Stark}}$  (kcal/mol) and with the Approximate Formula, eq 11

	$\alpha$	$[\langle E^2 \rangle - \langle E \rangle^2]10^5$	$\Delta W_{\text{Stark}}$	$\Delta W_{\text{Stark}}$ (eq 11)
methanol 1000	19.45	3.87	0.44	0.20
methanol 100	19.45	4.10	0.40	0.20
ethanol	31.00	2.77	0.51	0.25
propanol	42.11	1.15	0.29	0.15
formaldehyde	14.40	6.60	0.4	0.28
acetaldehyde	25.28	3.24	0.4	0.25
acetone	35.58	4.18	0.78	0.45

also do its fluctuations. These two factors, polarizability and electric field fluctuations, seem to mutually compensate, resulting in values of the Stark component that are almost independent of the polarizability.

Last, Table 7 presents the values obtained for the Stark component when the approximate formula, eq 11, was used. In general, the approximate values are one-half the exact values. The reason is that eq 11 neglects the effect of terms of higher order than the dipole. If the Stark component obtained with the approximate formula is added to the mean field value,  $E_{\text{MFA}}$ , the difference from the true energy,  $\langle E \rangle$ , reduces to 0.35–0.1 kcal/mol, representing less than 1.5%. These results, together with the small differences found in the dipoles, 0.02 D ( $\approx 1\%$ ), are evidence for the validity of the mean field approximation in the study of molecules in solution.

## V. Conclusions

The use of the mean field approximation has become standard in solvent effect theories. Its main advantage is that it requires notably fewer quantum calculations. The cost to pay is the loss of any information associated with the fluctuations of the solvent electric field. Furthermore, it is not clear to what extent the mean energy of different solute–solvent configurations can be replaced by the energy of an averaged configuration. In this work, we have shown that, for all of the systems studied, the mean field approximation introduces only small errors in the interaction energies, and negligible differences in dipole moments. We found that, independent of the system studied, the Stark component is close to 0.5 kcal/mol, a value that can be neglected when compared with the rest of the approximations (neglect of intermolecular electron exchange, parametrization of the Lennard-Jones potential, etc.) introduced in most QM/MM methods. Furthermore, the small errors in the energy introduced by this approximation can easily be corrected through the use of the approximate formula, eq 11, whose behavior is very good, especially given that is a very easy and rapid way to obtain an estimate of the Stark component.

**Acknowledgment.** This research was sponsored by the Dirección General de Investigación Científica y Técnica (Project No. BQU2000-0243) and by the Consejería de Educación Ciencia Y Tecnología de la Junta de Extremadura (Project No. 2PR01A010).

## References and Notes

- (1) (a) Tapia, O.; Gosinski, O. *Mol. Phys.* **1975**, *29*, 1653. (b) Tapia, O.; Sussman, F.; Poulain, E. *J. Theor. Biol.* **1978**, *71*, 49.
- (2) (a) Tomasi, J.; Bonaccorsi, R.; Cammi, R.; Olivares del Valle, F. *J. J. Mol. Struct. (THEOCHEM)*, **1991**, *234*, 401. (b) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
- (3) (a) Rivail, J. L.; Rinaldi, D. In *Computational Chemistry: Review of Current Trends*; Leszczynski, J., Ed.; World Scientific Publishing: Singapore, 1995. (b) Cramer, C. J.; Truhlar, C. J. In *Reviews in*

## Theoretical Calculation of the Stark Component

J. Phys. Chem. B, Vol. 106, No. 18, 2002 4817

*Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH Publishers: New York, 1995; Vol. VI, p 1.

(4) Warshel, A. *Computer Modelling of Chemical Reactions in Enzymes and Solutions*; Wiley-Interscience Publ.: New York, 1991.

(5) (a) Ten-no, S.; Hirata, F.; Kato, S. *Chem. Phys. Lett.* **1993**, *214*, 391. (b) Ten-no, S.; Hirata, F.; Kato, S. *J. Chem. Phys.* **1994**, *100*, 7443. (c) Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *Chem. Phys.* **1995**, *240*, 199. (d) Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *J. Phys. Chem.* **1996**, *100*, 1111. (e) Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Comput. Chem.* **1997**, *18*, 1320. (f) Akiyama, R.; Hirata, F. *J. Chem. Phys.* **1998**, *108*, 4904. (f) Sato, H.; Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 9463.

(6) (a) Sánchez, M. L.; Aguilar, M. A.; Olivares del Valle, F. J. *J. Comput. Chem.* **1997**, *18*, 313. (b) Sánchez, M. L.; Aguilar, M. A.; Olivares del Valle, F. J. *J. Mol. Struct. (THEOCHEM)* **1998**, *426*, 181. (c) Sánchez, M. L.; Martín, M. E.; Aguilar, M. A.; Olivares del Valle, F. J. *Chem. Phys. Lett.* **1999**, *310*, 195. (d) Sánchez, M. L.; Martín, M. E.; Aguilar, M. A.; Olivares del Valle, F. J. *J. Comput. Chem.* **2000**, *21*, 705. (e) Martín, M. E.; Sánchez, M. L.; Olivares del Valle, F. J.; Aguilar, M. A. *J. Chem. Phys.* **2000**, *113*, 6308. (f) Martín, M. E.; Sánchez, M. L.; Aguilar, M. A.; Olivares del Valle, F. J. *J. Mol. Structure (THEOCHEM)* **2000**, *537*, 213.

(7) (a) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227. (b) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700. (c) Luzhkov, V.; Warshel, A. *J. Comput. Chem.* **1992**, *13*, 199. (d) Gao, J. *J. Phys. Chem.* **1992**, *96*, 537. (e) Wei, D.; Salahub, D. R. *Chem. Phys. Lett.* **1994**, *224*, 291. (f) Tuñón, I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-López, M. F.; Rivail, J. L. *J. Comput. Chem.* **1996**, *17*, 19. (g) Wesolowski, T. A.; Warshel, A. *J. Phys. Chem.* **1993**, *97*, 8050; **1994**, *98*, 5183. (h) Stanton, R. V.; Little, L. R.; Merz, K. M. *J. Phys. Chem.* **1995**, *99*, 17344. (i) Moriarty, N. W.; Karlström, G. *J. Phys. Chem.* **1996**, *100*, 17791. (j) Tu Y.; Laaksonen, A. *J. Chem. Phys.* **1999**, *111*, 7519.

(8) Linder, B. *Adv. Chem. Phys.* **1967**, *12*, 225.

(9) Karlstrom, G.; Halle, B. *J. Chem. Phys.* **1993**, *99*, 8056.

(10) (a) Miertus, S.; Bartos, J.; Trebaticka, M. *J. Mol. Liquids* **1987**, *33*, 139. (b) Aguilar, M. A.; Olivares del Valle, F. J. *Chem. Phys.* **1989**, *129*, 439.

(11) (a) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: London, 1987. (b) McCammon, J. A.; Harvey, J. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987.

(12) (a) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 1276. (b) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638.

(13) (a) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007. (b) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796. (c) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.

(14) Lawrence, R. B.; Sbrandberg, M. W. P. *Phys. Rev.* **1951**, *83*, 363.

(15) Kilb, R. W.; Lin, C. C.; Wilson, E. B., Jr. *J. Chem. Phys.* **1957**, *26*, 1695.

(16) Swalen, J. D.; Costain, C. C. *J. Chem. Phys.* **1959**, *31*, 1562.

(17) Widmark, P.-O.; Malmquist, P. Å; Roos, B. O. *Theor. Chem. Acta* **1990**, *77*, 291.

(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*; Gaussian, Inc.: Pittsburgh, PA, 1995.

(19) Refson, K. *Moldy User's Manual*, Rev. 2.10; University of Oxford: Oxford, U.K., 1996; ftp.earth.ox.ac.uk/pub.

(20) (a) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765. (b) Weinwe, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

(21) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(22) Berthaut, F. *J. Phys. Radium* **1952**, *13*, 499.

(23) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.

(24) (a) Coutinho, K.; de Oliveira, M. J.; Canuto, S. *Int. J. Quantum Chem.* **1998**, *66*, 249. (b) Coutinho, K.; Canuto, S.; Zerner, M. C. *J. Chem. Phys.* **2000**, *112*, 9874.

(25) Olivares del Valle, F. J.; Aguilar, M. A.; Tolosa, S. *J. Mol. Struct. (THEOCHEM)* **1993**, *279*, 223.

(26) Narten, A. H.; Sandler, S. I. *J. Chem. Phys.* **1979**, *71*, 2069.

(27) Nishikawa, K.; Iijima, T. *J. Phys. Chem.* **1990**, *94*, 6227.

(28) Maryott, A. A.; Buckley, F. U.S. National Bureau of Standards. No. 537, 1953.

(29) Applequist, J.; Fung, K. K. *J. Am. Chem. Soc.* **1972**, *94*, 2952.

(30) Johns, J. W. C.; McKellar, A. R. W. *J. Chem. Phys.* **1975**, *63*, 1682.

(31) Turner, P. H.; Cox, A. P. *Chem. Soc. J. Faraday Trans. II, Chem. Phys.* **1978**, *74*, 533.

(32) Peter, R.; Dreizler, H. Z. *Naturforsch* **1965**, *20A*, 301.

(33) Stuart, H. A. *Landolf-Börnstein Zahlenwerte und Funktionen*; Euclen, A., Hellwege, K. H., Eds.; Springer-Verlag: Berlin, 1951; Vol 1, Part 3, p 511.



## Geometry optimization of molecules in solution: Joint use of the mean field approximation and the free-energy gradient method

I. Fdez. Galván, M. L. Sánchez, M. E. Martín, F. J. Olivares del Valle, and M. A. Aguilar  
Departamento Química-Física, Universidad de Extremadura, Avda de Elvas s/n,  
06071 Badajoz, Spain

(Received 17 June 2002; accepted 7 October 2002)

The average solvent electrostatic potential/molecular dynamics (ASEP/MD) and the free-energy gradient methods are applied together with the multidimensional geometry optimization of molecules *in solution*. The systems studied were formamide in aqueous solution and water and methanol in liquid phase. The solute molecules were described through *ab initio* quantum mechanics methods (density functional theory or Møller–Plesset second order perturbation theory) while the solvent structure was obtained from Molecular Dynamics calculations. The method is very efficient; the increase in computation time is minimal with respect to previous ASEP/MD versions that worked at a fixed geometry. Despite the use of the mean field approximation in the calculation of the solvent reaction potential the agreement with previous theoretical calculations was satisfactory. Large changes were observed in the solute charge distribution induced by the solvent, and the solute polarization was accompanied by an increase in the solvent structure around the solute. © 2003 American Institute of Physics. [DOI: 10.1063/1.1525798]

### I. INTRODUCTION

The optimization of structures (minima or saddle points) *in solution* is a complicated problem because of the difficulty in properly accounting for the influence of the bulk solvent polarization and the thermal effects. In the case of the bulk solvent polarization, especially in polar liquids, it is compulsory to consider a great number of solvent molecules so as to adequately describe the electric field generated by the solvent in the volume occupied by the solute, while to include thermal effects a great number of configurations have to be considered, with the computational cost that these two factors imply. One way to simplify the problem is by having recourse to the Mean Field Approximation (MFA).<sup>1,2</sup> The introduction of the MFA has been a constant aspect of the development of solvent effect theories. The different quantum versions of continuum models,<sup>3</sup> for instance, make use of this approximation. More recently, several methods<sup>4,5</sup> have been proposed that combine MFA with a detailed description of the solvent structure obtained from an extended version of the reference interaction site model<sup>6</sup> or from Molecular Dynamics simulations.<sup>7</sup>

In previous papers<sup>5</sup> our group has developed a nonstandard Quantum Mechanics/Molecular Mechanics method (QM/MM) that makes use of the MFA. The method, denoted ASEP/MD, is based on the introduction of the Averaged Solvent Electrostatic Potential obtained from Molecular Dynamics simulation into the solute molecular Hamiltonian. As in other mean field theories the mean value of the energies of the different configurations is replaced by the energy of an average configuration that in our case is obtained from MD calculations. Its main advantage is that it permits reducing the number of quantum calculations from several thousand to 4–8 without introducing significant errors either in the energy or the molecular properties. The method has been suc-

cessfully applied to the study of liquids<sup>5(e),5(f)</sup> and to the determination of solvent shifts in the UV/VIS spectra.<sup>5(b)–5(d)</sup>

Several strategies can be used in QM/MM calculations to obtain optimal structures *in solution*. In traditional QM/MM methods,<sup>8</sup> for instance, the internal degrees of freedom are considered to be additional dynamical variables subject to thermal fluctuations. At each step, the total force (solute+solute contributions) over the solute atoms is calculated and a new geometry is obtained by integration of the equations of motion. A mean geometry can then be obtained by averaging over the different solute structures.

In a second set of methods, known as free-energy gradient methods<sup>9–11</sup> the forces felt by the solute atoms are obtained from QM/MM simulations where the solute molecule has a fixed structure. The free-energy surface (FES) is defined as the time average of the forces acting on each atom of a solute molecule over the equilibrium distribution for all solvent molecules.<sup>10</sup> From the mean gradient, a new geometry can be generated and the process is repeated until the gradient converges to a desired precision. While the main advantage of this method is that it permits obtaining both stable and transition states, it is at the expense of notably increasing the simulation time and hence the computational cost with respect to traditional QM/MM.

Because of its use of average quantities, the free-energy gradient method seems especially adequate to use together with the MFA. Their joint application permits a considerable saving of computation time. In what follows, we present a method of obtaining geometries and properties *in solution* that combines the FEG method with ASEP/MD. The method permits obtaining both stable and transition states. Because of the small number of quantum calculations that it involves, the solute description can be performed at exactly the same level as is used for *in vacuo* calculations. In this work we

restrict ourselves to the application of the ASEP/MD model to the problem of obtaining optimized structures *in solution*. We postpone the more complicated case of searching for transition states to a future paper.

The rest of the paper is organized as follows: In Sec. II the computational method is explained. Section III provides a technical details; specifically, we discuss the nature and magnitude of the gradient fluctuations and how they affect the optimization procedure. The results are discussed in Sec. IV. The final section presents the main conclusions.

## II. METHOD

The main characteristic of the ASEP/MD has been discussed elsewhere.<sup>5</sup> Here, we center on the special features of its application to the determination of critical points on potential energy surfaces.

As in traditional QM/MM methods,<sup>12</sup> in ASEP/MD, the energy and state function of the solvated solute molecules are obtained by solving the effective Schrödinger equation,

$$(\hat{H}_{QM} + \hat{H}_{QM/MM})\Psi = E\Psi. \quad (1)$$

The interaction term,  $\hat{H}_{QM/MM}$  takes the following form:

$$\hat{H}_{QM/MM} = \hat{H}_{QM/MM}^{\text{elect}} + \hat{H}_{QM/MM}^{\text{vdw}}, \quad (2)$$

$$\hat{H}_{QM/MM}^{\text{elect}} = \int dr \cdot \hat{\rho} \cdot \langle \hat{V}_s(r; \rho) \rangle, \quad (3)$$

where  $\hat{\rho}$  is the solute charge density and the brackets denote a statistical average. The term  $\langle \hat{V}_s(r; \rho) \rangle$  is the average electrostatic potential generated by the solvent at the position  $r$ , and is obtained from MD calculations where the solute molecule is represented by the charge distribution  $\rho$  and a geometry fixed during the simulation. The term  $\hat{H}_{QM/MM}^{\text{vdw}}$  is the Hamiltonian for the van der Waals interaction, in general represented by a Lennard-Jones potential. Given that the solvent structure, and hence the ASEP, is a function of the solute charge density, Eqs. (1) and (3) have to be solved iteratively. In general, only a few cycles of quantum calculation/molecular dynamics simulations are needed for convergence.

Next, we shall describe the application of the MFA to the determination of the gradient and Hessian. By way of comparison, and to determine the nature of the approximation introduced, we first present their expressions when the MFA is not used. In this case the force on the free-energy surface (FES) is<sup>9-11</sup>

$$F(r) = -\frac{\partial G(r)}{\partial r} = -\left\langle \frac{\partial V(r)}{\partial r} \right\rangle, \quad (4)$$

where  $G(r)$  is the free-energy,  $V$  is the sum of the contributions associated with the interaction with the other atoms of the solute molecule,  $V_i$ , and with the solute-solvent interaction energy,  $V_s$  and the brackets denote a statistical average. The Hessian is

$$H \left\langle \frac{\partial^2 V}{\partial r \partial r} \right\rangle - \beta \left\langle \frac{\partial V}{\partial r} \frac{\partial V^T}{\partial r} \right\rangle + \beta \left\langle \frac{\partial V}{\partial r} \right\rangle \left\langle \frac{\partial V}{\partial r} \right\rangle^T, \quad (5)$$

$$H = \left\langle \frac{\partial^2 V}{\partial r \partial r} \right\rangle - \beta [ \langle F^2 \rangle - \beta F^2 ], \quad (6)$$

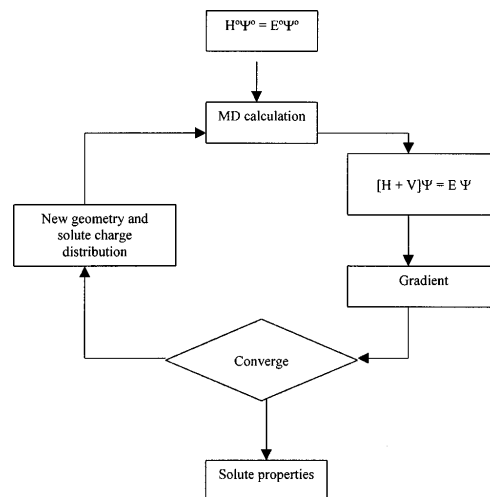


FIG. 1. Geometry optimization scheme.

where the superscript  $T$  denotes the transposition and  $\beta = 1/RT$ . The last term in Eq. (6) is related to the thermal fluctuation of the force.

Now we shall analyze how we can obtain the force and the Hessian in ASEP/MD.

Because we assume a fixed geometry and a fixed charge distribution of the solute during the simulation, the average value of the force can be replaced by the force of the mean configuration,

$$F(r) = -\frac{\partial \langle V \rangle}{\partial r}, \quad (7)$$

$$H = \frac{\partial^2 \langle V \rangle}{\partial r \partial r} - \beta \left\langle \frac{\partial V}{\partial r} \frac{\partial V^T}{\partial r} \right\rangle - \beta \frac{\partial \langle V \rangle}{\partial r} \frac{\partial \langle V \rangle^T}{\partial r}. \quad (8)$$

In this point we introduce an additional approximation: we neglect the force fluctuations. Given that the Hessian is used only to accelerate the optimization procedure, this approximation has no effect on the optimized geometries. In any case, preliminary estimations show that the errors introduced in the trace of the Hessian in the formamide-water systems when we neglect the fluctuation term is lower than 5%. With this approximation the Hessian reads

$$H = \frac{\partial^2 \langle V \rangle}{\partial r \partial r}. \quad (9)$$

Equations (7) and (9) are the equations traditionally used by mean field theories<sup>3,13</sup> and are the ones that we will follow in the present paper.

Figure 1 shows the scheme of the optimization procedure with ASEP/MD. The procedure is as follows:

- (1) We begin by obtaining through an *ab initio* calculation the *in vacuo* solute charge distribution that is used as input in the molecular dynamics calculation in the next step;



- (2) From the MD data, we obtain the ASEP,  $\langle \hat{V}_s(r) \rangle$ , and the solvent contribution to the gradient and Hessian;
- (3) The ASEP is introduced into the molecular Hamiltonian of the solute. The electronic wave function of the solute, now *in solution*, can be obtained by solving the associated effective Schrödinger equation;
- (4) The solute contributions to the gradient and Hessian are calculated and added to the solvent contribution. A new geometry is obtained;
- (5) With the new geometry the electronic wave function of the solute *in solution* is obtained and a new solute charge distribution is calculated. The new geometry and solute charge distribution are then used as an input in a new molecular dynamics calculation [step (2)]. This process is repeated until convergence in the free-energy is reached.

At each step of the self-consistent process, the solute charges used in the MD calculation were obtained by fitting the molecular electrostatic potential of the solute molecule in the presence of the solvent perturbation in the standard way. The CHELP program was used.<sup>14</sup>

The procedure for obtaining the new geometry from the old one, step (4), deserves more attention. We have checked three possibilities:

- (1) Using only the force to determine the position of the next point on the free-energy surface

$$q_{k+1} = q_k + F_k. \quad (10)$$

- (2) Finding the following point using the force and the Hessian,

$$q_{k+1} = q_k + H_k^{-1} F_k. \quad (11)$$

- (3) Performing the complete optimization of the solute geometry at each step of the cycle. This last option is the most expensive computationally, and does not seem worthwhile, especially if one takes into account that the optimized geometry will be distorted in the next cycle due to the fluctuation in the gradient.

In the three cases the forces and Hessians are calculated analytically. Compared with traditional QM/MM models the errors introduced by the above procedure are the following:

(1) The solute charge distribution obtained by solving the Schrödinger equation in the presence of a mean perturbation,  $\langle \rho \rangle$ , is different from the averaged value of the solute charge distribution obtained for each solvent configuration, i.e.,  $\bar{\rho} \neq \langle \rho \rangle$ . This difference appears because in traditional QM/MM methods the solute charge distribution is fitted at each step to the new solvent configuration, which is not permitted in MF theories. Given that  $\langle V_i \rangle$  and  $\langle V_s \rangle$  are both functions of  $\rho$ , this approximation affects the energy, the gradient, and the Hessian that are not calculated with the correct solute charge distribution. The magnitude of the errors introduced by the MFA in the calculation of the energy and the solute charge distribution has been addressed in previous papers,<sup>2,5</sup> and was evaluated to be <5% for the energy and 1% for the dipole moment. The magnitude of the errors in-

roduced by the MFA in the magnitude of the gradient is discussed below.

(2) In the calculation of the Hessian, the fluctuations of the force, the last term in Eq. (6), are neglected. Given that the Hessian is used only to accelerate the optimization procedure, this approximation, in principle, has not effect on the optimization of the geometry, and will not be considered in this paper.

### III. TECHNICAL DETAILS

In this paper, the model described above is applied to the study of several molecules in the condensed phase. We studied both liquids and solutions, and employ different levels of calculation in order to check the performance of the method in different situations. As an example of a solution, we took formamide in an aqueous solution. The basis set<sup>15</sup> used was NCO(711/411/1)H(41/1) and the solute wave function was obtained with DFT techniques. During the DFT calculation the density-gradient-corrected correlation functional proposed by Perdew<sup>16</sup> and Becke's<sup>17</sup> exchange functional were used.

As examples of liquids, we took water and methanol. In both cases, the basis set used was the aug-cc-pVDZ from Dunning *et al.*<sup>18</sup> The level of calculation was the second-order Møller–Plesset (MP2) perturbation theory for the water and DFT for the methanol. The functional was the same as used for formamide.

In the three systems the initial solute geometry was that optimized in the gas phase at the level and for the basis sets indicated above. However, in each case a different procedure was used to obtain the next solute geometry. In formamide, the geometry was completely optimized at each cycle of the ASEP/MD procedure. In methanol, the gradient and Hessian were used to determine the next geometry. Lastly, in water, only the gradient was used. All the quantum calculations were done using the GAUSSIAN98 package.<sup>19</sup>

The MD calculations were performed using the program MOLLY.<sup>20</sup> In each case, 128 (methanol) or 215 (water and formamide) molecules were simulated at a fixed intramolecular geometry by combining Lennard-Jones interatomic interactions with electrostatic interactions. In the formamide–water and water–water systems, the 214 solvent molecules were simulated by the TIP3P (Ref. 21) model at a fixed intramolecular geometry. The formamide–water potential parameters were taken from Jorgensen and Swenson.<sup>22</sup> The geometry and parameters of the classical methanol molecules were taken from Jorgensen *et al.*<sup>23</sup> Periodic boundary conditions were applied, and spherical cutoffs were used to truncate the molecular interactions at 9.0 Å. A time step of 0.5 fs was used. The electrostatic interaction was calculated with the Ewald method.<sup>22</sup> The temperature was fixed at 298 K by using the Nosé–Hoover<sup>24</sup> thermostat. Except in the cases expressly indicated, each MD calculation simulation was run for 150 000 time steps (50 000 equilibration, 100 000 production).

TABLE I. Mean value and standard deviation (in  $10^{-3}$  hartree/bohr) of the total Cartesian free energy gradient of formamide in aqueous solution, as well as their respective rms. The ASEP was calculated every 50 or 100 ps and the gradient was obtained.

		50 ps		100 ps	
		Mean	Std. dev.	Mean	Std. dev.
N1	x	-0.765	1.154	-0.706	0.173
	y	-0.892	0.372	-0.798	0.262
	z	-0.237	0.367	-0.212	0.045
H2	x	-0.054	0.625	-0.084	0.175
	y	0.616	0.348	0.632	0.132
	z	0.182	0.342	0.151	0.098
H3	x	-0.316	0.169	-0.288	.097
	y	-0.424	0.545	-0.498	0.343
	z	0.110	0.141	0.101	0.074
C4	x	-3.009	2.004	-3.040	0.261
	y	-0.682	0.557	-0.662	0.273
	z	0.003	0.046	-0.011	0.028
O5	x	-1.820	1.510	-1.695	0.496
	y	-1.389	0.727	-1.400	0.179
	z	0.164	0.105	0.245	0.136
H6	x	-0.054	0.063	-0.051	0.048
	y	0.102	0.135	0.107	0.036
	z	-0.002	0.063	0.000	0.007
rms		0.970	0.736	0.959	0.202

#### IV. RESULTS AND DISCUSSION

Prior to any geometry optimization, there are several numerical questions to consider. First, we have to determine the criterion of convergence of the gradient for optimizations *in solution* using the FEG method. The problem arises because, *in solution*, to determine the solvent contribution to the gradient we need to perform a thermodynamic average. Given the finite time of the simulations, calculations performed under the same conditions but starting from different points of the configuration space will yield different values of the gradient (and any other property, obviously). As a consequence, *in solution*, the precision of a geometry optimization is limited by the gradient fluctuations associated with the finite time of the simulation. Table I gives the mean value of the gradient on the free-energy surface of formamide in aqueous solution, and the magnitude of the largest standard deviation of the gradient (lsdg) as a function of the length of the simulation for a nonoptimized geometry. The root mean square (rms) gradient is also shown. To obtain these values we performed a MD calculation of 300 ps with a time step of 0.5 fs and fixed solute geometry. Next, averages of the gradient were calculated making use of the MFA at intervals of 50 ps and 100 ps. The lsdg can be taken as a measure of the precision at which the optimization can be performed. With 50 ps simulations, the lsdg is 0.0020 hartree/bohr. This value reduces to 0.0005 hartree/bohr when we double the simulation time. Figure 2 shows the molecular Cartesian frame of the formamide molecule.

Next, we shall analyze the magnitude of the errors that the MFA introduces in the gradient evaluation. Table II compares the gradient values for formamide in aqueous solution when the MFA is used and when it is not. From a 75 (25 + 50) ps simulation, we chose 1000 configurations. For each

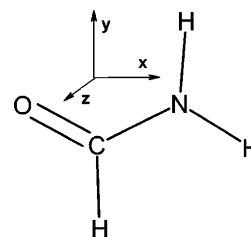


FIG. 2. Molecular Cartesian frame of formamide molecule.

configuration the gradient is obtained, and then the average value. This value is compared with the result obtained when the MFA is used. In this case, from the same set of 1000 configurations we calculate the ASEP, that is introduced into the solute molecular Hamiltonian from which the gradient can be obtained. The magnitude of the largest error is 0.00023 hartree/bohr, representing a relative error close to 3% with respect to the mean gradient. For 75 ps simulations, this error is an order of magnitude lower than the gradient fluctuations associated with the finite size of the simulations (0.0020 hartree/bohr), and we can expect that the error in the gradient had a negligible effect on the optimized geometry. However, for 150 ps simulations or longer the error associated with the MFA can become dominant, and sets a limit to the precision at which an optimization can be performed when the MFA is used.

Next, we shall consider the results for the optimization of several molecules *in solution*. The first system we studied was formamide in aqueous solution. We chose this system because the *in solution* geometry for this system had previously been considered by different authors using both continuum models and QM/MM methods.<sup>8</sup> Furthermore, it can be seen as the simplest example of a peptide bond. To check

TABLE II. Total Cartesian gradient of the free energy (in  $10^{-3}$  hartree/bohr) of a molecule of formamide in aqueous solution. The Lennard-Jones contribution was not included.

		Mean of 1000 configurations	Average configuration
N1	x	12.914	12.827
	y	3.858	3.831
	z	0.084	0.085
H2	x	-3.365	-3.132
	y	3.049	2.928
	z	0.053	0.049
H3	x	-1.116	-1.081
	y	-5.347	-5.081
	z	-0.006	-0.004
C4	x	-26.864	-26.816
	y	5.473	5.487
	z	-0.002	-0.010
O5	x	18.255	17.862
	y	-10.052	-9.788
	z	-0.041	-0.026
H6	x	1.567	1.571
	y	-2.136	-2.240
	z	-0.004	0.002
rms		8.898	8.807



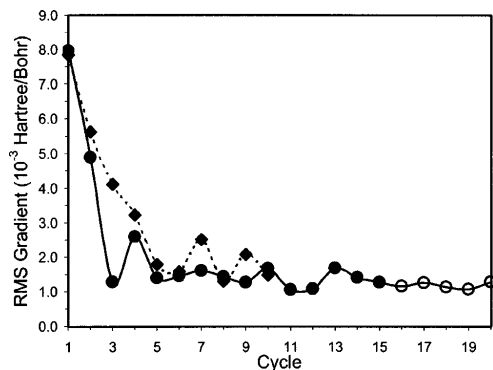


FIG. 3. Root mean square change in free-energy surface of formamide in aqueous solution. 50 ps. (filled circles), 150 ps. (open circles), and 25 ps. (dotted line, diamonds) production simulations.

the effect of the gradient fluctuation on the optimization procedure, we performed many more cycles than strictly necessary. Figure 3 displays the root mean square change of the gradient on the free-energy surface during the optimization of formamide in aqueous solution obtained with 75 (25+50) ps simulations. Convergence is reached in 5–6 cycles. From this point, the gradient begins to fluctuate around a rms gradient value of about 0.0015 hartree/bohr. This value is somewhat lower than the values obtained by other authors who also used the FEG method. For instance, Okuyama-Yoshida *et al.*<sup>10</sup> in the optimization of glycine obtain 0.0025 hartree/bohr, and Hirao *et al.*<sup>11</sup> in the transition state of the Menshutkin reaction between ammonia and methyl chloride obtain 0.010 hartree/bohr. The larger values obtained by these authors are probably related to the shorter duration of the simulations, 15+10 ps and 10+10 ps, respectively. Our gradient value is one order of magnitude larger than the thresholds used for *in vacuo* calculations (in Gaussian,<sup>19</sup> for instance, the maximum and rms gradients are 0.00045 and 0.0003 hartree/bohr, respectively). Contrary to what was expected,

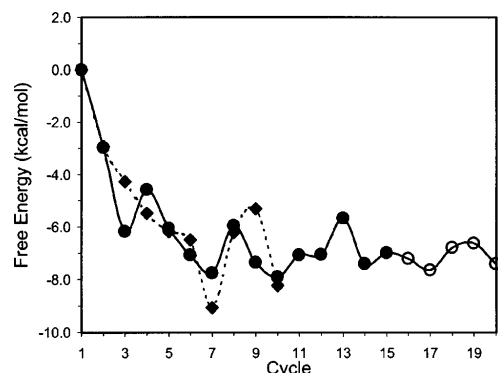


FIG. 4. Polarization free-energy change of formamide in aqueous solution during the optimization procedure. Symbols: see Fig. 3.

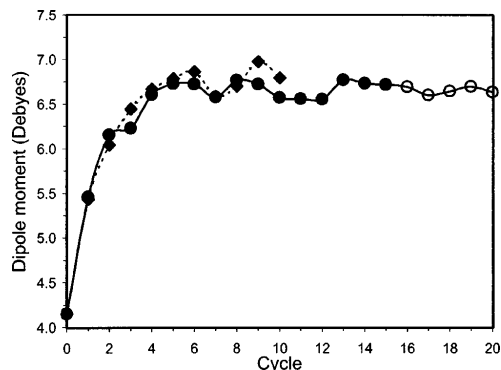


FIG. 5. Dipole moment change of formamide in aqueous solution during the optimization procedure. Symbols: see Fig. 3.

the rms gradient value, Fig. 3, is not reduced when the size of the simulation is increased. Only its fluctuations decrease. This fact is probably an artifact of the ASEP/MD method, related to the different representations of the solute charge distribution during the simulation (classical, through the point charges) and the optimization (quantum, through the wave function).

Figure 4 shows the evolution of the solute polarization free-energy during the optimization procedure. As before, convergence is reached in 5–6 cycles, when  $\Delta G$  begins to fluctuate. The free energies were calculated by the free-energy perturbation method.<sup>25</sup> As expected, the size of the fluctuations decreases with increasing length of the simulation time, but the average value is almost the same in all the cases ( $-7.1 \pm 1.4$  kcal/mol,  $-7.0 \pm 0.7$  kcal/mol, and  $-7.1 \pm 0.4$  kcal/mol for the 25, 50, and 150 ps simulations, respectively). Only for the longer simulations are the fluctuation less than the fluctuation due to thermal effects,  $k_B T$ , which is of the order of 0.6 kcal/mol at 300 K. The same oscillatory behavior is observed for the dipole moment, see Fig. 5. The averaged induced solute dipole moment is 2.5 D, represent-

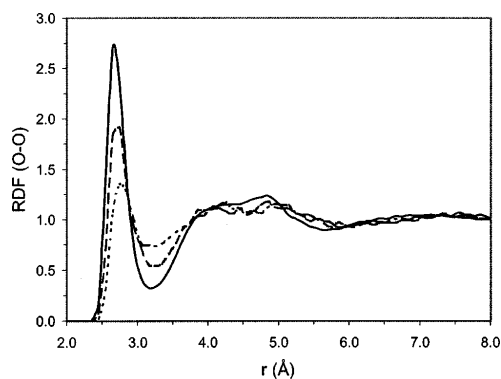


FIG. 6. O(formamide)-O(water) radial distribution function of methanol for the first (dotted line), second (dashed line), and last (full line) cycles of the ASEP/MD procedure.

TABLE III. Optimization of the geometry of formamide. Cycle 0 is the *in vacuo* calculation. The average of the converged cycles and the difference between the *in vacuo* and *in solution* geometries are also shown. For comparison the results obtained in Ref. 8 are displayed as DFT/MM. Distances in angstrom, angles in degrees.

Distance or angle	ASEP/MD (50 ps)							
	0 ( <i>in vacuo</i> )	1	2	3	4	5	10	15
N1-H2	1.017	1.022	1.025	1.028	1.030	1.031	1.030	1.029
N1-H3	1.020	1.025	1.029	1.029	1.030	1.032	1.032	1.033
N1-C4	1.359	1.343	1.335	1.334	1.331	1.329	1.331	1.329
C4-O5	1.229	1.240	1.248	1.248	1.253	1.254	1.252	1.254
C4-H6	1.114	1.109	1.107	1.106	1.106	1.105	1.106	1.105
H2-N1-H3	117.7	117.7	117.6	117.7	117.5	117.8	117.7	117.4
H2-N1-C4	122.6	121.9	121.5	121.6	121.1	120.6	121.0	120.6
N1-C4-O5	125.4	125.5	125.5	125.1	125.6	125.6	125.3	125.6
N1-C4-H6	112.1	113.0	113.5	113.8	113.7	113.8	113.9	113.8

Distance or angle	ASEP/MD				DFT/MM		
	50 ps		150 ps		<i>in vacuo</i>	<i>in solution</i>	Variation
	Average	Variation	Average	Variation			
N1-H2	1.029	0.013	1.030	0.013	1.016	1.040	0.024
N1-H3	1.033	0.013	1.033	0.013	1.018	1.041	0.023
N1-C4	1.329	-0.029	1.330	-0.029	1.366	1.338	-0.028
C4-O5	1.253	0.024	1.253	0.024	1.227	1.262	0.035
C4-H6	1.106	-0.008	1.106	-0.008	1.114	1.112	-0.002
H2-N1-H3	117.6	-0.1	117.5	-0.1	119.3	119.7	0.4
H2-N1-C4	120.8	-1.9	120.8	-1.9	121.7	119.8	-1.9
N1-C4-O5	125.5	0.2	125.6	0.2	124.4	124.6	0.2
N1-C4-H6	113.8	1.8	113.8	1.7	112.5	114.2	1.7

ing an increase of almost 60% with respect to the *in vacuo* value. The polarization of the solute is accompanied by a parallel increase in the structure of the solvent around the solute. This is clearly observed in the evolution of the oxygen-oxygen radial distribution function with the number of cycles of the self-consistent process. As polarization progresses, the height of the first peak increases and moves toward shorter distances. During this process, the solute charge distribution and the solvent structure become mutually equilibrated (Fig. 6).

The computed geometries in gas phase and solution are listed in Table III. The final *in solution* geometries are average values over the different simulations (we give the averages for 50 ps and 150 ps simulations). For comparison, we also include results obtained using the DFT/MD method developed at Nancy. The first conclusion is that the optimized geometry depends very little on the size of the simulation. The averages obtained with the 50 ps and 150 ps simulations are almost identical. Furthermore, despite the gradient fluctuations associated with the simulation size, the solute geometries are very stable: the standard deviations are  $<0.001$  Å for the bond lengths and  $0.2^\circ$  for the bond angles. Although our values are not directly comparable with the results obtained by Chalmet and Ruiz-López<sup>8</sup> due to small differences in the basis set and the density functional, they are very similar qualitatively and quantitatively: the solvent induces a decrease in the CN distance, an increase in the CO distances, and negligible variations in the bond angles. These results can be interpreted simply by the stabilization of the zwitterionic electronic configuration by the electrostatic solute-

solvent interactions, as was confirmed by the behavior of the net atomic charges obtained by a CHELP (Ref. 14) population analysis. As expected, the electronic density of the N atom decreases from gas phase to solution, whereas it increases on the oxygen atom. The strong polarization of the formamide molecule in aqueous solution is manifested in the change of the dipole moment from 4.1 D *in vacuo* to 6.6 D *in solution*.

Next, we studied the water molecule. In this case, in order to check the performance of the proposed method when different calculation levels are used, the calculations were performed at the MP2 level with an aug-cc-pVDZ basis

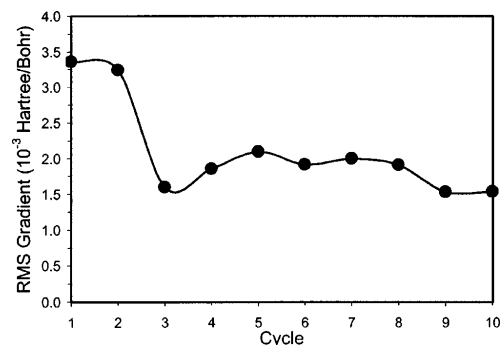


FIG. 7. Root mean square change in free-energy surface of water.



FIG. 8. Polarization free-energy change of a liquid water molecule during the optimization procedure.

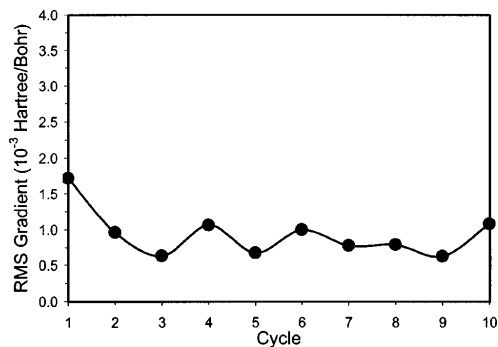


FIG. 10. Root mean square change in free-energy surface of methanol.

set.<sup>17</sup> As was the case for formamide, convergence is reached in about 4–5 cycles of the ASEP/MD procedure (see Fig. 7). The rms gradient value is also very similar, about 0.0015 hartree/bohr. The fluctuations in the free-energy, Fig. 8, are about 0.2 kcal/mol, clearly less than the fluctuations due to thermal effects. The water molecule undergoes a strong polarization during the solution procedure. The induced dipole moment, 0.77 D, compares well with the results obtained by other authors,<sup>26</sup> and with the estimated experimental<sup>27</sup> value, 0.70–1.0 D. The geometry changes are displayed in Fig. 9. The OH distance increases by  $0.015 \pm 0.002$  Å. The bond angle does not change appreciably.

Finally, we studied liquid methanol. The basis set was of aug-cc-pVDZ (Ref. 17) quality. The calculations were performed at the DFT level. Convergence was reached in 3–5 cycles, Fig. 10,<sup>13</sup> and the rms gradient was somewhat lower in this case, 0.0008 hartree/bohr, than for formamide and water. The free-energy fluctuations, Fig. 11, are however, similar to those of water, about 0.3 kcal/mol. The *in solution* dipole moment was 2.30 D and the induced dipole moment 0.55 D. The methanol induced dipole moment is clearly lower than that obtained for water. This is so even though water's polarizability is almost half that of methanol. The explanation is to be found in the number of hydrogen bonds that each system forms. Thus, while methanol can form two hydrogen bonds, this number increases to four in the case of water. The changes in the geometry originated by the solvent are also less in methanol than in water. Only the OH bond is increased by about 0.008 Å. The rest of the molecule, Fig. 12, does not change appreciably. As was the case for the water and formamide, the fluctuations in the gradient have

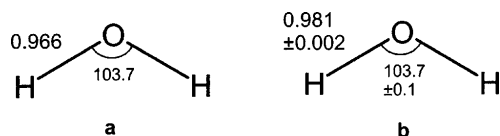


FIG. 9. Optimized geometries of water in the gas phase (a) and *in solution* (b).

no appreciable effect on the geometry: the standard deviations are 0.001 Å for the bonds and  $0.1^\circ$ – $0.4^\circ$  for the angles.

The number of cycles necessary to reach convergence is very similar in the three systems considered, although in each case a different procedure was used to calculate the next solute geometry in the ASEP/MD cycle (calculation of the gradient, gradient and Hessian, or of the complete optimized geometry). Given that the computational times are very different depending on the method used, the best option is to use the simplest procedure, i.e., to calculate the new geometry by using only the gradient.

## V. SUMMARY

In this article we have proposed a method for the optimization of molecules *in solution* that makes use of the mean field approximation and the free-energy gradient method. The method yields optimized stable geometries *in solution* in a very efficient way: the increase in computational time is minimal with respect to previous versions of the ASEP/MD that worked at a fixed geometry. As a new ingredient, we need only calculate the gradient, or the gradient and the Hessian, at each cycle of the ASEP/MD procedure. In general 5–6 cycles were enough to reach convergence. We have also

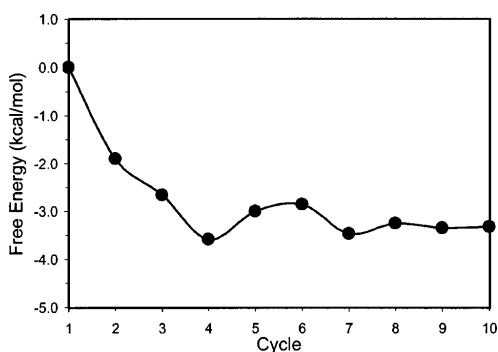


FIG. 11. Polarization free-energy change of a liquid methanol molecule during the optimization procedure.

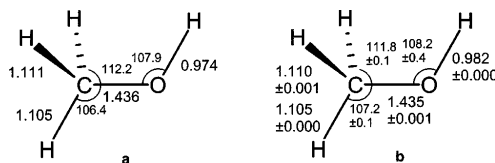


FIG. 12. Optimized geometries of methanol in the gas phase (a) and in solution (b).

shown that the fluctuation in the gradient and free-energy are related to the time of simulation. Except for formamide, where it was necessary to perform longer simulations (150 ps), for water and methanol simulations of 75 ps yielded to free-energy fluctuations that were lower than the thermal contribution. In the three molecules studied, and despite the gradient and free-energy fluctuations, the geometries obtained were very stable. The standard deviations were  $<0.001$  Å for the bond length and  $<0.2^\circ$  for the bond angles.

It is interesting to note that, for the molecules considered, the solvent only induces small changes in the solute geometry. In fact, the influence that the basis set or the level of calculation has on the geometry is larger than the solvent influence. For instance, in the water system, the *in vacuo* OH bond distance varied by about 0.022 Å from the HF to the MP2 level, and by 0.031 from the HF to the DFT level. A similar behavior was presented by the other two systems. The solvent, however, induced large changes in the solute charge distributions, which became strongly polarized. Increases in the dipole moment between 30%–60% were found. At the same time, the polarization of the solute was accompanied by an increase in the structure of the solvent around the solute. The first peak of the radial distribution function shifted to shorter distances and its height increased by several units.

In the present ASEP/MD version there are two possible sources of error: the use of the mean field approximation and the different representations of the electrostatic solute charge distribution during the simulation (classical) and the optimization (quantum mechanical). In this and a previous paper, we have shown that the MFA only introduces small errors in the magnitudes evaluated (dipole moment, interaction energies, gradients, etc.) while it allows one to greatly reduce the number of quantum calculations to perform. The errors associated with the classical representation of the solute during the MD calculation can be reduced by improving the solute charge representation through the inclusion of additional charges on the solute molecule. We are working in this direction. In any case, these errors seem negligible when compared with the errors introduced by the rest of the approximations used in most QM/MM methods (neglect of intermolecular electron exchange, parametrization of the Lennard-Jones potential, etc.)

#### ACKNOWLEDGMENTS

This research was sponsored by the Dirección General de Investigación Científica y Técnica (BQU2000-0243) and by the Consejería de Educación y Juventud de la Junta de Extremadura (Project No. 2PR01A010).

<sup>1</sup>O. Tapia, in *Theoretical Treatment of Large Molecules and Their Interactions*, edited by Z. B. Maksic (Springer-Verlag, Berlin, 1991), Vol. 4, p. 435.

<sup>2</sup>M. L. Sánchez, M. E. Martín, I. Fdez. Galván, F. J. Olivares de Valle, and M. A. Aguilar, *J. Phys. Chem.* **106**, 4813 (2002).

<sup>3</sup>(a) J. Tomasi, R. Bonaccorsi, R. Cammi, and F. J. Olivares de Valle, *J. Mol. Struct.: THEOCHEM* **234**, 401 (1991); (b) J. Tomasi and M. Persico, *Chem. Rev.* **94**, 2027 (1994); (c) J. L. Rivail and D. Rinaldi, in *Computational Chemistry: Review of Current Trends*, edited by J. Leszczynski (World Scientific, Singapore, 1995); (d) C. J. Cramer and C. J. Truhlar, in *Reviews in Computational Chemistry*, edited by K. B. Kipkowitz and D. B. Boyd (VCH, New York, 1995), Vol. VI, p. 1.

<sup>4</sup>(a) S. Ten-no, F. Hirata, and S. Kato, *Chem. Phys. Lett.* **214**, 391 (1993); (b) S. Ten-no, F. Hirata, and S. Kato, *J. Chem. Phys.* **100**, 7443 (1994); (c) M. Kawata, S. Ten-no, S. Kato, and F. Hirata, *Chem. Phys.* **240**, 199 (1995); (d) M. Kawata, S. Ten-no, S. Kato, and F. Hirata, *J. Phys. Chem.* **100**, 1111 (1996); (e) M. Kinoshita, Y. Okamoto, and F. Hirata, *J. Comput. Chem.* **18**, 1320 (1997); (f) R. Akiyama and F. Hirata, *J. Chem. Phys.* **108**, 4904 (1998); (g) H. Sato, A. Kovalenco, and F. Hirata, *ibid.* **112**, 9463 (2000).

<sup>5</sup>(a) M. L. Sánchez, M. A. Aguilar, and F. J. Olivares del Valle, *J. Comput. Chem.* **18**, 313 (1997); (b) M. L. Sánchez, M. E. Martín, M. A. Aguilar, and F. J. Olivares del Valle, *Chem. Phys. Lett.* **310**, 195 (1999); (c) M. E. Martín, M. L. Sánchez, J. Olivares del Valle, and M. A. Aguilar, *J. Chem. Phys.* **113**, 6308 (2000); (d) M. E. Martín, M. L. Sánchez, M. A. Aguilar, and F. J. Olivares de Valle, *J. Mol. Struct.: THEOCHEM* **537**, 213 (2001); (e) M. L. Sánchez, M. E. Martín, M. A. Aguilar, and F. J. Olivares del Valle, *J. Comput. Chem.* **21**, 705 (2000); (f) M. E. Martín, M. L. Sánchez, F. J. Olivares del Valle, and M. A. Aguilar, *J. Chem. Phys.* **116**, 1613 (2002).

<sup>6</sup>(a) F. Hirata and P. J. Rossky, *Chem. Phys. Lett.* **83**, 329 (1981); (b) F. Hirata, P. J. Rossky, and B. M. Pettitt, *J. Chem. Phys.* **78**, 4133 (1983); (c) D. Chandler and H. C. Andersen, *ibid.* **57**, 1930 (1972).

<sup>7</sup>(a) M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, London, 1987); (b) J. A. McCammon and J. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge, 1987).

<sup>8</sup>S. Chalmet and M. F. Ruiz-López, *J. Chem. Phys.* **111**, 1117 (1999).

<sup>9</sup>N. Okuyama-Yoshida, M. Nagaoka, and T. Yamabe, *Int. J. Quantum Chem.* **70**, 95 (1998).

<sup>10</sup>N. Okuyama-Yoshida, K. Kataoka, M. Nagaoka, and T. Yamabe, *J. Chem. Phys.* **113**, 3519 (2000).

<sup>11</sup>H. Hirao, Y. Nagae, and M. Nagaoka, *Chem. Phys. Lett.* **348**, 350 (2001).

<sup>12</sup>(a) A. Warshel and M. Levitt, *J. Mol. Biol.* **103**, 227 (1976); (b) U. C. Singh and P. A. Kollman, *J. Comput. Chem.* **7**, 718 (1986); (c) M. J. Field, P. A. Bash, and M. Karplus, *ibid.* **11**, 700 (1990); (d) V. Luzhkov and A. Warshel, *ibid.* **13**, 199 (1992); (e) J. Ga. in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (VCH, New York, 1996), Vol. 7, p. 119.

<sup>13</sup>(a) R. Cammi and J. Tomasi, *J. Chem. Phys.* **101**, 3888 (1994); (b) E. Cancès, B. Mennucci, and J. Tomasi, *ibid.* **109**, 260 (1998); (c) V. Barone, M. Cossi, and J. Tomasi, *J. Comput. Chem.* **19**, 404 (1998).

<sup>14</sup>L. E. Chirlian and M. M. Francl, *J. Comput. Chem.* **8**, 894 (1987).

<sup>15</sup>St. Amant and D. R. Salahub, *Chem. Phys. Lett.* **169**, 387 (1990).

<sup>16</sup>J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).

<sup>17</sup>A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).

<sup>18</sup>(a) T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989); (b) R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *ibid.* **96**, 6796 (1992); (c) D. E. Woon and T. H. Dunning, *ibid.* **98**, 1358 (1993).

<sup>19</sup>GAUSSIAN 98, M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, Gaussian, Inc., Pittsburgh, Pennsylvania, 1998.

<sup>20</sup>K. Refson, *MOLDY User's Manual Revision 2.10* (University of Oxford Press, Oxford, 1996) ftp.earth.ox.ac.uk/pub

<sup>21</sup>W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).

<sup>22</sup>W. L. Jorgensen and C. J. Swenson, *J. Am. Chem. Soc.* **107**, 569 (1985).

<sup>23</sup>(a) W. L. Jorgensen, *J. Phys. Chem.* **90**, 1276 (1986); (b) L. Jorgensen, J. D. Madura, and C. J. Swenson, *J. Am. Chem. Soc.* **106**, 6638 (1984).

<sup>24</sup>W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).

<sup>25</sup>R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954); U. C. Singh, F. K. Brown, P. A. Bash, and P. A. Kollman, *J. Am. Chem. Soc.* **109**, 1607 (1987).

<sup>26</sup>(a) I. Tuñón, M. T. C. Martins-Costa, C. Millot, M. F. Ruiz-López, and

J.L. Rivail, *J. Comput. Chem.* **17**, 19 (1996); (b) P. Ahlström, A. Wallqvist, S. Engström, and B. Jönsson, *Mol. Phys.* **68**, 563 (1989); (c) K. Laasonen, M. Sprik, M. Parrinello, and R. Car, *J. Chem. Phys.* **99**, 9080 (1993); (d) G. Jansen, F. Colonna, and J. G. Angyán, *Int. J. Quantum Chem.* **58**, 251 (1996); (e) P. L. Silvestrelli and M. Parrinello, *J. Chem. Phys.* **111**, 3572 (1999).

<sup>27</sup> (a) D. Eisenberg and W. Kauzmann, *The Structure and Properties of Water* (Oxford University Press, New York, 1969); (b) C. A. Coulson and D. Eisenberg, *Proc. R. Soc. London, Ser. A* **291**, 445 (1966); (c) E. Whalley, *Chem. Phys. Lett.* **53**, 449 (1978); (d) Y. S. Badyal, M. L. Saboungi, D. L. Price, S. D. Shasri, D. R. Haefner, and A. K. Soper, *J. Chem. Phys.* **12**, 9206 (2000).





ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)



Computer Physics Communications 155 (2003) 244–259

Computer Physics  
Communications

[www.elsevier.com/locate/cpc](http://www.elsevier.com/locate/cpc)

## ASEP/MD: A program for the calculation of solvent effects combining QM/MM methods and the mean field approximation <sup>☆</sup>

I. Fdez Galván, M.L. Sánchez, M.E. Martín, F.J. Olivares del Valle, M.A. Aguilar <sup>\*</sup>

*Dpto Química Física, Univ. de Extremadura, Avda de Elvas s/n, 06071 Badajoz, Spain*

Received 28 March 2003

### Abstract

ASEP/MD is a computer program designed to implement the Averaged Solvent Electrostatic Potential/Molecular Dynamics (ASEP/MD) method developed by our group. It can be used for the study of solvent effects and properties of molecules in their liquid state or in solution. It is written in the FORTRAN90 programming language, and should be easy to follow, understand, maintain and modify. Given the nature of the ASEP/MD method, external programs are needed for the quantum calculations and molecular dynamics simulations. The present version of ASEP/MD includes interface routines for the GAUSSIAN package, HONDO, and MOLDY, but adding support for other programs is straightforward. This article describes the program and its usage.

### Program summary

*Title of program:* ASEP/MD

*Catalogue identifier:* ADSF

*Program Summary URL:* <http://cpc.cs.qub.ac.uk/summaries/ADSF>

*Program obtainable from:* CPC Program Library, Queen's University of Belfast, N. Ireland

*Computer for which the program is designed:* it has been tested on Intel-based PC and Sun

*Operating systems under which the program has been tested:* Red Hat Linux 7.2 and SunOS 5.6

*Programming language used:* FORTRAN90

*Memory required to execute with typical data:* greatly depends on the system

*No. of processors used:* 1

*Has the code been vectorized or parallelized?:* no

*No. of bytes in distributed program, including test data, etc.:* 44 544

*Distribution format:* tar gzip file

*Keywords:* Solvent effects, QM/MM methods, mean field approximation, geometry optimization

*Nature of physical problem:* The study of molecules in solution with quantum methods is a difficult task because of the large number of molecules and configurations that must be taken into account. The quantum mechanics/molecular mechanics methods proposed to date either require massive computational power or oversimplify the solute quantum description.

<sup>☆</sup> This paper and its associated computer program are available via the Computer Physics Communications homepage on ScienceDirect (<http://www.sciencedirect.com/science/journal/00104655>).

<sup>\*</sup> Corresponding author.

*E-mail address:* [maguilar@unex.es](mailto:maguilar@unex.es) (M.A. Aguilar).

*Method of solution:* A non-traditional QM/MM method based on the mean field approximation was developed where a classical molecular dynamics simulation is coupled with a quantum calculation. The average electrostatic potential generated by the solvent over the solute is calculated from the simulation and introduced into the quantum calculation as an external field. This process can be performed iteratively. Standard external programs are used for the molecular dynamics simulations and for the quantum calculations. The present program acts as an interface and controls the flow of the calculation.

*Restrictions on the complexity of the problem:* At present, only pure liquids and binary dilute solutions (a single solute molecule) can be studied. For the molecular dynamics only MOLLY is implemented, while GAUSSIAN and HONDO are available for the quantum calculations. Restrictions of the aforementioned programs apply.

*Typical running time:* Running time depends on the nature of the chemical system and the options passed to the external programs, which are usually by far the longest part of the calculations.

*Unusual features of the program:* Uses SYSTEM and GETARG calls.

© 2003 Elsevier B.V. All rights reserved.

## 1. Introduction

Quantum Mechanics/Molecular Mechanics (QM/MM) methods [1] are now widely used in the study of molecules in solution. The main advantage of these methods is that they combine a quantum description of the solute, allowing chemical processes to be studied, with a detailed description of the solvent, obtained from simulation techniques. In most QM/MM methods the solute Schrödinger equation has to be solved for each solvent configuration, which implies several thousand quantum calculations. This imposes a limitation on the quality of the quantum description of the solvent (basis set and calculation level) and on the significance of the results (number of configurations considered).

In previous papers [2] we have developed a non-traditional QM/MM method that makes use of the mean field approximation [3] (MFA). In this approximation, the average value of the energies of the different solute-solvent configurations is replaced by the energy of the average configuration. Our method is based on the calculation of the Average Solvent Electrostatic Potential (ASEP) from Molecular Dynamics data (MD); this average potential is introduced into the molecular Hamiltonian of the solute and the Schrödinger equation is solved. This approximation, named ASEP/MD, reduces drastically the number of quantum calculations from several thousands to half a dozen, and introduces no significant inaccuracies [3] in the solute-solvent interaction energy or the solute dipole moment. This reduced number of quantum calculations permits one to use higher quality methods, while having an adequate sampling of the solvent configurations through the molecular dynamics simulations.

This paper presents of the computer program developed to implement our method. The quantum calculations and the molecular dynamics simulations are performed by external programs, intentionally left out of the present implementation. This dependence on external programs allows the users to employ whatever program they are accustomed to or that is best suited to their needs, although specific subroutines for the purpose may have to be written.

## 2. Method

The main characteristics of the ASEP/MD method have been discussed elsewhere [2]. The following is a short description. As in traditional QM/MM methods, in ASEP/MD the energy and wavefunction of the solvated solute molecule are obtained by solving the effective Schrödinger equation

$$(\hat{H}_{\text{QM}} + \hat{H}_{\text{QM/MM}})|\Psi\rangle = E|\Psi\rangle. \quad (1)$$

The interaction term,  $\hat{H}_{\text{QM/MM}}$ , takes the following form:

$$\hat{H}_{\text{QM/MM}} = \hat{H}_{\text{QM/MM}}^{\text{elect}} + \hat{H}_{\text{QM/MM}}^{\text{vdw}}, \quad (2)$$

$$\hat{H}_{\text{QM/MM}}^{\text{elect}} = \int \rho(V_s(r; \rho)) dr, \quad (3)$$



where  $\rho$  is the solute charge density, obtained from a quantum calculation, and the term  $\langle V_s(r; \rho) \rangle$  is the average electrostatic potential generated by the solvent (ASEP) at the position  $r$ , and is obtained from MD calculations where the solute molecule has fixed geometry and charge distribution  $\rho$ —the angle brackets denote a statistical average. The term  $\widehat{H}_{\text{QM/MM}}^{\text{vdw}}$  is the Hamiltonian for the van der Waals interaction, usually represented by a Lennard-Jones potential. Since the solvent structure, and hence the ASEP, is a function of the solute charge density, Eqs. (1) and (3) have to be solved iteratively. In general, only a few cycles of quantum calculation/molecular dynamics simulations are needed for convergence.

In order to facilitate its implementation in a standard quantum program package, the ASEP is represented by means of a set of point charges. The method used to obtain the ASEP and the point charges that approximate it is as follows. First a three-dimensional grid of points is created inside the volume occupied by the solute. Then, for every solvent configuration considered, the electrostatic potential created by the solvent at these points is calculated and averaged over all the configurations. The solvent charges inside a certain cut-off radius around the solute are explicitly kept (their values divided by the total number of configurations considered) and, to avoid having too many charges, they are grouped together if the distance between them is less than a certain value. The electrostatic potential created by these explicit charges is subtracted from the one calculated before. In this way one obtains the contribution of the solvent molecules placed outside the cut-off radius. We named this contribution ASEPI. Finally, an external set of point charges, arranged in two spherical shells, is obtained such that the potential generated by them is the closest to the calculated ASEPI. The positions of these charges are set previously and their values are obtained through a least squares fit, so that

$$\sum_i (V'_i - V_i)^2, \tag{4}$$

$$V'_i = \sum_a \frac{q_a}{r_{ai}} \tag{5}$$

is minimized, where  $V_i$  is the ASEPI calculated at point  $i$  and  $V'_i$  is the potential generated by the fitted charges at the same point,  $q_a$  is the value of the charge  $a$  and  $r_{ai}$  is the distance between charge  $a$  and point  $i$ . These point charges are then introduced into the solute Hamiltonian. In this way, the solvent is represented by two sets of point charges, one closer to the solute and keeping some information about the solvent structure and a second, farther from the solute, completing the fit to the ASEP. In the current implementation the molecular dynamics is performed with the MOLDY program [4] and the quantum calculation with the GAUSSIAN [5] or HONDO [6] packages.

In some cases it is desirable to make a calculation with polarizable solvent, for example to study electron transitions in the solute, which are too rapid to allow a reorganization of the solvent nuclei but permit the polarization of the electron clouds of the solvent molecules. In this case, instead of performing a full polarizable molecular dynamics—which would be too expensive—the polarization is calculated after the simulation is done, using the obtained solute–solvent configurations. In each configuration, the induced dipole moment is calculated for every solvent molecule,

$$\mu'_i = \alpha_i E_i, \tag{6}$$

where  $\mu'_i$  is the induced dipole moment of molecule  $i$ ,  $\alpha_i$  its polarizability tensor, and  $E_i$  the *total* electric field in the center of mass of the molecule  $i$ , including contributions from the solute and the charges and induced dipoles of the other solvent molecules. Since the dipole moment on each molecule depends on the dipole moments on the rest of molecules, the calculation is done iteratively until convergence. Once the dipole moments have been obtained, the ASEP is calculated as above.

To perform a geometry optimization of the solute with the ASEP/MD method [7] we use the free energy gradient (FEG) method [8], in which the forces on the nuclei are

$$F(x) = -\frac{\partial G(x)}{\partial x} = -\left\langle \frac{\partial V(x)}{\partial x} \right\rangle \tag{7}$$

and the Hessian

$$H(x) = \frac{\partial F(x)}{\partial x} = \left\langle \frac{\partial^2 V(x)}{\partial x^2} \right\rangle - \beta (\langle F(x)^2 \rangle - \langle F(x) \rangle^2), \quad (8)$$

where  $G(x)$  is the free energy associated with the solute geometry  $x$ ,  $V(r)$  the potential energy of the solute including the interaction with the solvent molecules, and  $\beta = 1/RT$ . The last term in Eq. (8) is related to the thermal fluctuation of the force and, since the Hessian is not essential for geometry optimization, we neglect it and assume

$$F(x) = -\frac{\partial \langle V(x) \rangle}{\partial x}, \quad (9)$$

$$H(x) = -\frac{\partial^2 \langle V(x) \rangle}{\partial x^2}, \quad (10)$$

where we replace the statistical average of the gradient by the gradient of the “average configuration” given by the ASEP, and similarly for the Hessian. Once the gradient and Hessian have been obtained, a standard geometry optimization (steepest descent, Newton-like methods, RFO) is performed in each cycle of the process.

### 3. Program description

The ASEP/MD program is written with readability in mind. Since most processor and memory consuming tasks are those performed by the external programs (quantum calculations and molecular dynamics simulations), code optimization was considered as only secondary. Once compiled and linked, the program is available as a single executable file. To launch ASEP/MD the user needs to provide a general input file and “template” files, which will be used as input for the external programs (obviously, these external programs must have been installed previously). The external programs currently supported are GAUSSIAN and HONDO for quantum calculations and MOLDFY for molecular dynamics simulations.

Fig. 1 shows the simplified flow chart of the program. The most important steps are those in bold face: molecular dynamics simulation, calculation of the ASEP, and the quantum calculation.

In a typical run, the first thing the program does is to read the values of some variables from the input file supplied by the user. It then reads the system definition from the molecular dynamics template (nature and number of solvent and solute molecules, interaction potential, etc.). With this information, and if needed, the external quantum calculation program is called to perform a calculation of the isolated solute molecule, using the options given in the quantum calculation template (method, basis set, etc.). The output of this calculation is read to extract the energy, dipole moment and atomic charges of the solute *in vacuo*.

Now the main ASEP/MD cycle is carried out. The solute and solvent geometry and atomic charges are inserted into the molecular dynamics input file and the appropriate program is called to perform the simulation. Next, a set of solvent–solute configurations is read from the output of the molecular dynamics program and the average electrostatic potential created by the solvent over the solute (ASEP) is calculated from them, as well as several interaction energies. The ASEP is then approximated by a set of point charges. A new quantum calculation, this time with the external point charges representing the solvent, is performed. The output gives again the energy, dipole moment, and atomic charges of the solute. A check for convergence is made and, if it has not been reached, the solute charges are inserted again into the molecular dynamics input file and the main cycle starts anew.

So far, the program flow has been described when a basic calculation is requested. There are two other main types of calculations which require some modifications of this scheme: polarizable solvent and geometry optimization.

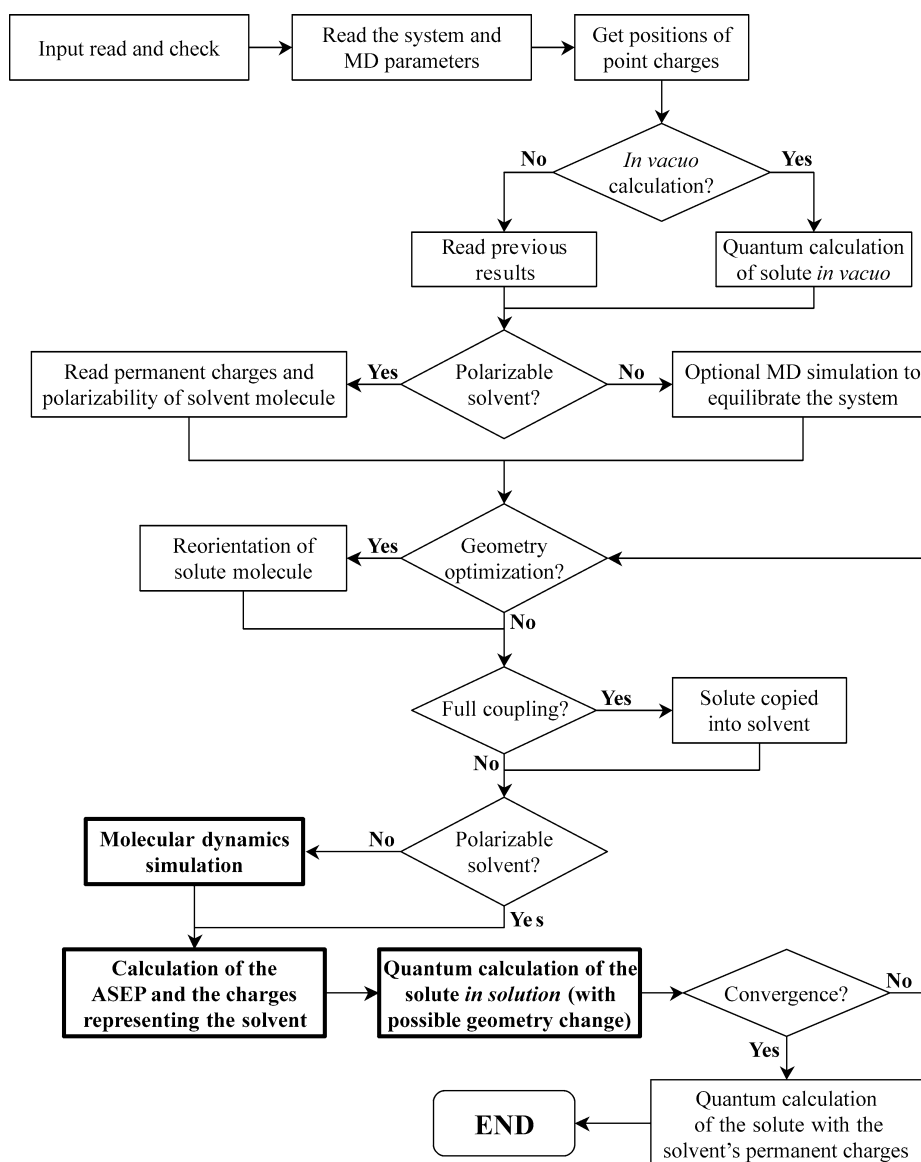


Fig. 1. Simplified flow chart of the ASEP/MD program.

### 3.1. Polarizable solvent

In the current implementation [2], the solvent polarization calculation is performed after an equilibrium sample of solvent–solute configurations has been obtained with a non-polarizable solvent. Once a normal run of the program, as described above, has finished, another calculation is carried out. This time no more molecular dynamics simulations are made, but the solvent configurations are read from the final simulation of the previous run. An additional input is given with the permanent *in vacuo* atomic charges of the solvent molecule and its polarizability tensor. Then, for each one of the configurations, all the solvent molecules are polarized by the solute charge distribution and by the other solvent molecules. The calculated ASEP now includes contributions from the permanent charges and the induced dipole moments of the solvent molecules. This potential is approximated by the two sets of point charges explained above and introduced into the quantum calculation. The charge distribution of the solute is used to obtain the solvent polarization again, and this process is repeated until convergence.

This whole calculation can be performed for different solute electronic states, keeping always the same set of solvent configurations. In this way, it is possible to study the solvent effect on electronic transitions in the solute, which take place without appreciable reordering of the solvent nuclei, but with a major response of the solvent electrons.

### 3.2. Geometry optimization

The geometry optimization method implemented in this program is based on the Free Energy Gradient (FEG) method [8], in which the effective forces on the nuclei are taken as the average force over a set of equilibrium solvent–solute configurations. Once the molecular dynamics simulation has been made and the ASEP obtained, a geometry optimization calculation is performed using the rational function optimization (RFO) method [9]. In this way, a new solute geometry is obtained, which is introduced into the next molecular dynamics simulation. The process is repeated until convergence.

The gradient and Hessian, needed for the optimization process, are given by the quantum calculation program, but the appropriate van der Waals components have to be added. These components are calculated by the ASEP/MD program as an average over all the solute–solvent configurations considered.

There are a number of options affecting the optimization method used. In each cycle of the ASEP/MD process, the optimization may be complete (a stationary point is found and this geometry is used in the next cycle) or partial (a fixed number of optimization steps is performed, and while the geometry used in the next cycle is not a stationary point, it is closer to one). Other options are related to the algorithm of the optimization procedure.

## 4. Program organization

The ASEP/MD source code is divided into different files, each of them containing subroutines and definitions grouped according to the subject. The program uses modules to make variables and arrays globally available. The following is a short description of the contents of the different files:

### *Main.f90*

This file contains the main program, global definitions, and subroutines related directly to the main program.

- Module definitions.
- ASEP\_MD: This is the main program that controls the flow of the calculations.
- LeerEntrada: This subroutine reads the input provided by the user.
- LeerContinuacion: If a calculation is resumed, this subroutine reads the “continue” file.
- EscribePrin: Writes the output of the main program.

*Execute.f90*

The subroutines in this file are related to the external programs used by ASEP/MD to perform the quantum mechanics calculations and molecular dynamics simulations. Most of these subroutines are program-independent: they simply call the appropriate subroutine or process depending on the specific program used.

- *LeerSistema*: Calls another subroutine to read the system definition (solute, solvent, force field, etc.)
- *LeerAbinitio*: Calls another subroutine to read the output of a quantum calculation.
- *EjecutarAbInitio*: Runs the appropriate quantum calculation through a SYSTEM call.
- *EjecutarDinamica*: Runs the appropriate MD simulation through a SYSTEM call.
- *LeerConfig*: Extracts a solute–solvent configuration from the output of the MD program.

*MoldySubroutines.f90*

The subroutines included in this file deal with the input and output of the molecular dynamics program MOLDY.

- *LeerControlMoldy*: Reads the MOLDY control file for some necessary parameters.
- *LeerSistemaMoldy*: Reads the system definition from the MOLDY input file.
- *ModificarMoldy*: Modifies the MOLDY input files to run a new simulation.

*GaussianSubroutines.f90*

The subroutines included in this file deal with the input and output of the quantum calculation program GAUSSIAN.

- *ModificarGaussian*: Modifies the GAUSSIAN input file to perform a new quantum calculation.
- *LeerSalidaGaussian*: Reads the output of GAUSSIAN.

*HondoSubroutines.f90*

The subroutines included in this file deal with the input and output of the quantum calculation program HONDO.

- *ModificarHondo*: Modifies the HONDO input file to perform a new quantum calculation.
- *LeerSalidaHondo*: Reads the output of HONDO.

*Calculations.f90*

This file is the core of the program. It contains the subroutines that calculate the average solvent electrostatic potential and the point charges that reproduce it. It also contains some other auxiliary subroutines.

- *PosicionesCargas*: Gets the positions of the outer point charges representing the average solvent electrostatic potential.
- *CalcularCargas*: This subroutine calls other subroutines to calculate the values of the point charges that will represent the solvent in the quantum calculations.
- *PuntosPotencial*: Calculates a grid of points inside the solute in which the ASEP will be calculated.
- *CalcularPotencial*: Calculates the electrostatic potential and interaction energies for a given solute–solvent configuration.
- *ReducirCapa*: Reduces the number of point charges in the solvent representation.

- `CalcularCampo`: Calculates the fluctuations of the solvent electric field.
- `AjusteLagrange`: Fits the outer point charges by the Lagrange multipliers method.
- `ResolverLU`: Solves a system of linear equations by the LU factorization method.
- `GirarMolecula`: Rotates a molecule (solute or solvent) to orient it along its main inertia axes.

### *Optim.f90*

This file contains all the subroutines needed to perform a geometry optimization.

- `OptimizarGeometria`: This is the main subroutine for geometry optimization. It controls the flow of the program in this section.
- `CalcularGradHessPot`: Calculates the contribution of the van der Waals potential to the gradient and Hessian.
- `BuscarMinimo`: Performs a linear search to find a minimum.
- `EnergPunto`: This is a function that returns the total energy for a given solute geometry.
- `ActualizarHessiana`: Updates the Hessian using an appropriate formula.
- `CalcularIncremento`: Calculates the increment to obtain the new solute geometry.
- `EscribirOptim`: Writes output related to the optimization process.

## 5. Usage

The compilation of this program is straightforward. Just feed the names of the different source files to the compiler. For example, with `pgf90` use (this is a single line):

```
[user@localhost]$ pgf90 Main.f90 Execute.f90 MoldySubroutines.f90  
GaussianSubroutines.f90 HondoSubroutines.f90 Calculations.f90  
Optim.f90 -o asepm
```

No makefile or configure script is needed.

Once the program has been successfully compiled and linked, the resulting binary file can be executed. It accepts a single command-line option: the name of the global input file. This file contains the values of the different options affecting the calculation. The input is formatted as a namelist called `Input`, whose syntax may depend on the compiler and platform.

In addition to the global input file, the user must have access to suitable external programs for the quantum calculations and for the molecular dynamics simulations. For these external programs the user must supply template files from which ASEP/MD can build the appropriate input files as needed. Currently, there are some limitations on the options and units that can be used in the template files, but they can be easily circumvented by modifying the launching scripts or the parser subroutines. Future versions may eliminate these limitations.

As the calculation progresses, input and output files for the external programs will be created for each cycle of the process. All files will be created in the current working directory, so it is recommended to create a dedicated directory for each calculation and run the program from there.

### 5.1. Input

The following is a short description of the different options in the `Input` namelist. For each option, its type (string, integer, or float) is indicated:

- `DirDynamics`: String. Must point to the directory where the molecular dynamics binaries reside, without a trailing slash.
- `ProgAbInitio`: String. Location of a script or binary file that will run the quantum calculation program. It will be launched as `'ProgAbInitio input-file output-file'`, so the user is advised to write a script that runs the program in this way. The value of this option should include the full path.
- `ProgDynamics`: String. Location of a script or executable that will run the molecular dynamics program. It will be launched as `'ProgDynamics input-file output-file'`, so the user is advised to write a script that runs the program in this way. The value of this option should include the full path.
- `AbInitio`: String. This option specifies the program that will be used for quantum calculations. Valid values are `'GAUSS'` for GAUSSIAN and `'HONDO'` for HONDO.
- `Dynamics`: String. This option specifies the program that will be used for molecular dynamics simulations. Since the only molecular dynamics program supported in this version is `MOLDY`, the value must be `'MOLDY'`.
- `MainOutput`: String. Name of the file where the general output will be written in text format.
- `ContFile`: String. Name of the file where the resume information will be written to and/or read from.
- `AbInitioFile`: String. Name of the template file for the quantum calculation input. This file will be read and used as a template for each quantum calculation performed. The system definition (solute and external point charges) will be added by ASEP/MD as needed, so only such options as basis set and quantum method have to be specified here. This file should be in the format expected by the quantum calculation program used. It will not be overwritten.
- `AbInitioOutput`: String. This is the generic name for the quantum calculation output files. The name will be modified for each cycle by appending `'cic.#'`, where `#` is the cycle number.
- `DynamicsFile`: String. Name of the template file for the molecular dynamics input. This file will be read and used as a template for each molecular dynamics simulation performed. This file should be in the format expected by the molecular dynamics program used, it will not be overwritten.
- `DynamicsOutput`: String. This is the generic name for the molecular dynamics output files. The name will be modified for each cycle by appending `'cic.#'`, where `#` is the cycle number.
- `InitialDynamics`: String. Sometimes the user may want to perform an initial molecular dynamics calculation with different conditions (smaller time step, lower temperature, etc.) to equilibrate the system. In that case the name of the input file for this simulation is specified here. If this option is left blank, no such equilibration dynamics will be performed.
- `EnergyConv`: Float. Value of the convergence criterium for the energy. If the solute energy difference between two consecutive cycles is less than `EnergyConv`, the energy is considered as converged.
- `ChargesConv`: Float. Value of the convergence criterion for the solute atomic charges. If the maximum difference in the atomic charges of the solute between two consecutive cycles is less than `ChargesConv`, the solute atomic charges are considered to have converged.
- `MaxDiffCharges`: Float. Maximum allowed variation in the solute atomic charges from cycle to cycle. The charges' variation is damped if it is larger than `MaxDiffCharges`.
- `ChargesType`: String. Type of atomic charges calculated for the solute. Allowed values are: `'MUL'` for Mulliken type, `'CHE'` for CHELP calculated by the quantum program, `'POT'` for charges fitted to the electrostatic potential by the ASEP/MD program.
- `MaxIter`: Integer. Maximum number of molecular dynamics + quantum calculation cycles. If this number is reached, the program terminates.
- `MaxIterPol`: Integer. Maximum number of solvent polarization cycles if a solvent polarization calculation is requested.
- `Start`: Integer. Initial cycle for the ASEP/MD calculation. If `Start = 0`, the program will start with a quantum calculation of the isolated solute. For other values, the file specified in `ContFile` should contain valid data for the previous cycles.

- **Coupling**: String. Type of coupling between solute and solvent. Allowed values are: 'NON' for no coupling, a single quantum calculation for the solute in solution will be performed; 'PAR' for partial coupling, several cycles are made until convergence is achieved or `MaxIter` is reached; 'TOT' for full coupling, the solute is copied into the solvent (this is only for pure liquids).
- **Polarization**: String. 'YES' if a solvent polarization calculation should be made; 'NO' otherwise.
- **DoDynamics**: String. The ASEP/MD cycles start with a molecular dynamics simulation, then a quantum calculation, and then another cycle begins. If a previous run is being resumed and the next molecular dynamics simulation has already been done, setting `DoDynamics` = 'NO' will skip the first molecular dynamics run and start with the following quantum calculation. This only skips the *first* molecular dynamics simulation needed. The following cycles will be made normally. If all simulations are to be performed, `DoDynamics` should be 'YES'.
- **NumConfig**: Integer. Number of solute-solvent configurations considered for obtaining the ASEP.
- **NumIniConfig**: Integer. First configuration to consider in obtaining the ASEP.
- **RadiiFactor**: Float. Factor by which the van der Waals radii of the solute atoms will be multiplied when getting the grid of points to calculate the ASEP.
- **GridDiv**: Integer. Number of points in each dimension where the ASEP will be calculated. The maximum number of grid points is thus `GridDiv` [3], but it will typically be far less.
- **FirstShell**: String. If 'YES', the charges of the solvent molecules closest to the solute will be considered explicitly. If 'NO', they will not.
- **ShellRadius**: Float. Cut-off radius around the solute beyond which solvent molecules will not be considered explicitly.
- **Proximity**: Float. Minimum distance between point charges. If two charges are closer than this value, they will be added together.
- **PermCharges**: String. Name of the file with the solvent atomic charges and polarizability. This is only used when a solvent polarization calculation is performed.
- **IndDip**: String. Type of calculation for the induced dipole moments in the solvent molecules. 'NOR' triggers a one-pass calculation, 'CON' triggers a self-consistent polarization process until convergence.
- **MaxIterOpt**: Integer. Maximum number of optimization iterations within each ASEP/MD cycle. If `MaxIterOpt` = 0, optimization is turned off. If this number of iterations is reached, the optimization ends, but the ASEP/MD program continues normally.
- **GradientConv**: Float. Convergence criterion for the gradient in the optimization process. When the gradient's magnitude is less than this value, the optimization ends.
- **MinStep**: Float. Convergence criterion for the geometry variation in the optimization process. If the magnitude of the predicted change in geometry is less than this value, the optimization ends.
- **MaxStep**: Float. Maximum allowed change in the geometry between iterations of the optimization process. If the geometry change is larger, it will be damped accordingly.
- **CalcHessian**: Integer. Number of optimization iterations after which the Hessian will be calculated anew by the quantum calculation program; in other iterations it will just be updated by the appropriate formula. If `CalcHessian` = 1, the Hessian will be calculated in every iteration. If `CalcHessian` = 0, the Hessian will never be calculated; the identity matrix will be taken as the initial Hessian and updated in following iterations. If `CalcHessian` = 10, for example, the Hessian will be calculated again every 10 cycles.
- **LinearSearch**: String. 'YES' turns on the linear search algorithm to locate a minimum of the energy in the search direction. 'NO' turns it off.

The input file can be specified as an argument when running the program. If no extension is given, '.dat' will be assumed and appended to the name:



- `asepmd` will launch ASEP/MD and ask for an input file.
- `asepmd water` will launch ASEP/MD using `water.dat` as input.
- `asepmd water.input` will launch ASEP/MD using `water.input` as input.

## 5.2. Output

The output of this program, stored in `MainOutput`, will show, for each cycle of the ASEP/MD process, the values of the solute–solvent interaction energy, solvent dipole moment, and other useful energies. A short description of this output follows:

**MM ENERGIES:** Energies and properties calculated from the MD data only; the solute and the solvent are treated classically. The MM solute–solvent interaction energy is averaged over all the considered configurations and split into three components.

- `Charge(solute)-charge(solvent) interaction`: Interaction energy between the permanent charges of the solvent and the atomic charges of the solute.
- `Dipole(solvent)-charge(solvent) interaction`: Interaction energy between the induced dipoles of the solvent and the atomic charges of the remaining solvent molecules. This energy will only be non-zero if a polarizable solvent calculation has been made.
- `Van der Waals interaction`: Van der Waals interaction energy between solvent and solute, calculated by using the same potential (typically Lennard-Jones) given in the molecular dynamics input.
- `Electric field fluctuation`: Fluctuation of the electric field at the solute’s center of mass, calculated as the difference between the mean squared field and the squared mean field.
- `Dipole-electric field interaction`: Solute dipole–electric field interaction energy. This is an oversimplification of the electrostatic solute–solvent interaction energy.

**QM ENERGIES:** Solute energies given by the quantum calculation program.

- `In vacuo energy`: Vacuum energy of the solute, this will be the same for all cycles, as it is calculated just once.
- `Total energy in solution`: Total solute energy in “solution”, including the interaction energy between the solute and the solvent charges, but not including the self-energy of the external charges themselves.
- `Energy difference`: Energy difference between the two previous values.

**QM/MM ENERGIES:** Solute–solvent interaction energies where the solute is quantum and the solvent classical, as well as some energies derived from these. The solvent representation for these energies is the set of charges that was introduced into the quantum calculation.

- `Total interaction`: Total electrostatic solute–solvent interaction, calculated as the product of the solvent charges and the electrostatic potential generated by the solute over these charges.
- `Interaction with permanent charges`: Same as above, but considering only the permanent solvent charges (and not the contribution of the induced dipoles).
- `Interaction with induced charges`: Contribution of the solvent induced dipoles to the solute–solvent interaction energy, calculated as the difference between the two previous values.
- `Internal energy in solution`: Solute internal energy in solution, without the solute–solvent interaction energy.
- `Solute polarization energy`: Solute polarization or distortion energy: the difference between the internal energy in solution and the vacuum energy.

- **Electrostatic interaction energy:** Electrostatic interaction energy including the energy needed to polarize the solvent molecules.
- **Free energy variation:** Free energy variation between the previous cycle and the current one, calculated by the free energy perturbation method.

Some of this data, as well as short messages about what the program is doing, will be written to standard output, but this can be redirected to a file. For example:

```
asepmd water > water.std will launch ASEP/MD using water.dat as input and writing messages to water.std. The output file specified in MainOutput will still be generated.
```

Apart from errors in the input, the most common errors are those occurring in the external programs, so usually the first thing to do when the program stops unexpectedly is to check the output of the molecular dynamics or quantum calculation program. The last lines in `MainOutput` and in the standard output (or the file to which it has been redirected) can also help in finding the errors.

### 5.3. Example

This example shows a simple calculation for pure water. Input files are `water.dat`, `water.g` (input for GAUSSIAN), `water.ctr`, and `water.in` (input files for MOLDY). Output file, as specified by `MainOutput`, is `water.out`.

*water.dat*

&Input

```
DirDynamics = '/usr/local/MOLDY/'
ProgAbInitio = '/home/user/bin/gaussian'
ProgDynamics = '/home/user/bin/moldy'
```

```
AbInitio = 'GAUSSIAN'
Dynamics = 'MOLDY'
MainOutput = 'water.out'
ContFile = 'water.con'
AbInitioFile = 'water.g'
AbInitioOutput = 'water.ogau'
DynamicsFile = 'water.ctr'
DynamicsOutput = 'water.omol'
```

```
EnergyConv = 0.005
ChargesConv = 0.01
MaxDiffCharges = 10.0
ChargesType = 'CHELF'
MaxIter = 50
Start = 0
Coupling = 'TOTAL'
Polarization = 'NO'
DoDynamics = 'YES'
```

```
NumConfig = 500
NumIniConfig = 1
```

256

I.F. Galván et al. / Computer Physics Communications 155 (2003) 244-259

```
RadiiFactor = 0.7
GridDiv = 12
ShellRadius = 12.0
FirstShell = 'YES'
Proximity = 0.5
```

```
MaxIterOpt = 0
/
```

*water.g*

```
#P HF/6-311G
```

```
Water
  0 1
O  0.0000000 0.0000000 0.0000000
H  0.7569503 0.0000000 -0.5858822
H -0.7569503 0.0000000 -0.5858822
```

*water.ctr* (time-unit must be set to 4.8888213e-14)

```
title = Water
nsteps = 150000
step = 0.0005
sys-spec-file = water.in
lattice-start = 1
save-file = hf.save
text-mode-save = 1
density = 1.000
scale-interval = 1
const-temp = 1
temperature = 273
roll-interval = 1000
print-interval = 1000
begin-average= 50001
average-interval= 100000
begin-rdf = 50001
rdf-interval = 10
rdf-out = 100000
dump-file = water.dump
begin-dump = 50001
dump-interval = 100
dump-level = 3
ndumps = 500
time-unit = 4.8888213e-14
end
```

*water.in*

```
#Water
```

```
SOLVENT 214
1 0.0000000 0.0000000 0.0000000 16.0 -0.834 O(S)
```



258

I.F. Galván et al. / Computer Physics Communications 155 (2003) 244-259

```

-----
|                                     QM ENERGIES                                     |
| **In vacuo energy =                                                         -76.00939257 Eh           |
|   E(0) = <Y(0)|H(0)|Y(0)>                                                 -47696.61901494 Kcal/mol |
| **Total energy in solution =                                               -76.18251789 Eh           |
|   E = <Y|H(0)+V|Y>                                                         -47805.25679947 Kcal/mol |
| **Energy difference =                                                       -0.17312531 Eh           |
|   D(E) = E - E(0)                                                           -108.63778453 Kcal/mol |
-----

```

```

-----
|                                     QM/MM ENERGIES                                     |
| **Total interaction =                                                         0.20160412 Eh           |
|   U(qt,r) = SUM[qt*V(r)]                                                    -126.50850884 Kcal/mol |
| **Interaction with permanent charges =                                       -0.20160412 Eh           |
|   U(q,r) = SUM[q*V(r)]                                                      -126.50850884 Kcal/mol |
| **Interaction with induced charges =                                         0.00000000 Eh           |
|   U(p,r) = SUM[(qt-q)*V(r)]                                                 0.00000000 Kcal/mol |
| **Internal energy in solution =                                             -75.98091377 Eh           |
|   E(int) = <Y|H(0)|Y> = E - U(qt,r)                                       -47678.74829063 Kcal/mol |
| **Solute polarization energy =                                               0.02847881 Eh           |
|   E(pol)sol = E - E(0) - U(qt,r)                                           17.87072431 Kcal/mol |
| **Electrostatic interaction energy =                                         -0.20160412 Eh           |
|   U(elec) = U(qt,r) + 1/2*U(p,q) - 1/2*U(p,r) -126.50850884 Kcal/mol |
| **Free energy variation =                                                   0.00171818 Eh           |
|   D(G), D(A)                                                                1.07817397 Kcal/mol |
-----

```

Convergence reached.

The scripts /home/user/bin/gaussian and /home/user/bin/moldy are used to give the required syntax. Their contents are:

**/home/user/bin/gaussian**

(the standard g98 script modified so that the last line reads):

```
g98 < $argv[1] > $argv[2]
```

**/home/user/bin/moldy**

(runs MOLDY in parallel using mpirun)

```
#!/bin/bash
```

```
/usr/bin/mpirun -v -np 3 -nolocal -stdin $1 /usr/local/MOLDY/moldy >& $2
```

**Acknowledgements**

This research was sponsored by the Dirección General de Investigación Científica y Técnica (BQU2000-0243) and by the Consejería de Educación y Juventud de la Junta de Extremadura (Project 2PR01A010).

**References**

- [1] A. Warshel, M. Levitt, *J. Mol. Biol.* 103 (1976) 227;  
U.C. Singh, P.A. Kollman, *J. Comput. Chem.* 7 (1986) 718;  
M.J. Field, P.A. Bash, M. Karplus, *J. Comput. Chem.* 11 (1990) 700;  
V. Luzhkov, A. Warshel, *J. Comput. Chem.* 13 (1992) 199;  
J. Gao, in: K.B. Lipkowitz, D.B. Boyd (Eds.), *Reviews in Computational Chemistry*, Vol. 7, VCH Publishers, New York, 1996, p. 119.
- [2] M.L. Sánchez, M.A. Aguilar, F.J. Olivares del Valle, *J. Comput. Chem.* 18 (1997) 313;  
M.L. Sánchez, M.E. Martín, M.A. Aguilar, F.J. Olivares del Valle, *Chem. Phys. Lett.* 310 (1999) 195;  
M.E. Martín, M.L. Sánchez, J. Olivares del Valle, M.A. Aguilar, *J. Chem. Phys.* 113 (2000) 6308;  
M.E. Martín, M.L. Sánchez, M.A. Aguilar, F.J. Olivares del Valle, *J. Mol. Structure (THEOCHEM)* 537 (2001) 213;  
M.L. Sánchez, M.E. Martín, M.A. Aguilar, F.J. Olivares del Valle, *J. Comput. Chem.* 21 (2000) 705;  
M.E. Martín, M.L. Sánchez, F.J. Olivares del Valle, M.A. Aguilar, *J. Chem. Phys.* 116 (2002) 1613.
- [3] M.L. Sánchez, M.E. Martín, I.F. Galván, F.J. Olivares del Valle, M.A. Aguilar, *J. Phys. Chem. B* 106 (2002) 4813.
- [4] K. Refson, *Comput. Phys. Comm.* 126 (2000) 309.
- [5] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian 98, Gaussian, Inc., Pittsburgh, PA, 1998.
- [6] M. Dupuis, A. Marquez, E.R. Davidson, in: *HONDO 2000*, based on HONDO 95.3; Quantum Chemistry Program Exchange (QCPE), Indiana University, Bloomington, 2000.
- [7] I. Fdez Galván, M.L. Sánchez, M.E. Martín, F.J. Olivares del Valle, M.A. Aguilar, *J. Chem. Phys.* 118 (2003) 225.
- [8] N. Okuyama-Yoshida, M. Nagaoka, T. Yamabe, *Internat. J. Quantum Chem.* 70 (1998) 95;  
N. Okuyama-Yoshida, K. Kataoka, M. Nagaoka, T. Yamabe, *J. Chem. Phys.* 113 (2000) 3519;  
H. Hirao, Y. Nagae, M. Nagaoka, *Chem. Phys. Lett.* 348 (2001) 350.
- [9] A. Banerjee, N. Adams, J. Simons, R. Shepard, *J. Phys. Chem.* 89 (1985) 52;  
J. Baker, *J. Comput. Chem.* 7 (1986) 385;  
X. Prat-Resina, M. García-Viloca, G. Monard, A. González-Lafont, J.M. Lluch, J.M. Bofill, J.M. Anglada, *Theoret. Chem. Accounts* 107 (2002) 147.

*Regular article***An averaged solvent electrostatic potential from molecular dynamics study of the anomeric equilibrium of D-xylose in aqueous solution**

I. Fdez. Galván, F.J. Olivares del Valle, M.E. Martín, M.A. Aguilar

Dpto Química-Física, Univ. de Extremadura, Avda de Elvas s/n, 06071, Badajoz, Spain

Received: 20 February 2003 / Accepted: 30 April 2003 / Published online: 5 December 2003  
© Springer-Verlag 2003

**Abstract.** We applied the free-energy perturbation method together with the averaged solvent electrostatic potential from molecular dynamics (ASEP/MD) method to study the anomeric equilibrium of D-xylose in aqueous solution. The level of calculation, 6-311G++(2d,2p) basis set and density functional theory, permits one to explain the main characteristics of the anomeric equilibrium of D-xylopyranose: in vacuo, the anomeric effect predominates and the  $\alpha$  form is the stabler. In water, solvation leads to the  $\beta$  form being the stabler. A comparison between the performances of the ASEP/MD and polarizable continuum models is also presented.

**Keywords:** Solvent effects – QM/MM methods – ASEP/MD – Xylopyranose – Anomeric effect

**Introduction**

Effective Hamiltonian methods are now widely used in the study of molecules and processes in the liquid phase [1]. This kind of method focuses on a part of the system, the solute, which is described quantum mechanically, while the rest of the system, the solvent, is described in a simplified way. In this way, it is possible to get practically the same level of calculation and accuracy for molecules in solution as has already been achieved for molecules in vacuo. The effect of the solvent enters as an effective perturbation in the solute molecular Hamiltonian—hence its name. The different versions of the theory come from the different levels of description of the solvent. In the simplest models the solvent is described as a dielectric continuum [1, 2, 3, 4]. Pioneer

works in this field were due to the groups of Tapia [4], Rivail [3], and Tomasi [1]. In general, continuum models take the solute to be inside a spherical or an ellipsoidal cavity, and a multipole expansion truncated at a certain order is employed in the description of the solute–solvent interaction. In the continuum version developed by the Tomasi group [1], known as the polarizable continuum model (PCM), the cavity is adapted to the molecular shape of the solute and the solvent perturbation is represented through a set of apparent charges placed on the cavity surface. In this way, it is not necessary to truncate the interaction potential. Continuum models have been successfully applied to the description of solvent effects on spectra [5], reactions [6], or molecular properties [7]. Their main limitation is that they neglect the microscopic structure of the solvent around the solute molecule. Because of this, many semicontinuum versions [1] have been developed where a small number of solvent molecules are explicitly included in the quantum region.

A major step forward in solvent-effect theories was represented by the use of quantum mechanics (QM)/molecular mechanics (MM) methods [8]. Here, the solvent is described by MM. When combined with simulation techniques [9] these methods permit one to obtain an adequate description of both the solute charge distribution and the solvent structure around it.

Most QM/MM methods combine a molecular dynamics (MD) simulation step with a quantum calculation, and hence it is necessary to perform as many quantum calculations as time steps in the MD simulation. An alternative strategy is to introduce the mean field (MF) approximation [10]. In this case, one alternates complete MD simulations and quantum calculations for the solute molecule in the presence of the average perturbation generated by the solvent. This approximation neglects the correlation between the solute polarization and the solvent structure around it, i.e., the solute charge distribution responds to the average perturbation, not to the instantaneous solvent configura-

Contribution to the Jacopo Tomasi Honorary Issue

Correspondence to: M. A. Aguilar  
e-mail: maguilar@unex.es

ration. Previous studies [10] have shown that this correlation energy, known as the Stark component, represents less than 2% of the solute-solvent interaction energy and that its effect on the dipole moment is less than 1%. The main advantage associated with the use of the MF approximation is that it permits one to drastically reduce the number of quantum calculations to be performed, and hence it becomes possible to describe the quantum subsystem with computationally intensive *ab initio* wavefunctions. A MF theory where the solvent structure is obtained from integral theories has also been proposed [11].

In many chemical processes, a high-level description of the quantum subsystem is compulsory. Previous studies [12] of glucopyranose, for instance, have shown the importance of including diffuse functions and correlation energy to obtain a proper description of the conformational equilibrium in monosaccharides. In *vacuo*, the conformational energy surfaces of aldohexoses and aldopentoses are determined by the anomeric effect and the formation of intramolecular hydrogen bonds. In aqueous solution, the conformational analysis is complicated by the competition between intermolecular and intramolecular hydrogen bonds. In this kind of problem, where a high-level description of the solute must be combined with a detailed description of the solvent structure around it, it is especially well suited to the combined use of QM/MM theories and the MF approximation. As an example application, we here investigate structural aspects of the anomeric equilibrium of D-xylopyranose in water.

The outline of the paper is as follows. The details of the MF theory used are described in Sect. 2, with special attention being paid to the description of the optimization geometry algorithm based on the free-energy gradient method. The influence of the solvent on the solute geometry and charge distribution and on the solvent structure are analysed in Sect. 3. Lastly, some conclusions are drawn in Sect. 4.

#### Details of the computational scheme

We used a nontraditional QM/MM method that alternates MD and QM calculations in an iterative procedure (Fig. 1). The details of the method have been described in a number of publications [13, 14]. During the MD simulation, the geometry and charge distribution of the solute and solvent molecules are considered as fixed. From the MD data one obtains the averaged solvent electrostatic potential (ASEP), which is then introduced as a perturbation into the solute molecular Hamiltonian. In order to facilitate its use with standard quantum programs, the ASEP is represented by a set of charges obtained in the following way:

1. All the charges belonging to solvent molecules that, in any of the MD configurations, lie within a sphere of radius  $R$  are included.  $R$  is chosen in such a way that it includes the first solvation shell. The value of any charge is divided by the number of solvent configurations included in the calculation of the ASEP.
2. Next, a second set of charges is obtained by a least-squares fit to the values of the electrostatic potential originating from the solvent molecules lying beyond the first solvation shell.

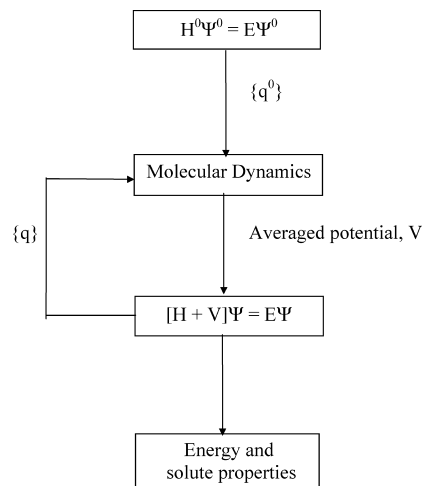


Fig. 1. Flow chart of the coupled averaged solvent electrostatic potential from molecular dynamics model

By introducing the ASEP into the solute Schrödinger equation we get a new solute charge distribution and geometry that serve as inputs for a new MD calculation. At each step the new geometry is obtained from the old one using the rational function optimization method [15]:

$$q_{k+1} = q_k + \Delta q_k, \quad (1)$$

where  $\Delta q_k$  gives the correction to the molecular geometry and is calculated by diagonalizing the augmented Hessian matrix. The force, which is the spatial derivative of the free-energy surface [16],  $G$ , is approximated as [14]

$$F(r) = -\frac{\partial G(r)}{\partial r} = -\left\langle \frac{\partial V}{\partial r} \right\rangle \approx -\frac{\partial \langle V \rangle}{\partial r}, \quad (2)$$

where  $V$  is the potential energy which is the sum of intramolecular and intermolecular contributions and the brackets denote a statistical average. As one can see, following the spirit of the MF approximation, we replace the average value of the gradient by the gradient of the average configuration. Using the same approximation, the Hessian is

$$H = \frac{\partial^2 \langle V \rangle}{\partial r \partial r}, \quad (3)$$

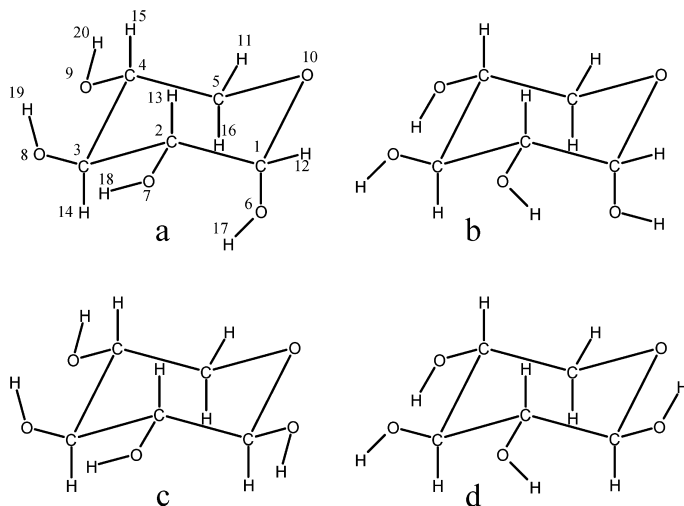
where we have furthermore neglected the terms associated with the thermal fluctuations of the gradient.

The ASEP/MD process terminates when convergence in the solute charges and in the solute energy is reached. The charges that represent the D-xylopyranose molecule during the MD were obtained from the wavefunction of the solute molecule in solution by using the CHELPG method [17].

All quantum calculations were performed with the program Gaussian 98 [18]. The 6-311G++G(2d,2p) basis set [19] was used. The use of an extended basis set is necessary to adequately describe the intramolecular hydrogen bonds. The energy and wavefunctions were calculated using density functional theory with the Becke three-parameter Lee-Yang-Parr functional [20].

The MD calculations were performed using the program MOLLY [21]. One xylose molecule in its pyranose conformation surrounded by 214 TIP3P [22] water molecules was simulated at a





**Fig. 2.** Four conformers of D-xylopyranose: **a** clockwise  $\alpha$ -xylopyranose, **b** counterclockwise  $\alpha$ -xylopyranose, **c** clockwise  $\beta$ -xylopyranose, **d** counterclockwise  $\beta$ -xylopyranose

fixed intramolecular geometry by combining Lennard-Jones interatomic interactions with electrostatic interactions in a cubic simulation box of side 18.8 Å. The pyranose-water Lennard-Jones potential parameters were taken from Ref. [23]. Periodic boundary conditions were applied, and spherical cutoffs were used to truncate the molecular interactions at 9.0 Å. A time step of 0.5 fs was used. The electrostatic interaction was calculated with the Ewald method. The temperature was fixed at 298 K by using a Nosé-Hoover [24] thermostat. Each MD calculation simulation was run for 250 000 time steps (50 000 equilibration, 200 000 production).

Having obtained the final charges for the two anomers with the ASEP/MD method, we performed free-energy simulations in a canonical ensemble ( $N, V, T$ ) to transform one of the two anomers into the other. To the free energy thus obtained, which is completely classical, one adds the difference in the internal energies of the two solutes calculated with ab initio methods:

$$\Delta E^{\text{internal}} = \langle \Psi_{\beta} | \mathbf{H}^0 | \Psi_{\beta} \rangle - \langle \Psi_{\alpha} | \mathbf{H}^0 | \Psi_{\alpha} \rangle, \quad (4)$$

where  $\Psi_{\beta}$  and  $\Psi_{\alpha}$  are the wavefunctions in solution of the two anomers considered, and  $\mathbf{H}^0$  is the gas-phase solute molecular Hamiltonian. The solution free-energy simulations were obtained using the free-energy perturbation method [25] with both single and dual topology. The coupling parameter,  $\lambda$ , was divided into 20 equally distributed intervals of 0.05 units.

## Results and discussion

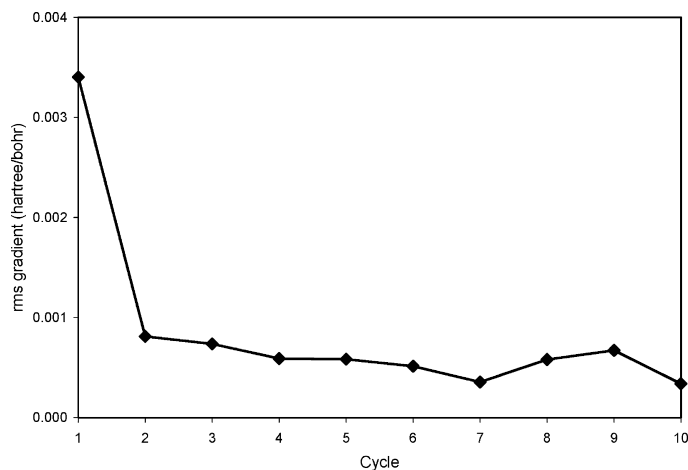
In solution, xylose cyclizes to produce five-membered (furanose form) or six-membered (pyranose form) rings, the pyranose form being the preferred form in solution. On cyclization, the achiral C1 carbon atom of an aldopentose becomes chiral. Depending on the position, axial or equatorial, of the OH group generated by cyclization,

there are two stereochemical species, anomers  $\alpha$  or  $\beta$ , for xylopyranose. Furthermore, for the isolated monomer, the hydroxyl groups prefer to orient in such a way as to yield a cooperative hydrogen bonding that is as efficient as possible. Two of these arrangements of the intramolecular hydrogen bonds are preferred: clockwise or counterclockwise. The four structures are displayed in Fig. 2. In the gas phase, Table 1, the counterclockwise orientation is preferred by 2.20 kcal/mol in the  $\alpha$  anomer and by 3.59 kcal/mol in the  $\beta$  anomer. In vacuo, the  $\alpha$  anomer is only slightly stabler than the  $\beta$  anomer (by less than 1 kcal/mol). In solution, the intramolecular hydrogen bonds can be broken and substituted by intermolecular hydrogen bonds with the water molecules, and the relative stability of the two anomers can be reversed.

We begin by analysing the changes introduced by the solvent in the geometry of the two anomers, both in their counterclockwise conformation. The root-mean-square change of the gradient on the free-energy surface during the optimization of the  $\beta$  anomer in aqueous solution obtained with 125 (25 + 100)-ps simulations is shown in Fig. 3. Convergence is reached in 5–6 cycles. From this point, the gradient begins to fluctuate around a root-mean-square gradient value of about 0.0005 hartree/bohr. This value is similar to that obtained in other compounds using the same technique, and is somewhat lower than the values obtained by other authors [16] who combined QM/MM and free-energy gradient methods but did not introduce the MF approximation.

**Table 1.** Relative stability of the different conformations of xylopyranose

	$\alpha$ -Xylopyranose		$\beta$ -Xylopyranose	
	Counterclockwise	Clockwise	Counterclockwise	Clockwise
Energy (au)	-572.854188	-572.850679	-572.852800	-572.847073
$\Delta E$ (kcal/mol)	0.0	2.20	0.87	4.46



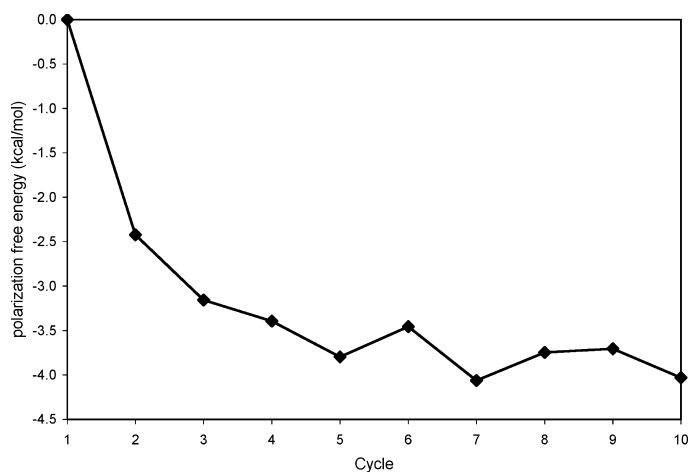
**Fig. 3.** Gradient root-mean-square (*rms*) change in free-energy surface of counterclockwise  $\beta$ -xylopyranose in aqueous solution

The evolution of the polarization free energy during the optimization procedure, i.e., the change in the solute-solvent interaction free energy when the solute charge distribution passes from its in vacuo values to the final charge distribution in solution, is shown in Fig. 4. As before, convergence is reached in 5–6 cycles, then  $\Delta G$  begins to fluctuate. The free energies were calculated by the free-energy perturbation method [25]. The final values of the energy (and of the rest of the properties) were obtained by averaging the results of the last five cycles. The behaviour of the gradient and free energy for the  $\alpha$  anomer is completely similar and is not displayed here.

In general, the variations in the geometry induced by the solvent, Table 2, are small. The C–C and C–H distances decrease by about 0.005 and 0.008 Å, respec-

tively. In contrast, the C–O hydroxyl distances increase probably owing to the participation of the oxygen atom in intermolecular hydrogen bonds. The largest difference between the two anomers appears in the C1–O anomeric distance: in the  $\beta$  anomers it does not vary with respect to the in vacuo value, but in the  $\alpha$  anomer it decreases by almost 0.01 Å. In angles, the largest variation appears in the COH angles, which increase by about 2.5°.

A summary of the thermodynamic results is given in Table 3. For comparison, we also give the results obtained with the PCM as implemented in Gaussian [18] and with a scale factor for the radius of each atomic sphere of 1.2. This model erroneously predicts that solution favours the  $\alpha$  anomers. In fact, the solvation free energy is 1.1 kcal/mol larger in the  $\alpha$  anomers than



**Fig. 4.** Polarization free-energy change of counterclockwise  $\beta$ -xylopyranose in aqueous solution during the optimization procedure

**Table 2.** Optimized geometry for the two anomers in vacuo and in solution. Distances in angstroms, angles in degrees

	$\alpha$ -D-Xylopyranose		$\beta$ -D-Xylopyranose	
	In vacuo	In solution	In vacuo	In solution
C1–C2	1.536	1.529	1.530	1.526
C2–C3	1.531	1.527	1.523	1.521
C3–C4	1.524	1.519	1.527	1.522
C4–C5	1.528	1.523	1.532	1.525
C5–O10	1.436	1.433	1.427	1.428
C1–O10	1.407	1.415	1.419	1.421
C1–O6	1.419	1.408	1.395	1.394
C2–O7	1.421	1.422	1.420	1.423
C3–O8	1.421	1.425	1.420	1.423
C4–O9	1.417	1.425	1.416	1.421
C1–H12	1.097	1.088	1.107	1.095
C2–H13	1.095	1.089	1.100	1.093
C3–H14	1.102	1.093	1.103	1.095
C4–H15	1.103	1.093	1.102	1.093
C5–H11	1.092	1.085	1.092	1.085
C5–H16	1.097	1.091	1.102	1.094
O6–H17	0.966	0.964	0.967	0.965
O7–H18	0.970	0.966	0.967	0.964
O8–H19	0.968	0.965	0.968	0.964
O9–H20	0.968	0.965	0.968	0.965
C1–O6–H17	108.2	109.5	108.2	109.7
C2–O7–H18	105.6	108.1	106.9	109.0
C3–O8–H19	106.3	108.7	106.7	109.3
C4–O9–H20	106.1	108.9	106.1	109.1

in the  $\beta$  ones. In contrast, the ASEP/MD predicts the correct trend: in solution the stabler form is the  $\beta$  anomer. Given that in vacuo the anomeric effect favours the  $\alpha$  anomer, the greater stability in solution of the  $\beta$  anomer must be due to a more favourable solvent interaction term. In the calculation of the free energy, excellent agreement between the two simulations, direct and reverse, is found. This fact shows the good precision of the simulations. The two free-energy methods used, single and dual topology, yield very similar results. Taking as a final value the average value of the direct and reverse calculations (0.5 and 0.7 kcal/mol for the dual and single topology methods, respectively) one obtains a relative stability that agrees very well with the experimental value [26], 0.4 kcal/mol.

That the solvation term favours the  $\beta$  anomer is confirmed by the results obtained from the separate ASEP/MD simulations of the two anomers, Table 4. In the two cases considered, the electrostatic component makes the largest contribution to the interaction energy, being more favourable to the  $\beta$  anomer by almost

**Table 3.** Calculated anomeric free energy differences for D-xylopyranose

Transformation direction	Polarizable continuum model	Free-energy perturbation		Experimental
		Single topology	Dual topology	
$\alpha \rightarrow \beta$		-0.9	-0.5	
$\beta \rightarrow \alpha$		-0.7	-0.7	
Mean	2.0	-0.8	-0.6	-0.4

**Table 4.** Solute–solvent interaction energy and its components (kcal/mol)

	$\alpha$ -Xylopyranose		$\beta$ -Xylopyranose	
	First cycle <sup>a</sup>	Final average <sup>b</sup>	First cycle <sup>a</sup>	Final average <sup>b</sup>
$\mu$ (D)	2.15	3.0	1.94	2.37
$E_{\text{elect}}$	-23.9	-34.1	-26.3	-37.1
$E_{\text{dist}}$	1.3	2.5	1.46	2.7
$E_{\text{LJ}}$	-11.1	-8.8	-10.8	-7.8
$E_{\text{int}}^c$	-33.7	-40.4	-35.26	-42.2

<sup>a</sup>The dipole moment in the first cycle is the in vacuo value

<sup>b</sup>The average values are calculated with the last five cycles of the averaged solvent electrostatic potential from molecular dynamics procedure

<sup>c</sup>The solute–solvent and inter action energy is calculated as the sum of the three previous contributions, where  $E_{\text{elect}}$  is the electrostatic component that includes the solute polarization,  $E_{\text{LJ}}$  is the Lennard-Jones component, and  $E_{\text{dist}}$  is the distortion energy or energy spent in polarizing the solute

3 kcal/mol. The interaction energies do not correlate with the dipole moment. In fact, the largest interaction energy corresponds to the  $\beta$  anomer, which has the lower dipole moment of the two anomers both in vacuo and in solution. The solvation of monosaccharides is clearly dominated by intermolecular hydrogen bonds, and continuum models cannot take this effect into account. Because of this, the PCM predicts greater solvation in the  $\alpha$  anomer. The rest of the components favour the  $\alpha$  form, and as a consequence the differences in the final solute–solvent interaction energy favour the  $\beta$  anomer by 1.8 kcal/mol. The  $\beta$  anomer is the stabler even if one uses the in vacuo charge distribution, i.e., if one does not permit the molecules to polarize as a consequence of the potential reaction generated by the solvent. In fact, if the in vacuo charges are used in the MD calculation then the difference of energy between the two anomers is 2.36 kcal/mol favourable to the  $\beta$  anomer.

During solution, the two anomers undergo strong polarization, which is greater in the  $\alpha$  anomer: the induced dipole moments represent 22% of the in vacuo values for the  $\beta$  anomer and almost 37% in the  $\alpha$  anomer. The CHLPG charges in vacuo and in solution are given in Table 5. In general, the two anomers have a similar charge distribution except on the C1 atom, where, in vacuo, the charge in the  $\beta$  anomer is almost twice that in the  $\alpha$  anomer. This difference is somewhat reduced in solution. There also exist appreciable differences in the charges in solution on the anomeric oxygen and on the hydroxyl oxygen bonded to C2.

The oxygen (anomeric)–oxygen (water) radial distribution functions (RDFs) for the two anomers are shown in Fig. 5. The first peak is higher in the  $\beta$  anomer than in the  $\alpha$  anomer, indicating a preferential solvation of the  $\beta$  anomer. This contrasts with what was recently found with an ab initio MD calculation of glucose [27], where the RDF of the  $\beta$  anomer was broader and lower than the RDF of the  $\alpha$  anomer. The study of the coordination numbers of the anomeric oxygen also confirms the preferential solvation of the  $\beta$  anomer—4.43 for the

**Table 5.** CHELPG charge distribution of the two anomers

	$\alpha$ -D-Xylopyranose		$\beta$ -D-Xylopyranose	
	In vacuo	In solution	In vacuo	In solution
C1	0.233	0.343	0.464	0.474
C2	0.262	0.200	0.188	0.189
C3	0.132	0.293	0.199	0.298
C4	0.348	0.293	0.324	0.321
C5	0.169	0.237	0.176	0.176
O6	-0.641	-0.687	-0.659	-0.742
O7	-0.627	-0.711	-0.679	-0.803
O8	-0.670	-0.783	-0.704	-0.805
O9	-0.709	-0.836	-0.714	-0.833
O10	-0.480	0.573	-0.517	-0.572
H11	0.069	0.066	0.076	0.093
H12	0.080	0.078	-0.016	0.029
H13	0.057	0.073	0.059	0.081
H14	0.051	0.049	0.029	0.030
H15	-0.007	0.029	0.004	0.028
H16	0.042	0.039	0.015	0.035
H17	0.442	0.480	0.429	0.489
H18	0.385	0.429	0.425	0.496
H19	0.422	0.476	0.454	0.504
H20	0.440	0.503	0.446	0.510

$\alpha$  anomer and 5.43 for the  $\beta$  anomer. The O–O RDF for the hydroxyl group of C3 is shown in Fig. 6. The differences between the two anomeric forms are minimal, with the first peak being slightly higher in the  $\alpha$  anomer. The rest of the hydroxyl groups in xylopyranose have similar RDFs. The coordination numbers in this case are now 3.46 for the two anomers. For all these reasons, it seems clear that the relative stabilization of the  $\beta$  anomers in solution is related to a better solvation of the anomeric hydroxyl group, the influence of the rest of the molecule being minimal.

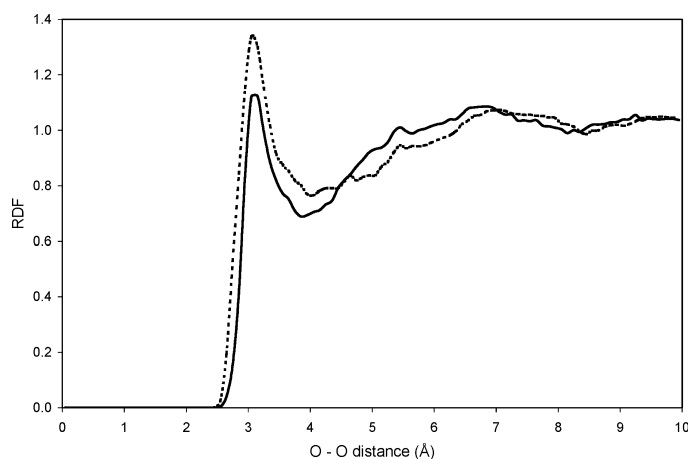
We finish by noting that there are remarkable differences between the RDF shown in Fig. 6 and those reported in Ref. [28]. Our O–O8 RDF peaks are far lower.

Our results in this case agree better with those obtained by Molteni and Parrinello [27] using ab initio MD.

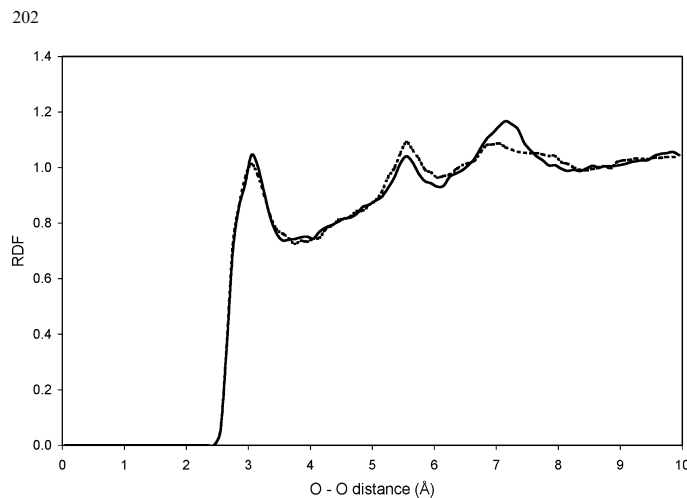
## Conclusions

We have demonstrated how, when combined with extended basis sets and correlated ab initio methods, the ASEP/MD method permits one to explain the main characteristics of the anomeric equilibrium of D-xylopyranose. In vacuo, the anomeric effect predominates and the  $\alpha$  form is the stabler. In water, solvation favours the  $\beta$  form, which becomes the stabler form. From the analysis of the O–O RDF functions and of the estimated coordination numbers it seems clear that there is a difference in the water structure around the anomeric hydroxyl group which increases hydrogen bonding in the  $\beta$  anomer. Our study predicts O–O RDFs that differ from those obtained by other authors. They also differ from a recent ab initio MD study [27] of glucose in that first peak of the oxygen (anomeric)–oxygen (water) is higher in the  $\beta$  anomer than in the  $\alpha$  anomer, indicating a preferential solvation of the  $\beta$  anomer. However, we do not find such high peaks for the O–O8 RDF as do Liu and Brady [28], and in this respect our result resembles more closely the findings of Molteni and Parrinello.

Unlike previous studies, the present ASEP/MD simulations assign different charges to the two anomers and permit the solutes to polarize in response to the reaction field generated by the solvent. While the two anomeric forms undergo strong polarization in aqueous solution, this is not the factor that explains the greater stability of the  $\beta$  form. Indeed, the  $\alpha$  form is more polarized in solution. The main factor that determines the stability of the  $\beta$  form is the difference in the in vacuo charge distribution of the two anomers, more specifically, the charges on the carbon and anomeric hydroxyls.



**Fig. 5.** Oxygen (water)–oxygen (anomeric) radial distribution function (RDF) for anomers  $\alpha$  (full line) and  $\beta$  (dashed line)



**Fig. 6.** Oxygen (water)–oxygen (bonded to C3) RDF for anomers  $\alpha$  (full line) and  $\beta$  (dashed line)

**Acknowledgements.** This research was sponsored by the Dirección General de Investigación Científica y Técnica (project no. BQU2000-0243) and by the Consejería de Educación Ciencia y Tecnología de la Junta de Extremadura (project no. 2PR01A010).

## References

- (a) Tomasi J, Bonaccorsi R, Cammi R, Olivares del Valle FJ (1991) *J Mol Struct (THEOCHEM)* 234:401; (b) Tomasi J, Persico M (1994) *Chem Rev* 94:2027
- Cramer CJ, Truhlar CJ (1995) In: Lipkowitz KB, Boyd DB (eds) *Reviews in computational chemistry*, vol VI. VCH, New York, p 1
- Rivaill JL, Rinaldi D (1995) In: Leszczynski J (ed) *Computational chemistry: review of current trends*. World Scientific, Singapore, pp
- Tapia O, Goschinski O (1975) *Mol Phys* 29:1653
- (a) Aguilar MA, Olivares del Valle FJ, Tomasi J (1993) *J Chem Phys* 98:7375; (b) Cossi M, Barone V (2000) *J Phys Chem A* 104:10614; (c) Mennucci B, Cammi R, Tomasi J (1998) *J Chem Phys* 109:2798
- (a) Bonnacorsi R, Cimiraglia R, Tomasi J, Miertus S (1983) *J Mol Struct (THEOCHEM)* 94:11; (b) Lee S, Hynes JT (1988) *J Chem Phys* 88:6953; (c) Karlström G (1988) *J Phys Chem* 92:1315; (d) Alemán C, Maseras F, Lledós A, Duran M, Bertrán J (1989) *J Phys Org Chem* 2:611
- (a) Ruiz López MF, Rinaldi D, Rivaill J-L (1986) *J Mol Struct (THEOCHEM)* 148:61; (b) Cammi R, Cossi M, Tomasi J (1996) *J Chem Phys* 104:4611; (c) Cammi R, Mennucci B, Tomasi J (1998) *J Phys Chem A* 102:870
- (a) Warshel A, Levitt M (1976) *J Mol Biol* 103:227; (b) Singh UC, Kollman PA (1986) *J Comput Chem* 7:718; (c) Field MJ, Bash PA, Karplus M (1990) *J Comput Chem* 11:700; (d) Luzhkov V, Warshel A (1992) *J Comput Chem* 13:199; (e) Gao J (1996) In: Lipkowitz KB, Boyd DB (eds) *Reviews in computational chemistry*, vol 7. VCH, New York, p 119
- (a) Allen MP, Tildesley DJ (1987) *Computer simulation of liquids*. Oxford University Press, London; (b) Frenkel D, Smit B. (2002) *Understanding molecular simulations*, 2nd edn. Academic, New York
- (a) Tapia O (1991) In: Maksic ZB (ed) *Theoretical treatment of large molecules and their interactions*, vol 4. Springer, Berlin Heidelberg New York, p 435; (b) Sánchez ML, Martín ME, Fdez Galván I, Olivares del Valle FJ, Aguilar MA (2002) *J Phys Chem B* 106:4813
- (a) Ten-no S, Hirata F, Kato S (1993) *Chem Phys Lett* 214:391; (b) Ten-no S, Hirata F, Kato S (1994) *J Chem Phys* 100:7443; (c) Kawata M, Ten-no S, Kato S, Hirata F (1995) *Chem Phys* 240:199; (d) Kawata M, Ten-no S, Kato S, Hirata F (1996) *J Phys Chem* 100:1111; (e) Kinoshita M, Okamoto Y, Hirata F (1997) *J Comput Chem* 18:1320; (f) Akiyama R, Hirata F (1998) *J Chem Phys* 108:4904; (g) Sato H, Kovalenco A, Hirata F (2000) *J Chem Phys* 112:9463
- Lii J-H, Ma B, Allinger NL (1999) *J Comput Chem* 20:1593
- (a) Sánchez ML, Aguilar MA, Olivares del Valle FJ (1997) *J Comput Chem* 18:313; (b) Martín ME, Sánchez ML, Olivares del Valle FJ, Aguilar MA (2000) *J Chem Phys* 113:6308; (c) Sánchez ML, Martín ME, Aguilar MA, Olivares del Valle FJ (2000) *J Comput Chem* 21:705; (d) Martín ME, Sánchez ML, Olivares del Valle FJ, Aguilar MA (2002) *J Chem Phys* 116:1613; (e) Sánchez ML, Martín ME, Fdez Galván I, Olivares del Valle FJ, Aguilar MA (2002), *J Phys Chem B* 106:4813
- Fdez Galván I, Sánchez ML, Martín ME, Olivares del Valle FJ, Aguilar MA (2003) *J Chem Phys* 118:255
- Prat-Resina X, García-Viloca M, Monard G, González-Lafont A, Lluh JM, Bofill JM, Anglada JM (2002) *Theor Chem Acc* 107:147
- (a) Okuyama-Yoshida N, Nagaoka M, Yamabe T (1998) *Inte J Quantum Chem* 70:95; (b) Okuyama-Yoshida N, Kataoka K, Nagaoka M, Yamabe T (2000) *J Chem Phys* 113:3519; (c) Hirao H, Nagae Y, Nagaoka M (2001) *Chem Phys Lett* 348:350
- (a) Chirlian LE, Francl MM (1987) *J Comput Chem* 8:894; (b) Breneman CM, Wiberg KB (1990) *J Comput Chem* 11:316
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (1998) *Gaussian 98*. Gaussian, Pittsburgh, PA
- (a) McLean AD, Chandler GS (1980) *J Chem Phys* 72:5639; (b) Krishnan R, Binkley JS, Seeger R, Pople JA (1980) *J Chem Phys* 72:650
- Becke AD (1993) *J Chem Phys* 98:5648

21. Refson K (1996) Moldy user's manual revision 2.10. University of Oxford, Oxford.<ftp://ftp.earth.ox.ac.uk/pub>
22. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) *J Chem Phys* 79:926
23. Sookhee N, Giammona A, Field M, Brady JW (1988) *Carbohydr Res* 180:207
24. Hoover WG (1985) *Phys Rev A* 31:1695
25. (a) Zwanzig RW (1954) *J Chem Phys* 22:1420; (b) Singh UC, Brown FK, Bash PA, Kollman PA (1987) *J Am Chem Soc* 109:1607
26. Höög C, Widmalm G (2001) *J Phys Chem B* 105:6375
27. Molteni C, Parrinello M (1998) *J Am Chem Soc* 120:2168
28. (a) Schmidt RK, Karplus M, Brady JW (1996) *J Am Chem Soc* 118:541; (b) Liu Q, Brady JW (1996) *J Am Chem Soc* 118:12276; (c) Liu Q, Brady JW (1997) *J Phys Chem B* 101:1317

## A New Method to Locate Saddle Points for Reactions in Solution by Using the Free-Energy Gradient Method and the Mean Field Approximation

I. FDEZ. GALVÁN, M. E. MARTÍN, M. A. AGUILAR

Dpto Química Física, Universidad de Extremadura, Avda. de Elvas s/n, 06071, Badajoz, Spain

Received 24 November 2003; Accepted 22 February 2004

DOI 10.1002/jcc.20048

Published online in Wiley InterScience (www.interscience.wiley.com).

**Abstract:** A new method for calculating saddle points of reactions in solution is presented. The main characteristics of the method are: (1) the solute–solvent system is described by the averaged solvent electrostatic potential/molecular dynamics method (ASEP/MD). This is a quantum mechanics/molecular mechanics method (QM/MM) that makes use of the mean field approximation (MFA) and that permits one to simultaneously optimize the electronic structure and geometry of the solute molecule and the solvent structure around it. (2) The transition state is located by the joint use of the free-energy gradient method and the mean field approximation. An application to the study of the Menshutkin reaction between  $\text{NH}_3$  and  $\text{CH}_3\text{Cl}$  in aqueous solution is discussed. The accuracy and usefulness of the proposed method is checked through comparison with other methods.

© 2004 Wiley Periodicals, Inc. J Comput Chem 25: 1227–1233, 2004

**Key words:** solvent effects; Menshutkin reaction; QM/MM; ASEP/MD

### Introduction

The determination of saddle points on free-energy surfaces for a solution reaction is an open question in current computational chemistry. In solution, any successful model must combine an accurate description of the chemical reaction and the solute–solvent interactions. In most cases a full quantum description of the solute–solvent system is impractical or simply unfeasible due to the large number of solvent molecules (several hundred) and configurations (several thousand) that it is necessary to consider to get statistically significant results. Things become even worse if one tries to determine free-energy changes along the reaction path, that is, the potential of mean force. In this case, the number of calculations increases drastically and a full quantum representation, type Car–Parrinello,<sup>1</sup> of the solute–solvent system is impossible except in the simplest cases. For these reasons, in the study of chemical reactions in solution one is compelled to introduce approximations either in the description of the solute or in the description of the solvent.

Two general methods are used in the study of solvent effects. In the first, one simplifies the solvent description, which is implicitly represented through a continuous dielectric medium.<sup>2</sup> In this case, one focuses on the solute molecule that can be described at exactly the same level as is usually employed for *in vacuo* calculations. Examples are the PCM<sup>2a,b</sup> or COSMO<sup>2c,f</sup> methods. The main criticism of these methods is that they neglect the micro-

scopic structure of the solvent around the solute. Furthermore, from a technical point of view it is not clear which are the values of the radii that one must use to construct the solute cavity. The main advantage of continuum models is their low computational cost, similar to *in vacuo* calculations.

In the second, the solvent structure is obtained from simulations. Here, we include Warshel's empirical valence bond (EVB) method<sup>3</sup> and QM/MM methods.<sup>4–6</sup> In both cases one obtains a detailed description of the microscopic structure of the solvent around the solute. EVB represents a reaction in terms of resonance structures and requires a process of parametrization for each chemical reaction. This method permits one to obtain, in an easy way, free energies of molecules and processes in solution. In traditional QM/MM methods one performs a quantum calculation for each solvent configuration generated, so that the number of quantum calculations is greatly increased in this approximation. As a consequence, most calculations to date have been performed at a semiempirical level,<sup>4</sup> although in recent years, and thanks to increasing computational capacity, several models have been proposed where the quantum mechanical calculations are performed at the DFT<sup>5</sup> or HF<sup>6</sup> level.

**Correspondence to:** M. A. Aguilar; e-mail: maguilar@unex.es

Contract/grant sponsor: Consejería de Educación y Juventud de la Junta de Extremadura; contract/grant number: ZPRO3A071

The low computational cost of continuum models comes in part from the introduction of the mean field approximation<sup>7f</sup> (MFA). In this approximation the average of the energies of the different solvent configurations is replaced by the energy of an average configuration. The MFA can also be applied to QM/MM methods. This approximation permits one to greatly reduce the number of quantum calculations, and in consequence, to increase the quality of the quantum description of the solute molecule. To date, two QM/MM methods that make use the MFA have been proposed. The first, known as RISM/SCF, was developed by the Hirata group.<sup>8</sup> Its main characteristic is that the solvent structure is obtained by using an integral theory: the reference interaction site method (RISM). The second, known as ASEP/MD,<sup>7</sup> was developed in our laboratory, and is based on the introduction of the averaged solvent electrostatic potential (ASEP) obtained from molecular dynamics (MD) simulations into the solute molecular Hamiltonian.

In this article we address the problem of the determination of transition states (TS) for reactions in solution. To locate the TS on the free-energy surface (FES) we use a variant<sup>9</sup> of the free-energy gradient method<sup>10</sup> where the derivatives of the free energy are simplified by using the mean field approximation. In this way, both the first and second derivatives of the free-energy surface can be calculated analytically, increasing the computational efficiency of the method.

The proposed method presents several advantages. First, during the ASEP/MD procedure the solute charge distribution and the solvent structure around it become mutually equilibrated. Second, one can use the same (or almost the same) level of quantum mechanical theory as for *in vacuo* calculations, that is, bond-forming and breaking process can be adequately described. Third, once one has determined the solvent structure (through molecular dynamics calculations), the free-energy derivatives are calculated analytically. This greatly increases the ability to explore the free-energy surface and to characterize stationary points. Finally, the model provides detailed information on the solvent structure and on its change during the reaction, so that it is possible to elucidate whether specific solvent molecules influence the reaction mechanism directly.

As an application, in this work we study the Menshutkin reaction (MR)  $\text{NH}_3 + \text{CH}_3\text{Cl} \rightarrow \text{CH}_3\text{NH}_3^+ + \text{Cl}^-$ . From a mechanistic point of view the MR is a special  $\text{S}_{\text{N}}2$  reaction where the reactants are neutral. Along the reaction coordinate there is a creation of two ions of opposite sign. This process is very unfavorable in the gas phase. In solution, however, the solvent significantly reduces the energy barrier and stabilizes the products. One advantage of this reaction is that it has been studied with RISM/SCF,<sup>8g</sup> QM/MM,<sup>11,12</sup> and continuum<sup>13-15</sup> models, and hence, it is possible to compare the performance of the different methods.

The rest of the article is organized as follows: In the next section the computational method is explained, and especial attention is paid to the definition of the free-energy surface and to the approximations introduced into the calculation of the first and second derivatives of the FES. The results are then discussed, and then we present the main conclusions.

## Method

In this section we describe the extension of the ASEP/MD approach to studying reactions in solution. We begin by defining the nature of the free-energy surface. We estimate the approximate standard free-energy difference between the reactant and transition state in solution as

$$\Delta G_s^0 = \Delta E + \Delta G_{\text{int}} \quad (1)$$

where  $\Delta E$  is the *ab initio* difference between the two QM models (transition state and reactant in our case) and  $\Delta G_{\text{int}}$  is the difference in the solute-solvent interaction free energy. Although formally this equation takes the same form as in the QM-FE approach of Jorgensen et al.<sup>16</sup> the meaning of the  $\Delta E$  term is different, first because the geometries of the transition state and reactant are optimized in solution, and second, because the internal energy and charge distribution of the solute are determined in the presence of the solvent (see below). Obviously, the change in the geometry and charge distribution of the solute also affects the calculation of the  $\Delta G_{\text{int}}$  term.

The saddle-point and minimum structures were located on this approximate free-energy surface,  $G_s^0$ . Once the stationary points on the approximate FES have been obtained, the total activation free-energy changes were calculated by adding the zero-point energy and the entropy and thermal contributions to  $\Delta E$ . It is important to note that these contributions are added *a posteriori* and that, in our case, the vibrational frequencies and molecular geometries necessary for the calculation of the vibrational, rotational, and translational partition functions of the solute were calculated in solution.

The different energies, geometries, vibrational frequencies, and wave functions necessary to calculate  $\Delta G_s^0$  and the zero-point and entropy corrections were obtained by using the ASEP/MD method. This is an iterative procedure that alternates molecular dynamics with quantum mechanics calculations. During the MD simulation the geometry and charge distribution of the solute molecule are considered as fixed. From the MD data we obtain the ASEP that is introduced as a perturbation into the solute molecular Hamiltonian. By solving the associated Schrödinger equation we get a new solute charge distribution and geometry that serves as input for the next MD calculation. The process terminates when convergence in the solute charges, energy, and geometry is reached. The procedure is illustrated in Figure 1.

### Calculation of $\Delta E$

The *ab initio* energy difference between the two QM structures is defined as

$$\Delta E = E_B - E_A = \langle \Psi_B | \hat{H}_B^0 | \Psi_B \rangle - \langle \Psi_A | \hat{H}_A^0 | \Psi_A \rangle \quad (2)$$

Here,  $\hat{H}_X^0$  is the *in vacuo* Hamiltonian of the structure  $X$ , and  $\Psi_X$  is the wave function of the structure  $X$  calculated in the presence of the perturbation due to the solvent.  $\Psi_X$  is obtained by solving the effective Schrödinger equation:



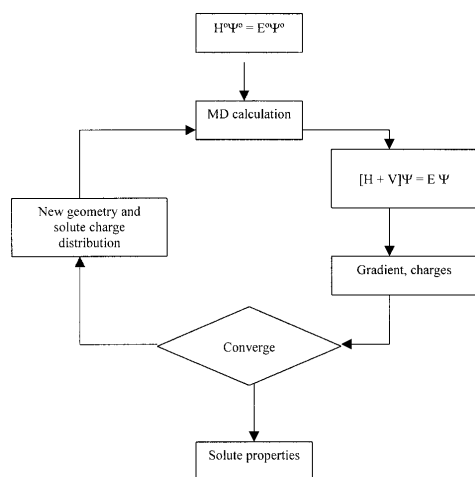


Figure 1. ASEP/MD scheme.

$$(\hat{H}_x^0 + \hat{H}_{QM/MM})|\Psi\rangle = E|\Psi\rangle \quad (3)$$

The interaction term,  $\hat{H}_{QM/MM}$ , takes the following form:

$$\hat{H}_{QM/MM} = \hat{H}_{QM/MM}^{\text{elect}} + \hat{H}_{QM/MM}^{\text{vdw}} \quad (4)$$

$$\hat{H}_{QM/MM}^{\text{elect}} = \int dr \cdot \hat{\rho} \cdot \langle \hat{V}_s(r; \rho) \rangle \quad (5)$$

where  $\hat{\rho}$  is the solute charge density and the brackets denote a statistical average. The term  $\langle \hat{V}_s(r) \rangle$  is the averaged electrostatic potential generated by the solvent at the position  $r$ , and is obtained from MD calculations where the solute molecule is represented by a set of atomic charges and geometry that are kept fixed during the simulation. The term  $\hat{H}_{QM/MM}^{\text{vdw}}$  is the Hamiltonian for the van der Waals interaction, in general represented by a Lennard-Jones potential. Given that the solvent structure, and hence, the ASEP, is a function of the solute charge density, these two quantities have to be determined iteratively. At each step of the ASEP/MD cycle, the solute charges used in the MD calculation were obtained by fitting the molecular electrostatic potential (MEP) of the solute molecule in the presence of the solvent perturbation. The CHELPG program<sup>17</sup> was used.

In our study, the  $\text{NH}_3$  and  $\text{CH}_3\text{Cl}$  molecules were described using quantum mechanical methods. The rest of the system consisting of the solvent molecules was described through classical mechanics. The basis set used in the description of the quantum part was the aug-cc-pVDZ.<sup>18</sup> The level of calculation was DFT. Following the article of Truong et al.<sup>13</sup> we used a combination of the hybrid Becke's half-and-half functional for exchange and Lee-Yang-Parr functional for correlation (BH&HLYP). From a comparison of several methods those authors conclude that this func-

tional gives the best overall performance. All QM calculations were performed with the Gaussian98<sup>19</sup> suite of programs.

#### Calculation of $\Delta G_{\text{int}}$

This magnitude is necessary to calculate the activation free energy but not in the determination of the gradient and Hessians that are calculated directly as the derivatives of the potential energy (see below). In fact,  $\Delta G_{\text{int}}$  is calculated only after the in solution structures of the TS and reactants have been determined.

The free-energy perturbation method was used to determine the free-energy change from the TS to the reactants. The solute geometry was assumed to be rigid and a function of the perturbation parameter ( $\lambda$ ) while the solvent was allowed to move freely.

The contribution of the fluctuations of the QM subsystem to the total free energy, the vibrational entropy of the solute, was calculated with the harmonic contribution. When  $\lambda = 0$ , the solute geometry and charges and the solute-solvent Lennard-Jones parameters correspond to the reactants. When  $\lambda = 1$ , the charges, Lennard-Jones parameters, and geometry are those of the transition state. For intermediate values a linear interpolation is applied. The free-energy difference between the states at  $\lambda$  and  $\lambda + \Delta\lambda$  calculated through free-energy perturbation (FEP) theory is<sup>20,21</sup>

$$\Delta G_\lambda = -RT \ln \left\langle \exp \left( - \frac{\hat{H}_{QM/MM}(\lambda + \Delta\lambda) - \hat{H}_{QM/MM}(\lambda)}{RT} \right) \right\rangle_\lambda \quad (6)$$

where  $R$  is the gas constant,  $T$  is the absolute temperature and  $\langle \rangle_\lambda$  denotes the ensemble average at state  $\lambda$ . The total free-energy change between reactants and transition state is thus

$$\Delta G_{\text{int}} = \sum_{\lambda=0}^{\lambda=1} \Delta G_\lambda \quad (7)$$

A value of  $\Delta\lambda = 0.025$  was used. That means that a total of 40 separate molecular dynamics simulations were carried out to determine the free energy. To test the convergence of the calculation, the difference in interaction free energies calculated forward—from the minimum to the saddle point—and backward—from the saddle point to the minimum—are compared. For all the results reported below, the backward and forward activation free energies agree to within less than 5%.

In eq. (6), the solute-solvent interaction energy,  $\hat{H}_{QM/MM}$ , is calculated classically. This point needs clarification. In the determination of the energies, geometries, and charge distribution of reactants and transition state in solution the solute is quantum mechanically represented. However, once one has determined these magnitudes, the calculation of  $\Delta G_{\text{int}}$  is performed through molecular dynamics simulations where the solute is represented by a set of point charges. In principle, no improvement is expected from replacing the classical by the quantum representation. This is first because the solute charges used in the MD calculation were obtained by fitting the MEP, and hence, they reproduce the QM/MM solute-solvent electrostatic energy, and second, because the free energy is a state function, and hence, its value depends only on the initial and final states (which are quantum mechani-

cally determined) and not on the particular set of charges used in the intermediate steps. For instance, the ASEP/MD method could have been used to obtain the solute charges appropriate for the intermediate geometries. These charges would be different from the ones actually used in the simulations, which are interpolated between the initial and final states. However, as was indicated above, the final energy must be independent of the particular set of charges used in the intermediate steps.

The MD calculations were performed using the program MOLLY.<sup>22</sup> In each case, 214 TIP3P<sup>23</sup> water molecules were simulated at fixed intramolecular geometry. The solute–water potential parameters for the reactants were taken from Carlson et al.<sup>24</sup> (CH<sub>3</sub>Cl) and Ferrario et al.<sup>25</sup> (NH<sub>3</sub>), and the parameters for the TS were taken from Gao and Xia.<sup>11</sup> Periodic boundary conditions, an appropriate cutoff (9.0 Å), and a cubic simulation box of 18.6 Å side were assumed. A time step of 0.5 fs was used. The electrostatic interaction was calculated with the Ewald method. The temperature was fixed at 298 K by using a Nosé–Hoover<sup>26</sup> thermostat. Each MD calculation simulation was run for 150,000 time steps (50,000 equilibration, 100,000 production).

#### Quantum Mechanical Determination of Minimum and Saddle Points

In this section we describe the calculation of the gradient and Hessian on the FES. The force on the free-energy surface (FES) is defined as:<sup>10</sup>

$$F(r) = -\frac{\partial G(r)}{\partial r} = -\left\langle \frac{\partial V(r)}{\partial r} \right\rangle \quad (8)$$

where  $G(r)$  is the free-energy,  $V$  is the sum of the contributions associated with the interaction between the atoms of the solute molecule,  $V_s$ , and with the solute–solvent interaction energy,  $V_{s'}$ , and the brackets denote a statistical average.

The Hessian is:

$$H = \left\langle \frac{\partial^2 V}{\partial r \partial r} \right\rangle - \beta \left\langle \frac{\partial V}{\partial r} \frac{\partial V^T}{\partial r} \right\rangle + \beta \left\langle \frac{\partial V}{\partial r} \right\rangle \left\langle \frac{\partial V}{\partial r} \right\rangle^T \quad (9)$$

$$H = \left\langle \frac{\partial^2 V}{\partial r \partial r} \right\rangle - \beta [\langle F^2 \rangle - \beta \langle F \rangle^2] \quad (10)$$

where the superscript  $T$  denotes the transpose and  $\beta = 1/RT$ . The last term in eq. (10) is related to the thermal fluctuation of the force.

At this point we introduce two approximations. First, and following the spirit of the mean field approximation used in ASEP/MD, we replace the average value of the force by the force of the mean configuration:

$$F(r) = -\frac{\partial \langle V \rangle}{\partial r} \quad (11)$$

$$H = \frac{\partial^2 \langle V \rangle}{\partial r \partial r} - \beta \left\langle \frac{\partial V}{\partial r} \frac{\partial V^T}{\partial r} \right\rangle + \beta \frac{\partial \langle V \rangle}{\partial r} \frac{\partial \langle V \rangle^T}{\partial r} \quad (12)$$

Second, we neglect the force fluctuations. With this approximation the Hessian is

$$H = \frac{\partial^2 \langle V \rangle}{\partial r \partial r} \quad (13)$$

Because the stationary points are defined as points where the gradient vanishes, this last approximation has no effect on the geometry of these points. The Hessian is used only to accelerate the search procedure.

The effect of these two approximations has been analyzed in a previous article.<sup>9</sup> We showed that the use of the MFA introduces only small errors into the dipole moment, energies, and gradients. Furthermore, a detailed analysis of the fluctuation term contribution showed that the errors introduced in the trace of the Hessian in the formamide–water systems when we neglected the fluctuation term was less than 5%. A similar error is introduced into the computed frequencies for methanol in the liquid state.

Once the gradient and Hessian are available, the positions of the minimum and saddle point on the FES are determined by the RFO<sup>27</sup> algorithm. It must be stressed that in our method, the solvent is in equilibrium with the charge distribution of the solute at each step of the optimization. As a consequence, nonequilibrium contributions to the activation free energy are completely neglected and, if necessary, must be included *a posteriori*.

## Results and Discussion

To examine the performance of the proposed method we consider the Menshutkin NH<sub>3</sub> + CH<sub>3</sub>Cl reaction in water. The system was partitioned into a QM subsystem (the NH<sub>3</sub> and CH<sub>3</sub>Cl molecules) and an MM system (214 water molecules). The iterative optimization ASEP/MD procedure based on the use of the free-energy gradient method was applied to obtain the optimized structure for the separated reactants and transition state in solution. Then, the activation free-energy change between reactants and TS was determined. In principle, our method yields a true stationary point on the free-energy surface in solution. The computed force constants confirm this.

The Menshutkin reaction has already been studied by different workers using different methods. In the two QM/MM studies<sup>11,12</sup> that have been performed, the solute was represented at the AM1 level. Gao and Xia<sup>11</sup> construct a two-dimensional surface, and the transition state was located on this reduced surface. Hirao et al.<sup>12</sup> locate the TS on a multidimensional free-energy surface obtained by using a variant of the free-energy gradient method proposed by Okuyama–Yoshida et al.<sup>10a</sup> The studies performed with the continuum model, PCM<sup>14,15</sup> and GCOSMO,<sup>13</sup> used different levels of *ab initio* calculations, HF, MCSCF, MP2, MP4, DFT. Truong et al.<sup>13</sup> conclude that, taking as reference the MP4 value, DFT is computationally cheaper and yields more accurate results than MP2 and MCSCF. Consequently, it is our choice in this study. The results of these authors at the DFT level are similar to those obtained by QM/MM methods. Finally, Naka et al.<sup>8b</sup> estimate the free-energy profile by the RISM-MP2 method at the Hartree–Fock optimized geometries along a distinguished reaction coordinate.

**Table 1.** Geometries, Dipole Moments, Atomic Charges, and Solute–Solvent Interaction Energies of the Reactants and Products of the Menshutkin Reaction in the Gas Phase and in Aqueous Solution.

	<i>In vacuo</i>	In solution
CH <sub>3</sub> Cl		
$E_{\text{int}}$	—	−9.36 kcal/mol
C—H	1.086 Å	1.084 Å
C—Cl	1.791 Å	1.798 Å
H—C—Cl	108.1°	107.7°
$q_{\text{C}}$	−0.162 <i>e</i>	−0.106 <i>e</i>
$q_{\text{H}}$	0.115 <i>e</i>	0.118 <i>e</i>
$q_{\text{Cl}}$	−0.183 <i>e</i>	−0.247 <i>e</i>
$\mu$	2.07 D	2.66 D
NH <sub>3</sub>		
$E_{\text{int}}$	—	−15.47 kcal/mol
N—H	1.009 Å	1.012 Å
H—N—H	107.3°	106.3°
$q_{\text{N}}$	−0.898 <i>e</i>	−1.185 <i>e</i>
$q_{\text{H}}$	0.299 <i>e</i>	0.395 <i>e</i>
$\mu$	1.54 D	2.22 D
(NH <sub>3</sub> CH <sub>3</sub> Cl) <sup>‡</sup>		
$E_{\text{int}}$	—	−51.01 kcal/mol
C—H	1.075 Å	1.073 Å
C—N	1.837 Å	2.186 Å
C—Cl	2.443 Å	2.183 Å
N—H	1.010 Å	1.014 Å
Cl—C—H	82.3°	93.8°
C—N—H	110.6°	111.0°
$q_{\text{C}}$	−0.022 <i>e</i>	−0.100 <i>e</i>
$q_{\text{N}}$	−0.361 <i>e</i>	−0.704 <i>e</i>
$q_{\text{Cl}}$	−0.710 <i>e</i>	−0.686 <i>e</i>
$q_{\text{H(C)}}$	0.126 <i>e</i>	0.198 <i>e</i>
$q_{\text{H(N)}}$	0.239 <i>e</i>	0.298 <i>e</i>
$\mu$	12.48 D	11.09 D

Table 1 gives the geometrical parameters, dipole moments, atomic charges, and solute–solvent interaction energies of the reactants and TS in gas phase and in solution. The calculation level used slightly overestimates by +0.18 D the experimental<sup>28</sup> *in vacuo* dipole moment, 1.87 D, of CH<sub>3</sub>Cl. The NH<sub>3</sub> charge distribution is described better. This point is important because in a previous article we showed the importance that a correct description of the solute charge distribution has on the prediction of thermodynamical and structural properties of solutions. Our results slightly improve those obtained by Truong et al.,<sup>13</sup> who use the same functional but a basis set of lower quality.

As shown in Table 1, solvation has an appreciable effect on the polar reactant molecules. The dipole moments of CH<sub>3</sub>Cl and NH<sub>3</sub> increase by 28 and 44%, respectively. In CH<sub>3</sub>Cl, the charge on the hydrogen atoms hardly changes, and the change in the dipole moment is associated to the polarization of the C—Cl bond. For CH<sub>3</sub>Cl, the C—Cl distance increases from 1.791 to 1.798 Å when one passes from vacuum to solution. The solution values are given as the average value of the last five ASEP/MD cycles. The standard deviations are minimal, less than 0.001 Å, 0.1°, and 0.05 D. This indicates that the size of the simulations (75 ps) is adequate: the fluctuations associated with the limited size of the simulation

have a negligible effect on the geometric parameters. The Cl—C—H angle compresses slightly from 108.1° to 107.7°. Similar effects are found for NH<sub>3</sub>; the N—H bond elongates by 0.003 Å and the H—N—H angle is compressed by about 1°.

The solvent effects on the transition state structure are larger: the C—Cl distance decreases from 2.443 to 2.183 Å when one passes from gas phase to solution, and the N—C distance increases from 1.837 to 2.186 Å. In solution, the TS structure is more symmetric, and as a consequence, the dipole moment is lower. The CHELPG charge for the leaving group (Cl) decreases from −0.71 to −0.69 when one passes from vacuum to aqueous solution. Simultaneously, the charge on the N atom increases from −0.36 to −0.70. Given that the charge separation increases along the reaction coordinate, these results indicate that the TS obtained in solution is earlier than the TS in the gas phase, which is in agreement with the Hammond postulate.

As for the free energy of activation, our *in vacuo* result, 44.9 kcal/mol at room temperature, compares very well with the Truong et al.<sup>13</sup> result, 45.7 kcal/mol, as expected from the similar level of calculation. These results are also in good agreement with the AM1 result of Gao and Xia,<sup>11</sup> 46.7 kcal/mol. Solvent effects decrease appreciably the activation free energy. In solution our model yields a value of 25.64 kcal/mol. This value lies between the GCOSMO<sup>13</sup> (24.8 kcal/mol) and AM1/MM values (26.3 kcal/mol) if a two-dimensional mapping<sup>11</sup> is used to locate the TS and 26.8–27.1 when the FEG is used<sup>12</sup>). Amovilli et al.<sup>14</sup> using PCM and a 6-311G\*\* basis set obtain 16.8 kcal/mol at the Hartree–Fock level and 20.5 kcal/mol at the CASSCF level. A similar value, 20.9 kcal/mol, is obtained by Naka et al.<sup>8e</sup> with their RISM-MP2 calculation. Experimental data are not available for this system. There is an experimental<sup>29</sup> result (23.5 kcal/mol) for a similar reaction where the CH<sub>3</sub>Cl is replaced by CH<sub>3</sub>I. The differences obtained between the different continuum models, GCOSMO and PCM, are due in part to the level of calculation, but also to the different choice of the cavity radius. In GCOSMO, the atomic radii are adjusted to reproduce free energies of hydration for a representative set of small molecules and ions. In PCM, they are calculated, in general, as 1.2 times the van der Waals radii. Given the great difference that exists between the different versions of the continuum model, a comparison with QM/MM values is complicated. In general, continuum models yield lower activation free energies than QM/MM methods. However, the differences are small and, as has been established in previous articles, we can conclude that in this kind of reaction a nonspecific electrostatic solvent–solute interaction makes the greatest contribution to the TS stabilization.

A direct comparison between our result and other QM/MM results is also complicated by the different level of calculation used (DFT in our case and AM1 in refs. 11 and 12). Furthermore, in the calculation of the activation free energy we include the entropy corrections, which are neglected in the AM1 calculations. However Gao and Xia<sup>11</sup> showed that the AM1 results without entropy corrections provide a reasonable approximation to the *ab initio* free-energy profile obtained at the MP4SDTQ/6-31+G(d) level with scaled frequencies. The same level of calculation (AM1) is used by Hirao et al.<sup>12</sup> If we take the AM1 results as valid, then the agreement between them and our results permits us to conclude that the use of the mean field approximation in the calculation of

**Table 2.** Activation Free Energy (In kcal/mol) and Its Components ( $\Delta V^0$  Stands for the Zero-Point Energy and Thermal Contributions; for the Meaning of the Rest of the Terms, See the Text), and Variation of the Solute–Solvent Interaction Energy from Transition State to Reactants,  $\Delta E_{\text{int}}$ , and Its Electrostatic,  $\Delta E^{\text{elec}}$ , and Lennard–Jones,  $\Delta E^{\text{L-J}}$ , Components.

	<i>In vacuo</i>	In solution
$\Delta E$	32.70	27.48
$\Delta V^0$	12.23	11.86
$\Delta G_{\text{int}}$	—	−13.70
$\Delta G_s^0$ (total)	44.93	25.64
$\Delta E^{\text{elec}}$	—	−30.50
$\Delta E^{\text{L-J}}$	—	4.32
$\Delta E_{\text{int}}$	—	−26.18

gradient and Hessian does not introduce significant errors into the determination of transition states.

Table 2 displays the different contributions to the activation free energy. As one can see, the most important contribution to the decrease of the activation energy when one passes from gas phase to solution comes from the differential solvation free energy. In aqueous solution the TS is better solvated than the reactants. This table also gives the differential solute–solvent interaction energy and its components. As can be seen, the solvent stabilization (−26.18 kcal/mol) is mainly due to the electrostatic (permanent + induced charges) component. In fact, the Lennard–Jones contribution is positive. It stabilizes the reactants in comparison to the TS. The interaction energy does not include the energy spent in polarizing the solvent (in this case, because nonpolarizable molecules are employed for the solvent, it is its structure that is polarized). In classical electrostatics this distortion energy can be obtained as  $-1/2$  of the interaction energy. If the two contributions are added together one obtains the real stabilization produced by the solvent, −13.09 kcal/mol, which agrees very well with the result obtained with the much more computationally demanding free-energy perturbation method, −13.70 kcal/mol.

The evolution of the solvation along the reaction coordinate can be characterized by the changes in the radial distribution functions (rdf), which are shown in Figure 2. The most remarkable fact is the displacement in the position of the first peak of the Cl–O and N–O rdfs from reactants to TS. Simultaneously, the height of the first peak increases, indicating a stronger interaction of the TS with the solvent. Our rdfs compare very well to those obtained by Hiraó et al., confirming the validity of the mean field approximation.

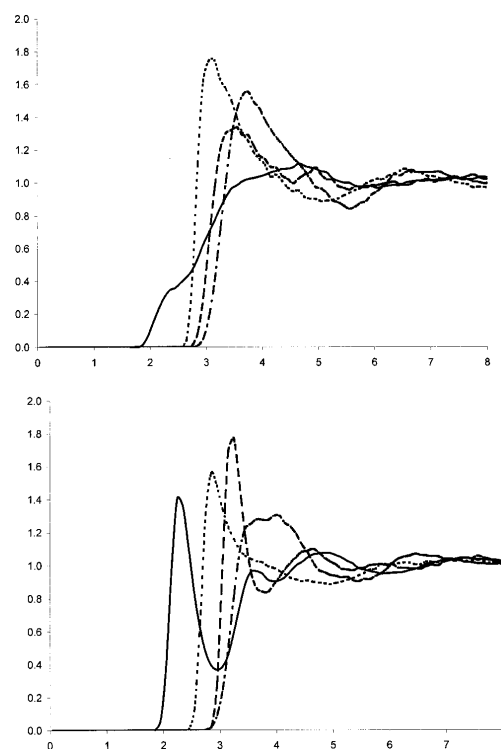
## Summary

A new method for determining saddle points in solution has been presented. It is based on the use of the mean field approximation, and permits one to sample the QM/MM potential energy surface properly and to obtain the reaction paths in the solvent environment and the free-energy changes associated with the reactions. The main characteristics of the method are: (1) high-level *ab initio*

quantum calculations are used in the description of the solute molecule; (2) the information about the solvent structure is obtained from MD simulations; (3) the free-energy derivatives are calculated analytically; (4) true stationary points are found, which avoids the definition of a distinguished reaction coordinate. The use of the MFA permits one to greatly reduce the number of quantum calculations and facilitates the calculation of the derivatives. These two points greatly increase the ability of the method to explore the free-energy surface and to find stationary points.

As application of the new method we carried out a study of the Menshutkin reaction  $\text{NH}_3 + \text{CH}_3\text{Cl} \rightarrow \text{CH}_3\text{NH}_3^+ + \text{Cl}^-$  in aqueous solution. In agreement with previous studies we found that the activation energy is greatly reduced by the solvent. In solution, the TS structure is more symmetric than in the gas phase. In agreement with the Hammond postulate, the TS obtained in solution is earlier than the TS in the gas phase.

It was shown, through comparison with the results provided by other methods, that the MFA does not introduce significant errors



**Figure 2.** Radial distribution functions for reactants (top) and transition state (bottom) of the Menshutkin reaction. Four functions are represented: Cl–H (continuous line), Cl–O (dashed line), N–O (dotted line), and C–O (dash-dotted line).

into either the position of the saddle point or the height of the barrier. We can conclude that the joint use of the ASEP/MD method in the determination of the electronic structure of the solute and the free-energy gradient method and MFA in the determination of the free-energy derivatives constitutes a valid and efficient method for the study of reactions in solution.

### Acknowledgments

The authors acknowledge Dr. José Carlos Corchado for fruitful discussions and help in some calculations.

### References

- Car, R.; Parrinello, M. *Phys Rev Lett* 1985, 55, 2471.
- (a) Tomasi, J.; Bonaccorsi, R.; Cammi, R.; Olivares del Valle, F. J. *J Mol Struct (Theochem)* 1991, 234, 401; (b) Tomasi, J.; Persico, M. *Chem Rev* 1994, 94, 2027; (c) Rivail, J. L.; Rinaldi, D. In *Computational Chemistry: Review of Current Trends*; Leszczynski, J., Ed.; World Scientific Publishing: Singapore, 1995; (d) Cramer, C. J.; Truhlar, D. G. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH Publishers: New York, 1995, p. 1, Vol. VI; (e) Klamt, A.; Schüürmann, J. *Chem Soc Perkin Trans II* 1993, 799; (f) Truong, T. N.; Stefanovich, E. V. *Chem Phys Lett* 1995, 240, 253.
- (a) Warshel, A.; Weiss, R. *J Am Chem Soc* 1980, 102, 6218; (b) Hwang, J.-K.; King, G.; Creighton, S.; Warshel, A. *J Am Chem Soc* 1988, 110, 5297; (c) Åqvist, J.; Warshel, A. *Chem Rev* 1993, 93, 2523.
- (a) Warshel, A.; Levitt, M. *J Mol Biol* 1976, 103, 227; (b) Field, M. J.; Bash, P. A.; Karplus, M. *J Comp Chem* 1990, 11, 700; (c) Luzhkov, V.; Warshel, A. *J Comp Chem* 1992, 13, 199; (d) Gao, J. *J Phys Chem* 1992, 96, 537; (e) Vasilyev, V. V.; Bliznyuk, A. A.; Voityuk, A. A. *Int J Quantum Chem* 1992, 44, 897; (f) Théry, V.; Rinaldi, D.; Rivail, J.-L.; Maigret, B.; Ferenczy, G. G. *J Comp Chem* 1994, 15, 269; (g) Thompson, M. A.; Glendenning, E. D.; Feller, D. *J Phys Chem* 1994, 98, 10465.
- (a) Wei, D.; Salahub, D. R. *Chem Phys Lett* 1994, 224, 291; (b) Tuñón, I.; Martins-Costa, M. T. C.; Millot, C.; Ruiz-López, M. F.; Rivail, J.-L. *J Comp Chem* 1996, 17, 19; (c) Wesolowski, T. A.; Warshel, A. *J Phys Chem* 1993, 97, 8050; *J Phys Chem* 1994, 98, 5183; (d) Wesolowski, T. A.; Muller, R. P.; Warshel, A. *J Phys Chem* 1996, 100, 15444.
- (a) Vaidehi, N.; Wesolowski, T. A.; Warshel, A. *J Chem Phys* 1992, 97, 4264; (b) Stanton, R. V.; Little, L. R.; Merz, K. M. *J Phys Chem* 1995, 99, 17344; (c) Moriarty, N. W.; Karlström, G. *J Phys Chem* 1996, 100, 17791; (d) Tu, Y.; Laaksonen, A. *J Chem Phys* 1999, 111, 7519.
- (a) Sánchez, M. L.; Aguilar, M. A.; Olivares del Valle, F. J. *J Comp Chem* 1997, 18, 313; (b) Sánchez, M. L.; Martín, M. E.; Aguilar, M. A.; Olivares del Valle, F. J. *Chem Phys Lett* 1999, 310, 1995; (c) Sánchez, M. L.; Martín, M. E.; Aguilar, M. A.; Olivares del Valle, F. J. *J Comp Chem* 2000, 21, 705; (d) Martín, M. E.; Sánchez, M. L.; Olivares del Valle, F. J.; Aguilar, M. A. *J Chem Phys* 2000, 113, 6308; (e) Muñoz Losa, A.; Fdez. Galván, I.; Martín, M. E.; Aguilar, M. A. *J Phys Chem B* 2003, 107, 5043; (f) Sánchez, M. L.; Martín, M. E.; Fdez. Galván, I.; Olivares del Valle, F. J.; Aguilar, M. A. *J Phys Chem B* 2002, 106, 4813.
- (a) Ten-no, S.; Hirata, F.; Kato, S. *Chem Phys Lett* 1993, 214, 391; (b) Ten-no, S.; Hirata, F.; Kato, S. *J Chem Phys* 1994, 100, 7443; (c) Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *Chem Phys* 1995, 240, 199; (d) Kawata, M.; Ten-no, S.; Kato, S.; Hirata, F. *J Phys Chem* 1996, 100, 1111; (e) Kinoshita, M.; Okamoto, Y.; Hirata, F. *J Comp Chem* 1997, 18, 1320; (f) Akiyama, R.; Hirata, F. *J Chem Phys* 1998, 108, 4904; (g) Sato, H.; Kovalenko, A.; Hirata, F. *J Chem Phys* 2000, 112, 9463; (h) Naka, K.; Sato, H.; Morita, A.; Hirata, F.; Kato, S. *Theor Chem Acc* 1999, 102, 165.
- Fdez. Galván, I.; Sánchez, M. L.; Martín, M. E.; Olivares del Valle, F. J.; Aguilar, M. A. *J Chem Phys* 2003, 118, 255.
- (a) Okuyama-Yoshida, N.; Nagaoka, M.; Yamabe, T. *Int J Quantum Chem* 1998, 70, 95; (b) Okuyama-Yoshida, N.; Kataoka, K.; Nagaoka, M.; Yamabe, T. *J Chem Phys* 2000, 113, 3519.
- Gao, J.; Xia, X. *J Am Chem Soc* 1993, 115, 9667.
- Hirao, H.; Nagae, Y.; Nagaoka, M. *Chem Phys Lett* 2001, 348, 350.
- Truong, T. N.; Truong, T.-T. T.; Stefanovich, E. V. *J Chem Phys* 1997, 107, 1881.
- Amovilli, C. A.; Mennucci, B.; Floris, F. M. *J Phys Chem B* 1998, 102, 3023.
- Fradera, X.; Amat, L.; Torrent, M.; Mestres, J.; Constans, P.; Besalú, E.; Martí, J.; Simon, S.; Lobato, M.; Oliva, J. M.; Luis, J. M.; Andrés, J. L.; Solà, M.; Carbó, R.; Duran, M. *J Mol Struct (Theochem)* 1996, 371, 171.
- (a) Chandrasekhar, J.; Smith, S. F.; Jorgensen, W. L. *J Am Chem Soc* 1985, 107, 154; (b) Chandrasekhar, J.; Jorgensen, W. L. *J Am Chem Soc* 1985, 107, 2974; (c) Jorgensen, W. L. *Acc Chem Res* 1989, 22, 184.
- (a) Chirlan, L. E.; Francl, M. M. *J Comp Chem* 1987, 8, 894; (b) Breneman, C. M.; Wiberg, K. B. *J Comp Chem* 1990, 11, 316.
- (a) Dunning, T. H., Jr. *J Chem Phys* 1989, 90, 1007; (b) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J Chem Phys* 1992, 96, 6796; (c) Woon, D. E.; Dunning, T. H., Jr. *J Chem Phys* 1993, 98, 1358.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
- Kollman, P. A. *Chem Rev* 1993, 93, 2395.
- Mark, A. E. In *Encyclopedia of Computational Chemistry*, Schleyer, P. V. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer, H. F., III; Schreiner, P. R., Eds.; Wiley and Sons: Chichester, 1998, p. 1070, Vol. 2.
- Refson, K. *Comp Phys Commun* 2000, 126, 310.
- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J Chem Phys* 1983, 79, 926.
- Carlson, H. A.; Nguyen, T. B.; Orozco, M.; Jorgensen, W. L. *J Comp Chem* 1993, 14, 1240.
- Ferrario, M.; Haughney, M.; McDonald, I. R.; Klein, M. L. *J Chem Phys* 1990, 93, 5156.
- Hoover, W. G. *Phys Rev A* 1985, 31, 1695.
- (a) Simons, J.; Jørgensen, P.; Taylor, H.; Ozmert, J. *J Phys Chem* 1983, 87, 2745; (b) Banerjee, A.; Adams, N.; Simons, J.; Shepard, R. *J Phys Chem* 1985, 89, 82; (c) Baker, J. *J Comp Chem* 1986, 7, 385.
- de Jongh, J. P.; Dijkerman, H. A. *J Mol Spectrosc* 1968, 25, 129.
- Okamoto, K.; Fukui, S.; Shingu, H. *Bull Chem Soc Jpn* 1967, 40, 1920.